

Онлайн- исследования в России: тенденции и перспективы

Москва
Online Market Intelligence

2016

Рецензенты:

д. соц. н. Г.Г. Татарова
д. полит. н. С.С. Бодрунова

Авторский коллектив:

Б. Адаму, Д. Беркс, И. Бурдейн, Н. Вееш, Т. Вейс, К.А. Гаврилов, С.Г. Давыдов,
И.Ф. Девятко, И.С. Искра, К.Р. Казарян, П. Квок, Ф. Келли, С.В. Кетов,
П.В. Колозариди, О. Кольцова, П.А. Лебедев, О.С. Логунова, А.М. Мавлетова,
Н.Г. Малошонов, И.В. Микаелян, С.П. Моисеев, К. Очоа, Д. Пулестон, М. Ревилла,
Е. Рождественская, Ю.К. Савинкова, Д. Сантус, Е.А. Терентьев, О.В. Ткачев,
А. Турбина, Ф. Тьюро, А. Уитли, С.С. Хитров, Х. Хофкирхнер, Д.А. Шурыгина.

Ответственные редакторы:

А.В. Шашкин, И.Ф. Девятко, С.Г. Давыдов

Онлайн-исследования в России: тенденции и перспективы /

Под редакцией Шашкина А.В., Девятко И.Ф., Давыдова С.Г. — Москва: Издательство МИК, 2016. — 555 с.

«Онлайн-исследования в России: тенденции и перспективы» — это четвертый сборник статей ведущих российских и зарубежных ученых и практиков, посвященных социологическим и маркетинговым исследованиям с использованием сети интернет. Первый сборник вышел в 2007 году, но многие темы, поднятые на его страницах девять лет назад, актуальны и сейчас. Цель сборника — представить российскому читателю как современные методические разработки, так и результаты проведенных за последние несколько лет исследований. В материалах представлены размышления ведущих специалистов из разных стран о том, как будут развиваться современные исследовательские методы, как изменятся подходы к поиску и отбору респондентов, как эволюционируют способы нахождения потребительских инсайтов, а также какие ключевые тенденции и точки роста существуют сейчас в исследовательской индустрии. Сборник будет интересен как профессионалам исследовательской индустрии, так и широкому кругу читателей.

Содержание

Предисловие	9
-------------------	---

Раздел 1

Методология онлайн-исследований: программные статьи

От «виртуальной лаборатории» до «социального телескопа»: метафоры тематических и методологических инноваций в онлайн-исследованиях.....	19
---	----

Девятко И.Ф.

Опыт прогнозирования развития российской онлайн-индустрии.....	35
--	----

Давыдов С.Г., Казарян К.Р., Логунова О.С.

Раздел 2

Методология онлайн-исследований: дизайн инструментария

Разработка опросов в стиле бонсай®	65
--	----

Пулестон Д.

Не всегда короче значит лучше.....	113
------------------------------------	-----

Бурдейн И.

Новое поколение онлайн-анкет?	
Использование преимуществ интерактивности интернета.....	133
<i>Ревилла М., Очоа К., Турбина А.</i>	

Как влияет приглашение на отклик в онлайн-опросах?	161
<i>Мавлетова А.М., Малошенок Н.Г., Терентьев Е.А.</i>	

Влияние знакомства на результаты рекрутинга респондентов для участия в онлайн-опросах: опыт эмпирического изучения.....	185
<i>Гаврилов К.А.</i>	

Раздел 3

Методология онлайн-исследований: построение выборок

Нужно ли оставлять выборку на волю случая?.....	203
<i>Сантус Д., Квок П., Келли Ф.</i>	

Применение выборки, направляемой респондентом, в онлайн-исследованиях: методологические вызовы и исследовательские перспективы.....	219
<i>Моисеев С.П., Савинкова Ю.К.</i>	

Раздел 4

Методология онлайн-исследований: онлайн панели

Эффект самоотбора при формировании онлайн-панели: опыт первого года лонгитюдного исследования студентов НИУ ВШЭ	237
<i>Малошенок Н.Г., Семенова Т.В., Терентьев Е.А.</i>	

Раздел 5

Методология онлайн-исследований: геймификация

Использование исследовательских игр вместо
геймифицированных опросов. Влияние метода
исследовательских игр на вовлеченность респондентов
и вероятность их будущего участия в подобных проектах.....267
Адаму Б., Беркс Д.

Предсказывая будущее: начала научного прогнозирования.....323
Пулестон Д., Уитли А., Хофкирхнер Х.

Раздел 6

Методология онлайн-исследований: качественные исследования

Качественные онлайн-методы.....391
Рождественская Е.

Проблема способа проведения качественных интервью:
скайп, онлайн, лицом к лицу.....417
Колозариди П.В.

Раздел 7

Практикум. Опросные методы и онлайн-панели

Онлайн-панель как инструмент телевизионных измерений.....433
Искра И.С., Кетов С.В.

Поиск реальной целевой аудитории своего бренда.....451
Ткачев О.В.

Использование методов онлайн-исследований для прогнозирования поведения потребителей в кризис.....	459
<i>Микаелян И.В., Хитров С.С.</i>	

Раздел 8

Практикум. Безопросные методы: Big Data, аналитика социальных сетей

Большие вопросы — большие ответы.....	471
<i>Вейс Т., Тьюро Ф., Вееш Н.</i>	

Классификация должностей: переход к парадигме Big Data.....	483
<i>Лебедев П.А., Шурыгина Д.А.</i>	

Автоматизированные методы анализа больших массивов интернет-текстов.....	505
<i>Кольцова О.</i>	

WAVE 8: контент и его ценность.

Новый язык общения производителя бренда и потребителя.....	521
<i>Бородин Д.А.</i>	

Аннотации и сведения об авторах/ Summaries and information about the authors.....	535
--	------------

Предисловие

С момента выхода первого сборника «Онлайн-исследования в России: тенденции и перспективы» прошло девять лет; сейчас перед вами уже четвертый сборник статей, посвященных социологическим и маркетинговым исследованиям, проводимым с использованием сети интернет. Когда читаешь предисловия, написанные к предыдущим сборникам, перечитываешь статьи, становится понятно, что многое из того, что было написано и опубликовано тогда, актуально и сейчас, а настоящий сборник — продолжение тем, начатых в 2007 году. Проследить развитие дискуссий с самого начала легко — все сборники можно найти и бесплатно скачать в сети.

Многие из нас размышляют о том, как будут развиваться современные методы исследований, как изменятся подходы к поиску и отбору респондентов, как эволюционируют способы нахождения потребительских инсайтов. Мы ездим на конференции по всему миру, общаемся с коллегами, не перестаем искать точки роста исследовательской индустрии вообще и нашего бизнеса в частности. Какие же ключевые тенденции существуют сейчас в области онлайн-исследований? Я бы выделил три такие тенденции:

1. Снижение интереса людей к опросам и желания участвовать в них. Интернет предлагает человеку множество способов провести свободное время — от общения с друзьями до шоппинга. Онлайн-опросы также вступают в конкуренцию за это время. Несмотря на то, что панельные проекты имеют не только социальную значимость (позволяют участникам влиять на мир вокруг себя), но и дают пользователям возможность заработать, их популярность

снижается. Ситуация усугубляется еще и тем, что сами заказчики исследований, а также исследовательские организации, сопротивляются изменениям инструментария, векторами которых могли бы быть уменьшение длительности интервью и интерактивизация/геймификация процесса опроса. Последствия развития этой тенденции можно проследить на примере развитого американского рынка: приглашенные на сцену участники американских панелей признались на конференции CASRO Digital в США в 2015 году, что участвуют во многих панелях и получают более сотни приглашений к опросам в неделю, заполняя анкеты одну за другой. Такая ситуация снижает доверие к данным, собираемым с использованием панелей, заставляет заказчиков экспериментировать с другими источниками выборок для онлайн-опросов (соцсети, поточные выборки — с баннеров сразу в анкету и т.п.), а исследователей — размышлять о возможностях разбивать анкеты на части, сокращая их длительность, и работать над инструментами, способствующими удержанию внимания респондентов.

2. Значимое увеличение распространенности и времени пользования мобильными устройствами, в частности — переход к массовому использованию мобильных устройств для доступа к сети интернет. Уже сейчас более 20% респондентов, приглашенных к онлайн-опросам в России, пытаются заполнить анкету с мобильного телефона, при этом доля анкет, адаптированных для такого заполнения, не превышает 5%. Пока вы читаете этот текст, сотни исследователей по всему миру размышляют о том, как адаптировать анкеты для мобильных устройств: что делать с табличными вопросами и шкалами, как найти оптимальную длительность интервью, как сохранить внимание респондента. Кроме того, важно понять, какой способ приглашения к опросу на мобильном устройстве является оптимальным, если традиционные e-mail и SMS стремительно сменяются сообщениям в мессенджерах и соцсетях. Кроме того, только мы придумали инструменты интерактивизации анкет и работы с изображениями, идеально подходящие для стандартных экранов десктопов и ноутбуков (такие, например, как drag-and-drop, click test и т.п.), как эти наработки потеряли свою значимость в связи с уменьшением размера экрана. Казалось бы, не так давно тесты ви-

зуальных концепций и рекламы переместились в онлайн, а теперь, по мере увеличения доли мобильного заполнения на небольшом экране, у них есть все шансы вернуться обратно в холл-тесты.

3. Big Data и рост возможностей для получения и использования пассивно собранных данных в принятии бизнес-решений. За последние несколько лет неопросные исследовательские методы поиска потребительских инсайтов получили широкое распространение. Это прежде всего связано с появлением источников данных, не предполагающих взаимодействия вида «вопрос-ответ»: мы имеем в виду неструктурированные источники информации, такие как социальные сети или другие технологии, которые производят большие массивы потенциально полезных текстов; новые перспективные типы данных, такие как биометрия, распознавание эмоций, eye-tracking — отслеживание движений глаз, и набирающий популярность «интернет-вещей»; новые платформы, такие как краудсорсинг и predictive markets (техника по определению потребительских предпочтений, симулирующая игру на бирже); современные этнографические техники; большие базы данных о поведении потребителей (Big Data). Если исследователи традиционно умеют работать с обычными данными, то в случае с «большими данными» преимущество получают технологические компании. Плюсы технологических компаний заключаются не только в возможности обрабатывать большие массивы информации, но и в наличие метаданных, которые можно использовать для синтеза методов, способного существенно улучшить качество принимаемых бизнесом решений, а также служить основой для более точных прогнозов и стратегий. Преимуществом исследователей в этом соперничестве является знание бизнеса клиентов, поэтому успех будет зависеть от скорости адаптации к технологическим изменениям.

Книга, которую вы держите в руках, является сборником статей ведущих российских и зарубежных ученых и практиков в области онлайн-исследований, отобранных редакторским коллективом по степени актуальности, научной и практической значимости. Цель сборника — представить российскому читателю как современные методические разработки, так и результаты проведенных за последние несколько лет исследований, необходимые для понимания ситу-

ации и векторов развития индустрии маркетинговых исследований, а также выработки стратегии поведения в новых реалиях (от индивидуального поиска и расстановки приоритетов между темами и компетенциями до корпоративного выбора возможностей для инвестирования и роста).

Мы разделили статьи на несколько смысловых блоков, чтобы читателю было удобно. Сборник начинается с двух «программных статей». Первая — статья одного из ведущих российских экспертов в области методов социологических исследований Инны Девятко (ВШЭ). Статья представляет собой аналитический обзор ключевых публикаций в социологии, психологии и когнитивной науке, позволяющий продемонстрировать эволюцию тем и методов исследований с использованием интернета. Важен фокус автора не только на возможностях, но и на проблемах, возникающих по мере роста популярности онлайн-исследований. Вторая статья написана авторским коллективом (Сергей Давыдов — ВШЭ, Ольга Логунова — ВШЭ, Карен Казарян — РАЭК) и посвящена развитию российской онлайн-индустрии. Прогнозы и сценарии, изложенные в статье, основаны на результатах исследовательского проекта «Экономика Рунета», проводимого ВШЭ совместно с РАЭК с 2012 года. Таким образом мы хотели сразу поместить читателя в контекст: помочь составить представление о развитии научных подходов, происходящем на фоне определенных социально-экономических реалий России.

Автор следующей статьи — Джон Пулестон, вице-президент по инновациям компании Lightspeed GMI, входящей в группу Kantar. Джон известен своими экспериментами в области геймификации онлайн-анкет (от способов представления до формулировок вопросов), которым способствовал доступ к многомиллионным онлайн-панелям по всему миру. Статья обязательна к прочтению для тех, кто понимает, что с современными анкетами что-то не так, и мы можем совсем лишиться респондентов, если ничего не предпримем по этому поводу.

Часто мы слышим, что сокращение длительности анкеты может решить все наши проблемы. Инна Бурдейн (NPD Group) показывает в своей статье, что с длительностью не все так однозначно. Важно, из каких блоков состоит анкета, какова сложность этих блоков,

а также каково их субъективное восприятие человеком, заполняющим анкету. Важным дополнением, расширяющим наше понимание проблематики онлайн-анкет, является коллективная статья Мелани Ревиллы (Университет Помпеу Фабра), Карлоса Очоа (Netquest) и Альберта Турбины (Netquest), фокусом которой являются различные подходы к интерактивизации анкет, проверенные путем тестирований и экспериментов. Помимо самой анкеты, на принятие решения об участии в опросе влияет приглашение к нему, а также факт знакомства респондента с отправителем. В статье исследователей из Высшей Школы Экономики Айгуль Мавлетовой, Натальи Малошенок и Евгения Терентьева рассматривается влияние различных элементов приглашения на отклик в онлайн-опросе, а в статье Кирилла Гаврилова (ВШЭ, ИС РАН) доказывается, что только личное знакомство с приглашающим, а вовсе не профессиональная солидарность, значимо влияет на отклик.

Две следующие статьи сборника посвящены построению выборок в онлайн-опросах. Раздел открывают признанные международные эксперты в области выборок Дэбра Сантус (Sample Strategies TNS Research International), Питер Квок (TNS Research International) и Франк Келли (Lightspeed GMI), которые рассуждают о методах квотирования выборок, доказывая, в частности, что квотирование на этапе рассылки приглашений, а не только квоты на завершенные интервью, даст лучшее соответствие респондентов, заполнивших анкету, целевой аудитории проекта. Статья российских авторов Станислава Моисеева (ВШЭ, «инФОМ») и Юлии Савинковой (Ипсос) рассматривает особенности выборки, направляемой респондентом (WebRDS) или, как ее иногда называют, выборки методом «снежного кома». Авторы описывают результаты ряда собственных экспериментов по созданию выборки для онлайн-опросов методом RDS, а также анализируют свежие методические наработки коллег в этой области.

Известны методические исследования, посвященные анализу различий между людьми, соглашающимися участвовать в опросах, и теми, кто в опросах не участвует. Авторы статьи, посвященной эффекту самоотбора при формировании онлайн-панели (Наталья Малошенок, Татьяна Семенова, Евгений Терентьев, ВШЭ), анализи-

руют опыт реализации лонгитюдного исследования со студентами НИУ-ВШЭ. Эффект самоотбора в работе определяется с помощью сравнения социально-демографических характеристик, успеваемости, социального и культурного капитала и других атрибутов студентов, согласившихся на участие в исследовании, с теми, кто от участия отказался.

Есть авторы, статьи и доклады которых неоднократно получали престижные награды международных исследовательских ассоциаций. К таким авторам относится Бетти Адаму (Research Through Gaming), которая совместно с Дэвидом Берксом (Бизнес-школа при Университете Винчестера) представила статью по исследовательским играм (не путать с геймификацией опросов!). Статья систематизирует опыт автора в области создания игр, дает ссылки на примеры таких игр и анализирует опыт и впечатления респондентов от участия в них. Вторая статья «орденоносного» автора Джона Пулестона (Lightspeed GMI), написанная в соавторстве с Александром Уитли (Lightspeed GMI) и Хубертусом Хофкирхнером (Prediki), посвящена прогностическим исследованиям. Частично результаты этого совместного труда презентовались в России в 2015 году на конференции ESOMAR Best of Russia; в настоящем сборнике публикуется полный текст статьи.

Одним из самых быстрорастущих направлений в онлайн-исследованиях является переход в онлайн качественных методов. В статье Елены Рождественской (ВШЭ) рассматриваются различные подходы к проведению качественных исследований онлайн, делаются выводы о возможностях и ограничениях этих подходов. Статья Полины Колозарики (ВШЭ) анализирует опыт проведения качественных интервью с помощью онлайн-инструментов, таких как Skype, текстовые сообщения и электронная почта; поднимаются вопросы, связанные с качеством данных, этикой исследователя, применимостью технологически опосредованных инструментов.

Важной частью сборника является раздел «Практикум», разбитый на два подраздела, посвященных опросным и неопросным методам (Big Data и аналитике соцсетей). Опросный блок включает статьи Ивана Искры и Сергея Кетова (ГфК) о телеизмерениях с использованием онлайн-панелей; Олега Ткачева (Wellhead) об опы-

те сегментации потребителей по итогам онлайн-опросов; Инги Микаэлян и Сергея Хитрова (РБК.research) об использовании онлайн-исследований для прогнозирования поведения потребителей в кризис. В блоке про безопросные методы представлена еще одна статья от коллектива авторов из компании GfK (Том Вейс, Франк Тюрро и Ник Вееш), посвященная возможностям использования «больших данных» для составления прогнозов, построения бизнес-моделей и отслеживания потребительского поведения; статья Павла Лебедева и Дарьи Шурыгиной (Superjob), описывающая эксперимент по работе с текстовыми данными (названиями должностей) с помощью алгоритмов, характерных для работы с Big Data; статья Олеси Кольцовой (ВШЭ), также посвященная анализу текстов и не только рассматривающая современные методы такого анализа, но и дающая обзор российских исследований, в которых использовались данные методы; статья Даниила Бородина (UM Russia), описывающая результаты глобального международного исследования WAVE о важности выбора каналов коммуникации и видов контента для построения эффективных отношений между производителями и потребителями.

Продуктивного вам чтения! Если у вас возникнут вопросы к авторам статей, вы можете связаться с ними по указанным адресам электронной почты. Вопросы и предложения инициаторам сборника можно отправлять на адрес: ashashkin@omirussia.ru.

*Александр Шашкин
(Online Market Intelligence)*

Раздел 1

Методология онлайн- исследований: программные статьи

От «виртуальной лаборатории» до «социального телескопа»: метафоры тематических и методологических инноваций в онлайн-исследованиях

*Девятко Инна Феликсовна
(Национальный исследовательский университет
«Высшая школа экономики»)*

Очевидным трендом первого десятилетия XXI века стал растущий интерес академических и прикладных исследователей к осмыслению последствий радикальной трансформации общепринятых методов социальных и поведенческих наук в результате глобальной цифровой революции, превратившей значительную часть населения Земли как предельной целевой аудитории научных изысканий в «виртуальных жителей», доступ к которым, с одной стороны, существенно упростился, а с другой, отчасти в силу упомянутого упрощения, стал подвержен новым рискам снижения качества получаемых данных — от «опросной усталости» до «профессионализации» участия в таких реактивных (предполагающих осознание факта участия в исследовании) процедурах, как опрос или эксперимент. Этот растущий интерес привёл к возникновению практически необозримого количества методических экспериментов и метааналитических исследований, посвящённых сравнению внешней и внутренней валидности данных, получаемых офлайн и онлайн, а также возможным подходам к повышению качества опросного инструментария и дизайна исследований с учётом специфических угроз валидности и надёжности измерения, а также угроз валидности статистического вывода, возникающих в онлайн-исследованиях.

Описанные тенденции привели к предсказуемому смещению фокуса многих аналитических обзоров в сторону описания и систематизации возникающих методологических новаций — например, последствий появления социальных сетей для традиционного сетевого анализа и для планирования и проведения масштабных онлайн-

экспериментов, или предсказуемых и непредсказуемых последствий «революции больших данных» для возможностей накопления и анализа предположительно нереактивных «цифровых отпечатков». Однако вполне оправданное пристальное внимание к методологической стороне дела неизбежно ведёт к пренебрежению теми постоянными изменениями, которые происходят в самой исследуемой интернет-реальности как своеобразной «движущейся цели» [1] и способствуют существенным сдвигам в тематическом репертуаре, а иногда и к переопределению дисциплинарных границ социальных исследований.

Таким образом, основная цель данного обзора — не столько систематизация и каталогизация инструментария онлайн-исследований (подобная предпринятой нами ранее [2]), сколько попытка продемонстрировать произошедшие с начала 2000-х гг. изменения в тематическом репертуаре онлайн-исследований в социологии и «родственных» дисциплинах (прежде всего — в психологии и когнитивной науке), в типах доступных данных, а также в понимании природы интернета как среды для академических и прикладных исследований. Разумеется, достижение этой цели потребует хотя бы краткого рассмотрения возникающих в ответ на новые технологические и теоретические вызовы методологических инноваций и сопряжённых с ними исследовательских возможностей и ограничений. Однако для того чтобы систематический анализ вновь возникающих технических решений не оттеснил на второй план содержательные задачи, мы попытаемся оттолкнуться в данном обзоре не от вышеупомянутой классификации методов онлайн-исследований [2], тесно связанной с устоявшимися типологиями методологических подходов, существующих в социологии, а от более широкой классификации областей интернет-исследований в психологии, исходно предложенной Л. Скитка и Э. Саргисом в 2006 г. [3; см. также: 1]. Указанные авторы использовали в проведённом ими метаанализе публикаций в журналах Американской психологической ассоциации за 2003–2004 гг., так или иначе опиравшихся на данные, полученные онлайн (преимущественно это были данные экспериментов), следующее разграничение видов интернет-исследований:

(а) *транслирующие* — исследования, в которых осуществляется адаптация устоявшихся методов к изучению существующих «офлайн» исследовательских тематик в интернет-среде (например, моди-

фикация и применение какой-то из версий включённого наблюдения для изучения онлайн-групп политических активистов, использование традиционных личностных опросников для масштабной проверки предсказаний тех или иных теорий на онлайн-аудитории или перенос в сеть лабораторных экспериментов по оценке скорости реакции или вынесению разными индивидами суждений относительно стилистической правильности тех или иных конструкций, что позволяет судить о различиях в социолектах);

(б) *феноменологические*— исследования, в которых основной интерес для изучения представляют новые феномены и процессы, возникающие в интернет-среде и ставшие возможными именно в результате тех технологических изменений, которые обеспечили создание инфраструктуры для этой среды (например, исследования сетевого краудсорсинга или процессов распространения слухов в сетевых сообществах);

(в) *новаторские*— исследования, использующие уникальные методологические возможности, открывающиеся лишь в интернет-среде (например, исследование точности обыденных оценок личностных качеств, выносимых на основании информации из личных аккаунтов пользователей сетей, исследование размерности субъективного пространства музыкальных предпочтений по данным изучения индивидуальных музыкальных библиотек на сайтах, предоставляющих доступ к музыкальным записям и т.п.).

Описанное аналитическое различие, конечно, весьма условно, поскольку новые методологические возможности тесно связаны и с новой «феноменологией», и с превращением инфраструктуры интернета в своего рода «виртуальную лабораторию» [4], где традиционные методы социальных наук могут быть реализованы (и существенно трансформированы) не только с меньшими издержками и с привлечением малодоступных ранее социальных групп вне зависимости от институциональных и географических границ, но и с использованием принципиально новых возможностей социальных сетей и открываемых доступом к разработке API-приложений новых подходов к мотивированию испытуемых, получающих обратную связь о своих результатах в текущем режиме (с возможностью их размещения в социальных сетях, сравнения с результатами других

испытуемых и т.п.). Относительно недавним примером такого «синкретического» исследования может служить совокупность проектов в областях социо- и психолингвистики, реализованная как исследовательская интернет-лаборатория “Games With Words” (<http://www.gameswithwords.org>, ранее — <http://coglanglab.org/>), предлагающая потенциальным испытуемым «узнать больше о языке и о себе, содействуя передовой науке».

Кроме того, очевиден и общий институциональный контекст для всех описанных видов исследований — радикальные изменения в уровне доступности и «онлайн-наблюдаемости» сотен миллионов индивидов: по данным Международного союза телекоммуникаций, предварительная оценка доли индивидов, имеющих интернет-доступ, в 2015 г. достигла 43,4% на глобальном уровне (совокупно в развитых странах, включая Российскую Федерацию, этот показатель составит по предварительной оценке 82,2%), тогда как в 2005 году соответствующие показатели составляли 15,8 и 50,9% соответственно (см. подробнее <http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>). В результате возникла беспрецедентная возможность получения научных данных, относящихся не только к студентам университетов (традиционному пулу испытуемых для экспериментов, имеющему специфические ограничения [5]) или к дорогостоящим национальным опросным выборкам жителей западных стран, образованных, индустриализированных, богатых и демократичных (после публикации обзора [6] за такими испытуемыми и респондентами закрепились аббревиатура WEIRD), но и к практически неограниченному диапазону разнообразных возрастных, культурных, образовательных и т.п. групп¹. Иллюстрацией такой возможности может служить широко цитируемая статья [7], в которой на основе сравнительного анализа большой интернет-выборки и доступных выборок, использовавшихся при проведении социально-психологических экспериментов и личностных опросников, результаты которых были опубликованы за год

¹ Примером сетевого исследовательского проекта, потенциально адресованного миллиардам жителей Земного шара как целевой совокупности и направленного на получение сравнительных данных о когнитивных функциях, предпочтениях, поведенческих показателях для различных культурных, возрастных и т.п. групп, может служить проект «Making Science Less Weird» (<http://lessweird.org/index.html>).

в одном из психологических журналов, эмпирически обосновывался тезис о том, что в совокупности интернет-выборка была более разнородной и лучше репрезентировала население США с точки зрения пола, социально-экономического статуса, местожительства и возраста, чем упомянутые доступные выборки (на том, как в результате изменились подходы к рекрутированию испытуемых для экспериментов и к каким новым методологическим вызовам это привело, мы вкратце остановимся дальше).

С учётом сделанных оговорок мы, однако, будем ориентироваться на вышеописанную типологию в целях структурирования нашего обсуждения.

К ярким образцам *трансляционных* исследований, раньше осуществлявшихся в форме «бумажного» опроса или лабораторного эксперимента, помимо приведённых выше примеров «виртуальных лабораторий», можно отнести, в частности, совокупность личностных опросов и психометрических тестов, реализованных в Фейсбук-приложении *my Personality* (<http://mypersonality.org/wiki/doku.php>), с помощью которого к началу 2012 года (период закрытия проекта, запущенного в 2007 году как личный сайт одного из организаторов) были собраны данные о личностных чертах (прежде всего речь о популярном личностном опроснике «Большая Пятёрка» [8], см. также <http://www.outofservice.com/bigfive/>) и некоторых других психометрических и биомедицинских показателях для более чем 4000000 индивидуальных пользователей Фейсбука, которые примерно в 40% случаях также давали согласие на доступ к данным своего профиля и информации о «лайках» в этой социальной сети («бизнес-ориентированная» часть проекта предполагает возможность получить данные о расширенном профиле личности и разместить их на своей странице за \$4 либо за выполнение предлагаемых разработчиками задач). База данных этого проекта доступна другим исследователям и послужила основой для ряда публикаций. Насколько новаторскими в целом являются получаемые в данном случае результаты? Если данные о надёжности-воспроизводимости и конкурентной валидности оценок личностных качеств, основанных на самоотчётах участников и оценках их друзей, в целом близки к результатам, полученным в предыдущих онлайн-исследованиях (с той существенной оговоркой, что

сам масштаб доступной совокупности наблюдений обеспечивает несопоставимую с большинством офлайн-исследований статистическую мощность, заведомо достаточную для обнаружения даже очень слабых связей, а возможность анализа данных о нереактивных поведенческих «цифровых следах», подобных «лайкам», позволяет прямо проверять модели типа «черты личности — предпочтения — поведение»), то данные о сравнительной точности основанных на информации из индивидуальных аккаунтов (лайков в Фейсбуке) машинных прогнозов индивидуального профиля личности, полученного по результатам заполнения индивидом теста «Большой Пятёрки», и аналогичных предсказаний, сделанных другими людьми на основании личного знакомства, пусть и требующие дальнейшей проверки, представляют собой принципиально новую исследовательскую перспективу [9].

На этом примере легко увидеть и методические преимущества, и проблемы, и вызовы, с которыми сталкиваются типичные трансляционные исследования. Разумеется, осуществляемые онлайн-опросы и эксперименты очевидным образом снижают издержки на тиражирование материалов, контроль за правильностью заполнения форм, ввод данных и т.п., а также позволяют использовать обратную связь в качестве эффективного нематериального стимула для масштабного привлечения участников — и в результате спонтанного влияния размещения результатов тестирования в социальных сетях и блогах на друзей испытуемого, и в результате использования метода «выборки, управляемой респондентами». Однако даже самые успешные из таких проектов сталкиваются с проблемой весьма разнообразного и широкого, однако существенно смещённого по ряду социально-демографических признаков самоотбора. Так, в проекте my Personality приняли участие более чем 1000 испытуемых в каждой из 40 наиболее «активных» стран за пределами США и Великобритании, однако если ориентироваться на базовые характеристики выборки по состоянию на 2010 год, в ней были сверхпредставлены женщины и возрастные группы до 30 лет (http://mypersonality.org/wiki/doku.php?id=about_mypersonality).

Говоря о переносе в интернет-среду опросов и экспериментов, следует упомянуть и растущее использование краудсорсинговых

платформ, прежде всего самой известной из них — Amazon Mechanical Turk (AMT), для рекрутирования испытуемых. Все более значительная доля экспериментов и квазиэкспериментов, публикуемых в англоязычных журналах, осуществляется с использованием последней. Испытуемые получают небольшое денежное вознаграждение или подарочный сертификат за выполнение заданий, а исследователи имеют возможность вводить ограничения, связанные с квалификацией первых (в том числе умениями, качеством исполнения или мотивацией участников, которые могут проверяться в предварительных тестовых заданиях). В целом выборки на основе AMT существенно смещены в сторону сверхпредставленности женщин, более молодых возрастных групп и, в случае США, белого населения, однако, как уже говорилось, они обладают явными преимуществами с точки зрения географического и культурного разнообразия [1; 7]. Вместе с тем существуют данные в пользу несколько более высокой доли ответов в онлайн-опросах в сравнении с офлайн-опросами, хотя эти различия зависят и от динамики уровня вознаграждения [10;11]. Однако высказываются и небезосновательные опасения относительно того, что оперирование уровнем вознаграждения может вести лишь к росту количественных показателей деятельности респондентов и испытуемых, выполняющих задания «быстрее и больше», но не обязательно более качественно. Насколько эквивалентны данные офлайн- и онлайн-экспериментов и факторных опросов с точки зрения получаемых результатов и качества измерения? Существуют подтверждения эквивалентности для ряда психологических экспериментов и факторных опросов — в частности, подтверждения эквивалентности результатов для популярных среди социологов факторных опросов, использующих экспериментальный план и технику виньеток — [10; 12; 13], однако различия в способе реализации (онлайн или офлайн) могут взаимодействовать с выявляемыми время от времени различиями между конкретными методиками. Некоторые из методик (например, эксперименты с виньетками и вынесением нормативных и фактических суждений о социальном мире), видимо, гарантируют высокий уровень «трансляционной» эквивалентности измерения, тогда как для личностных шкал эквивалентность результатов не так очевидна и должна обосновываться в каждом конкретном случае [14].

Существенным ограничением всех онлайн-экспериментов остаётся невозможность контролировать условия выполнения испытуемыми задачи, в частности, потенциального переключения между участием в эксперименте и другими видами деятельности, использования других источников информации и т.п. [4].

Широко обсуждаемой методической проблемой экспериментов и опросов на краудсорсинговых платформах остаётся и множественное заполнение, т.е. непропорциональное участие в онлайн-экспериментах и опросах «профессиональных» и далеко не наивных испытуемых и респондентов, которое не всегда может быть предотвращено с помощью предварительного скрининга и введения критериев участия, поскольку участники активно делятся сведениями о последних на многочисленных форумах. В частности, для АМТ было показано, что для почти полутора сотен исследований 11% заданий были выполнены 1% наиболее активных «турок» [15]. Иногда высказывается предположение, что наиболее активные участники опросов и экспериментов на краудсорсинговых платформах могут составить устойчивую основу для создания панелей, применяемых в лонгитюдных исследованиях, однако и оно нуждается в эмпирической проверке.

Среди наиболее популярных тематик *феноменологических* исследований, сфокусированных на принципиально новых (или не столь новых, но претерпевших существенную трансформацию в цифровую эпоху) феноменов и процессов, возникающих в интернет-среде, чаще всего выделяют исследования интернет-зависимости [1; 16; 17], кибер-травли (cyberbullying) [18], распространения сетевых мемов [19], сплетен и слухов (например, [20; 21]).

Последняя из перечисленных широких тематик — культурной диффузии и подражания, т.е. контактного распространения информации и убеждений — вполне традиционная для социологии, культурной антропологии и психологии, в последние два десятилетия буквально обрела «второе дыхание» в силу того, что интернет — это уникальный источник масштабных, имеющих временную, а часто и географическую разметку нереактивных данных, позволяющих проверять весьма сложные модели распространения влияния и передачи информации без необходимости обращаться к микроуровневым

данным, основанным на индивидуальных самоотчётах о поведении или на включённом наблюдении множества взаимодействий. Иными словами, интернет — это своего рода «социальный телескоп» [4], открывающий для наблюдения обширные области социальной реальности, но допускающий при этом и масштабирование, превращающее его скорее в «социальный микроскоп», пригодный для анализа индивидуальных единиц наблюдения и элементарных взаимодействий. Дополнительную, прикладную актуальность этому тематическому направлению придало активное использование данных квази-эпидемиологических исследований распространения информации с помощью социальных сетей и цифровых медиа в публичных дискуссиях о предполагаемой роли последних в инициировании и информационной поддержке социальных движений, которые в некоторых (но далеко не во всех) случаях привели к политическим изменениям на Ближнем Востоке, в Восточной Европе, Латинской Америке — см., например, [22] (однако в [23] представлены и серьёзные критические аргументы против переоценки роли социальных медиа в этих изменениях).

Выше обосновывался относительный характер различия между тремя типами исследований и, с учётом этой оговорки, к *новаторским* исследованиям можно отнести масштабные эксперименты с использованием инфраструктуры социальных сетей (например, масштабные эксперименты в Фейсбуке или Твиттере), новые подходы к анализу огромных массивов неактивных текстовых данных (тематическое моделирование и т.п.), и новации, связанные с использованием больших массивов поведенческих данных, представленных в форме «цифровых следов», в том числе получаемых с помощью сенсоров смартфонов и планшетов (например, геолокационных данных). Мы лишь кратко охарактеризуем эти новые подходы, останавливаясь, прежде всего, на новых, многообещающих методологических возможностях, которые здесь возникают, и вызовах, которые эти данные ставят перед традиционными науками о поведении.

Как справедливо отмечают С. Голдер и М. Мейси, Всемирная сеть позволила преодолеть некоторые ограничения, присущие более ранним попыткам проанализировать эффекты сетевой структуры на социетальном уровне на основе выборочных обследований эгоцентрических сетей, которые реализовывались по модели массовых опросов и не по-

зволяли ни учесть такие существенные атрибуты этой сетевой структуры, как связность, наличие кластеров или центральность, ни собрать параллельно данные об объективных индикаторах текущего социального взаимодействия [4, р. 131]. Ставшие возможными в результате растущей доступности интернета истинные рандомизированные эксперименты с манипулированием уровнем кластеризации или гомофилии в сетях индивидуальных пользователей, когда сам исследователь создаёт некое интернет-сообщество и случайно приписывает уровни названных атрибутов различным экспериментальным группам, позволяют в значительной мере решить старую проблему разделения эффектов общего окружения и гомофилии (которые ведут к одновременному восприятию некоторого воздействия в силу межличностного сходства и общности «фона»), с одной стороны, и, с другой стороны, собственно контактного «заражения» той или иной информацией или культурным образцом. В известном исследовании Д. Сентолы [24], например, с помощью таких «искусственно структурированных сообществ», которые имели, по воле экспериментатора, характеристики случайного или кластеризованного графа (т.е. онлайн-сообщества различались исключительно топологической структурой социальной сети), было показано, что контактное распространение поведенческих образцов (в частности, решения зарегистрироваться на некоем форуме, посвящённом вопросам здоровья) оказывается, вопреки ранее преобладавшей точке зрения, более эффективным в сетях с большей степенью кластеризации и большим диаметром, обеспечивающих повторное воздействие информации на индивида. Однако важно отметить, что существенным ограничением такого подхода, как это часто случается и с обычными лабораторными экспериментами, оказывается низкая экологическая валидность: в эксперименте Сентолы существенную роль играла невозможность для участников взаимодействовать друг с другом напрямую, вне заданных условиями эксперимента коммуникативной инфраструктуры, которая помогла избежать смешивающего эффекта спонтанных межличностных коммуникаций, однако именно такое естественное взаимодействие с многогранными идентичностями окружающих зачастую играет решающую роль в ситуациях, когда реальные люди принимают решения, связанные с вопросами их здоровья.

Несмотря на некоторые ограничения описанного нового подхода к экспериментальному сетевому анализу, он явно открывает широкое поле для изучения макросоциальных эффектов за пределами привычных опросных данных и экспериментов в масштабе первичной социальной группы. В частности, можно предположить, что интересные перспективы в области построения более сложных моделей глобальных эффектов сетевой структуры создаст применение этого подхода к изучению объяснительных возможностей формальных моделей «взрывного» распространения информации в сетях, в которых передача сигнала происходит в соответствии с принципом «с большей вероятностью — к менее популярному «узлу», обладающим некоторым правдоподобием с точки зрения ранее предложенных моделей иерархического упорядочения процессов передачи образцов в социокультурном цикле (см., например, [25]).

Ещё один яркий пример сверхмасштабного эксперимента, ставшего возможным исключительно благодаря социальным сетям и изменившего сложившиеся подходы к исследованию социального влияния, — это исследование Р. Бонда и соавторов, посвящённое методам мобилизации избирателей для участия в голосовании [26]. Эта группа исследователей показала, что доступ к информации об участии в голосовании, реализованный как доступная более чем 60 млн. пользователей Фейсбука в формате сообщения кнопка «Я проголосовал» (исходно речь шла о выборах в Конгресс США, состоявшихся в 2010 г.), оказала прямое и непосредственное влияние на реальное электоральное поведение не только самих участников, но и на их друзей, и на «друзей друзей», продемонстрировав «силу сильных связей»².

Важнейшим направлением инновационных интернет-исследований, как уже говорилось, является анализ больших массивов данных. Обычно под «Большими данными» понимаются не только нереактивные текстовые данные (архивы), но и любые сверхмасштабные массивы «цифровых отпечатков» онлайн-транзакций, геотрекинга и иных данных, в том числе получаемых с камер сле-

² Участники эксперимента получали специальное сообщение, содержавшее либо не содержавшее, помимо предложения воспользоваться кнопкой, информацию об уже проголосовавших друзьях испытуемого.

жения, а также с помощью сенсоров мобильных устройств, в частности, камер и бесконтактных датчиков смартфонов и т.п., однако в социальных науках основной интерес пока прикован к огромным объёмам оцифрованной текстовой информации, доступной во всемирной сети — от записей в блогах до архивов медиа, онлайн-чатов и т.п. Существуют значительные технические, организационные и этические проблемы доступа к таким данным (доступ ко многим из них, например, к переписке пользователей в чатах, возможен лишь в случае сотрудничества с крупными ИТ-компаниями, см. интересный и этически небесспорный пример подобного исследования [27]), однако здесь мы коротко остановимся лишь на некоторых принципиально новых перспективах для развития невозможных в «доцифровую» эпоху подходов к анализу текстов в социологии и сопряжённых с ними вызовах.

Большие объёмы личных записей в блогах, публикаций медиа и т.п. могут, безусловно, анализироваться с помощью усовершенствованных, но сохраняющих традиционную логику контент-анализа полуавтоматизированных процедур, предполагающих те или иные формы ручного кодирования на завершающих этапах. Например, в исследовании К.А. Гаврилова и А.Д. Толмача [28] были продемонстрированы значительные возможности оригинального подхода к итеративной разработке категорий виновных акторов для исследования текстов, размещённых в интернет-дневниках русскоязычных пользователей после теракта в аэропорту «Домодедово» в 2011 г. Однако всё большую популярность в последние годы приобретают методы автоматического анализа масштабных баз доступных в сети текстовых данных, предполагающих использование таких подходов, как автоматическая кластеризация и латентное тематическое моделирование текстов на основе размещения Дирихле [29], в том числе публикаций сетевых СМИ, записей в блогах, Твиттере и т.п. Обзоры сравнительных методологических преимуществ, сложностей и особенностей применения разных алгоритмов для анализа текстов, а также имеющегося опыта их применения в социологических и культурологических исследованиях представлены в работах [30; 31]. И всё же даже самые успешные примеры применения пока демонстрируют нерешённость многих проблем ав-

томатического текстового анализа. Так, в известном исследовании М. Бака и соавт. [32], посвящённом массовой динамике аффекта во время и после событий 11.09.2001, использовался автоматический анализ сообщений, отправленных на 85000 пейджеров, доступ к которым был получен посредством Wiki Leaks. Автоматический анализ сообщений, для которого использовалась кодировка переживаний с помощью распознавания обозначающих эмоции слов, объединённых в укрупнённые категории аффекта (например, для «печали» — «плач», «скорбь» и т.д.), позволил выявить некоторые тренды в эмоциональном компоненте социетальной реакции на трагические события — в частности, опережающий, но быстро возвращающийся к базовому уровню рост тревоги, постепенное возрастание гнева и т.п. К сожалению, более детальный анализ результатов автоматического кодирования выявил, что примерно 35% слов из категории «гнев» представляли собой автоматически сгенерированное сообщение о *критической* необходимости перезагрузки компьютера, передававшееся с возрастающей частотой на один пейджер [33].

Разумеется, описанные сложности, возникающие в этих и других новаторских исследованиях — это естественные «болезни роста», преодоление которых должно стать ближайшей задачей и для академических исследователей, и для исследователей-практиков. Однако, как справедливо отмечают авторы вышеупомянутого обзора [4], сама возможность участия социологов, психологов и других представителей наук об обществе и поведении в исследовательских проектах, использующих плоды «цифровой революции», существенным образом зависит от способности первых изменить не только свои привычные представления о тематике и методах исследований, но и овладеть новым техническим навыкам, необходимыми при работе с большими массивами полуструктурированных данных — от создания баз данных и работы с API до знакомства с алгоритмами машинного обучения, используемыми при анализе текстов.

ЛИТЕРАТУРА

1. Gosling S.D., Mason W. Internet research in psychology // *Annu. Rev. Psychol.* 2015. Т. 66. P. 877–902.
2. Девятко И.Ф. Инструментарий онлайн-исследований: попытка каталогизации // *Онлайн-исследования в России 3.0* / Под ред. Шашкина А.В. Девятко И.Ф., Давыдова С.Г. М.: Издательский дом «Кодекс». 2012. С. 17–30.
3. Skitka L.J., Sargis E.G. The Internet as psychological laboratory // *Annu. Rev. Psychol.* 2006. Vol. 57. P. 529–555.
4. Golder S.A., Macy M.W. Digital footprints: opportunities and challenges for online research // *Annu. Rev. Sociol.* 2014. Vol. 40. P. 129–152.
5. Sears D.O. College sophomores in the lab: influences of a narrow data base on social psychology's view of human nature // *J. Personal. Soc. Psychol.* Vol. 51. P. 515–530.
6. Henrich J., Heine, S.G., Norenzayan, A. The weirdest people in the world? // *Behav. Brain Sci.* 2010. Vol. 33. P. 61–135.
7. Gosling S.D., Vazire S., Srivastava S., John O.P. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires // *Am. Psychol.* 2004. Vol. 59. P. 93–104.
8. Goldberg, L.R. The development of markers for the Big-Five factor structure // *Psychological Assessment*. 1992. Vol. 4. P. 26–42.
9. Youyou W., Kosinski M., Stillwell D. Computer-based personality judgments are more accurate than those made by humans // *PNAS*. 27 January 2015 112 (4). P. 1036–1040.
10. Paolacci G., Chandler J., Ipeirotis P.G. Running experiments on Amazon Mechanical Turk // *Judgment and Decision Making*. 2010. Vol. 5 (5). P. 411–419.
11. Al Baghal T. (ed.). *Understanding Society Innovation Panel Wave 6: results from methodological experiments: Working Paper Series*. Colchester, Essex: Institute for Social and Economic Research. 2014. No. 2014–04.
12. Weigold A., Weigold I.K., Russel E.J. Examination of the equivalence of self-report survey-based paper-and-pencil and Internet data collection methods // *Psychological Methods*. 2013. Vol. 18. P. 53–70.
13. Некрасов С.И. Сравнение результатов онлайн- и офлайн-опросов (на примере анкет разной сложности) // *Социология: методология, методы, математическое моделирование*. 2011. № 32. С. 53–74.
14. Быков А.В. Метод факторных виньеток и шкала самооценки альтруизма: сравнение онлайн- и офлайн-опросов // *Социология: методология, методы, математическое моделирование*. 2014. № 39. С. 62–98.
15. Chandler J., Mueller P., Paolacchi G. Nonnaiveté among Amazon Mechanical Turk workers: consequences and solutions for behavioral researchers // *Behav. Res. Methods*. 2014. Vol. 46. P. 112–30.
16. Kuss D.J., Griffiths M.D., Karila L., Billieux, J. Internet addiction: a systematic review of epidemiological research for the last decade // *Curr. Pharm. Des.* 2014. Vol. 20 (25). P. 4026–4052.
17. Voiskounsky A.E. Internet addiction in the context of positive psychology // *Psychology in Russia: State of the Art*. 2010. Vol. 3. P. 541–549.
18. Kowalski R.M., Giumetti G.W., Schroeder A.N., Lattanner M.R. Bullying in the digital age: a critical review and meta-analysis of cyberbullying research among youth // *Psychol. Bull.* 2014. Vol. 140. P. 1073–1137.
19. Чистякова А.А. Миметическое воспроизведение культурных образцов (на примере сообщества imgur.com). Рукопись магистерской диссертации. Москва: НИУ ВШЭ, 2013.

20. Tanaka Y., Sakamoto Y., Matsuka T. Transmission of Rumor and Criticism in Twitter after the Great Japan Earthquake // *Proc. of Annu. Meet. Cogn. Sci. Soc.* 2012. P. 2387–2392.
21. Bordia P., DiFonzo N. Problem solving in social interactions on the Internet: rumor as social cognition // *Soc. Psychol. Q.* 2004. Vol. 67 (1). P. 33–49.
22. Tufekci Z., Wilson C. Social media and the decision to participate in political protest: observations from Tahrir Square // *J. Commun.* 2012. Vol. 62 (2). P. 363–379.
23. Rich F. Wallflowers at the revolution // *New York Times*. 2011.
URL: <http://www.nytimes.com/2011/02/06/opinion/06rich.html>.
24. Centola D. The spread of behavior in an online social network experiment // *Science*. 2010. September (329). P. 1194–1197.
25. Моль А. Социодинамика культуры / Перев. с фр. М.: Издательство АКИ. 2008. (3-е изд.).
26. Bond R.M., Fariss C.J., Jones J.J., Kramer A.D.I., Marlow C., Settle J.E., Fowler J.H. A 61-million-person experiment in social influence and political mobilization // *Nature*. 13 September 2012. (489). P. 295–298.
27. Das S., Kramer A. Self-Censorship on Facebook // *Proc. 7th International AAAI Conf. Weblogs and Soc. Media (ICWSM)*. Palo Alto, CA: AAAI, 2013. P. 120–127.
28. Гаврилов К.А., Толмач А.Д. Обыденное приписывание ответственности в ситуации теракта: опыт выделения категорий виновных акторов для контент-анализа // *Социология: Методология, методы, математическое моделирование*. 2012. № 35. С. 119–154.
29. Blei D.M., Ng A.Y., Jordan M.I. Latent Dirichlet Allocation // *Journal of Machine Learning Research*. 2003. Vol. 3. P. 993–1022.
30. Кольцова О.Ю., Маслинский К.А. Выявление тематической структуры российской блогосферы: автоматические методы анализа текстов // *Социология: методология, методы, математическое моделирование*. 2013. № 36. С. 113–139.
31. Mohr J.W., Bogdanov P. Introduction — topic models: what they are and why they matter // *Poetics*. 2013. Vol. 41. P. 545–569.
32. Back M.D., Küfner A.C.P., Egloff B. The emotional timeline of September 11, 2001. // *Psychol. Sci.* 2010. Vol. 21. P. 1417–1419.
33. Pury C.L.S. Automation can lead to confounds in text analysis: Back, Kufner, and Egloff (2010) and the not-so-angry Americans // *Psychol. Sci.* 2011. Vol. 22. P. 835–836.

Опыт прогнозирования развития русской онлайн-индустрии

*Давыдов Сергей Геннадьевич
(Национальный исследовательский университет
«Высшая школа экономики»)*

*Казарян Карен Размикович
(Российская ассоциация электронных
коммуникаций)*

*Логанова Ольга Сергеевна
(Национальный исследовательский университет
«Высшая школа экономики»)*

«Экономика Рунета»: методология исследования

Настоящая статья подготовлена по результатам исследовательского проекта «Экономика Рунета», совместно реализуемого Факультетом коммуникаций, медиа и дизайна Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ) и НП «Российская ассоциация электронных коммуникаций» (РАЭК) с 2012 г. Подробное обоснование исследовательского подхода, а также методология первой волны исследования нашли отражение в статье, опубликованной в предыдущем выпуске сборника «Онлайн-исследования в России»¹. Логика изложения предполагает, что ниже мы дадим краткую характеристику проекта в целом и остановимся более подробно на методологии второй и четвертой волн, результаты которых имеют прямое отношение к тематике данной публикации.

Основной задачей исследования «Экономика Рунета» является оценка текущего состояния и выявление тенденций развития рос-

¹ Давыдов С.Г., Кирия И.В. Экономика Рунета 2011–2012: методика и результаты измерения российских рынков / Онлайн-исследования в России 3.0. М. Online Marketing Intelligence. 2012. С. 297–323.

сийских онлайн-рынков. В период реализации проекта российская онлайн-экономика находилась в состоянии активного развития, появлялись новые активные игроки, новые ниши, новые продукты и целые рынки; все это привело к изменению структуры рынков. Кроме того, обсуждение результатов проекта среди экспертного сообщества индустрии помогло более четко структурировать данный вопрос. В первой волне исследования их было 11. В 2012 г. их число было увеличено до 14, причем все они были разделены на 4 сегмента. Принцип группировки рынков был сохранен в 3 и 4 волнах, но по мере появления были введены новые; списки изучаемых рынков приведены в Таблице 1. Как видно из последней, основные изменения следующие.

1. Рынок цифрового контента в итоге был разделен на 3 рынка (видео, музыка и книги) и перенесен из второго сегмента в четвертый, сформировав, таким образом, сегмент электронного контента.
2. Расширена дефиниция рынка Веб-разработки путем добавления мобильной разработки.
3. В сегмент «Электронная коммерция» добавлен сегмент «Прочие услуги».

Отметим, что в каждой из волн исследования были решены дополнительные исследовательские задачи, а именно:

- 2012 г. — исследование пяти конкурентных сил на онлайн-рынках (по методике М. Портера);
- 2013 г. — прогноз развития российских онлайн-рынков на 5 лет (до 2018 г.);
- 2014 г. — отношение индустриальных экспертов к изменениям в российском онлайн-законодательстве;
- 2015 г. — отношение индустриальных экспертов к оценке текущего состояния экономики («кризисный барометр»).

Полевая часть исследовательского проекта 2013 г. состояла из двух этапов.

1. 9 форсайт-сессий, проведенных в Москве в период с апреля по октябрь 2013 г. К участию в данном этапе проекта были привлечены 24 индустриальных эксперта, а также представители рабочей группы по реализации исследования из РАЭК и НИУ ВШЭ.

2. Опрос экспертов методом онлайн-анкетирования. Данный этап исследования был реализован с 16 сентября по 6 октября 2013 г. Объем выборки — 292 респондента.

В 2015 г. исследовательские процедуры были ограничены экспертным опросом методом онлайн-анкетирования. В данном случае в выборку вошли 153 эксперта.

В обоих случаях процедура онлайн-анкетирования была реализована при помощи портала Runet-ID.

Таблица 1. Сегменты и рынки исследования «Экономика Рунета» во 2 и 4 волнах (2013 и 2015 гг.)

Исследование 2013 г. (2 волна)		Исследование 2015 г. (4 волна)	
Сегменты	Рынки	Сегменты	Рынки
1. Маркетинг и реклама	1. Веб-разработка 2. Медийная реклама / Display 3. Контекстная реклама / Performance 4. Видеореклама 5. Маркетинг и коммуникации в социальных сетях 6. Поисковая оптимизация	1. Маркетинг и реклама	1. Веб-разработка и мобильная разработка 2. Контекстная реклама / Performance 3. Медийная реклама / Display 4. Видеореклама 5. Маркетинг в социальных медиа (SMM) 6. Поисковая оптимизация
2. Электронная коммерция	7. Онлайн-ритейл 8. Онлайн-туризм 9. Электронные платежи 10. Цифровой контент	2. Электронная коммерция	7. Онлайн-ритейл 8. Онлайн-туризм 9. Электронные платежи 10. Прочие услуги
3. Инфраструктура	11. Домены 12. SaaS 13. Хостинг	3. Инфраструктура	11. Домены 12. SaaS 13. Хостинг
4. Игры	14. Игры	4. Электронный контент	14. Игры 15. Видео 16. Музыка 17. Книги

Общее состояние онлайн-индустрии в России

Как мы уже отметили, состав и дефиниции измеряемых рынков от волны к волне разнились, поэтому напрямую сопоставлять интегральные показатели экономики Рунета, полученные в рамках исследования, было бы не вполне корректно. Данные исследования демонстрируют устойчивый рост большинства рынков, однако были отмечены и отрицательные тенденции развития. В четвертой волне исследования было зафиксировано снижение динамики и даже отрицательная динамика по ряду рынков. В то же время рост крупных рынков (прежде всего, речь идет о сегменте электронной коммерции) предопределил продолжающийся рост индустрии в целом.

Обобщенные данные об объемах онлайн-рынков в 2011–2015 гг., полученные в рамках исследования «Экономика Рунета», представлены в Таблице 1.

Таблица 2. Динамика российских онлайн рынков в 2011–2015 гг.

№	Сегмент	Рынок	2011		2012		2013		2014		2015 (прогноз)	
			млрд. руб.	%	млрд. руб.	%	млрд. руб.	%	млрд. руб.	%	млрд. руб.	%
1	Маркетинг и реклама	Медийная реклама (display)	15,83	41	19,2	21	23	20	22	-5	19,8	-10
2		Контекстная реклама (performance)	24,24	53	37,55	55	57	52	70	23	79,8	14
3		Видеоре-клама	0,8	53	1,74	118	2,9	67	3,8	30	4,45	17
4		Маркетинг в социальных медиа (SMM)	2,98	43	4,67	57	6,3	35	7,4	18	8,51	15
5		Поисковая оптимизация	8,56	18	10,24	20	12	17	13	9	12,6	-3
6		Мобильная и Web-разработка ¹	9,83	31	12,27	25	21	71	24,2	15	25,4	5

№	Сегмент	Рынок	2011		2012		2013		2014		2015 (прогноз)	
			млрд. руб.	%	млрд. руб.	%	млрд. руб.	%	млрд. руб.	%	млрд. руб.	%
7	Электронная коммерция	Онлайн ритейл	309,4	30	284,96	28	368	29	475	29	570	20
8		Онлайн туризм			153,2	40	195	27	279	43	315,3	13
9		Электронные платежи	166,7	34	268,7	61	350	30	476	36	571,2	20
10		Другие услуги	—	—	—	—	—	—	—	—		
11	Инфраструктура	Хостинг	5,32	25	4,38	26	4,9	12	5,6	15	6,0	8
12		Домены			2,18	18	2,5	15	2,6	5	2,6	0
13		ПО как услуга (SaaS)	1,89	46	3,22	70	4,3	34	5,4	26	6,3	17
14	Электронный контент	Игры ²	8,24	26	28,58	21	35	22	41	16	46,7	14
15		Музыка			5,07	32	2,0	10	2,2	8	2,33	6
16		Книги					0,6	60	0,96	60	1,69	77
17		Видео					3,2	10	3,5	10	3,8	9

¹ До 2013 г. измерялась только Web-разработка без мобильной разработки.

² В 2013 определение рынка «Игры» было изменено, с чем связано существенное увеличение оценки по сравнению с 2011 г.

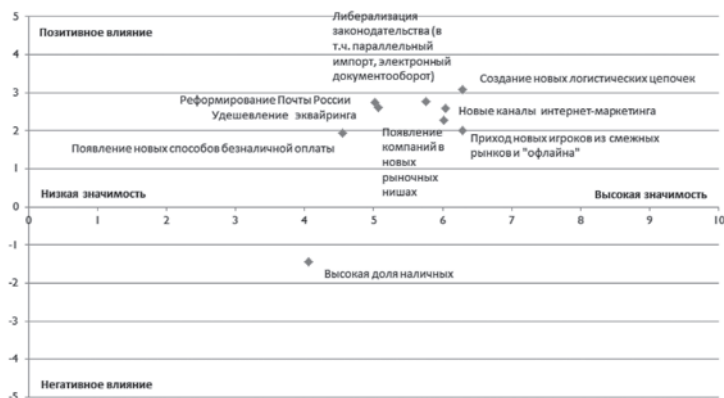
Пятилетний прогноз развития российской онлайн-индустрии (2013 г.)

Среднесрочное прогнозирование развития российской онлайн-индустрии было осуществлено следующим образом. На первом этапе были проведены форсайт-сессии по основным рынкам исследования, в рамках которых экспертами были сформулированы тенденции, ожидаемые в различных сегментах экономики, а также факторы, которые будут оказывать на них влияние. Также на данном этапе были сформулированы ко роткие формулировки прогнозов развития сегментов — оптимистический и пессимистический. Данные результаты были использованы при составлении анкеты количественного экспертного опроса. В рамках последнего тенденции и факторы влияния оценивались по основаниям значимости и характера влияния, а прогнозы — по вероятности их реализации. Далее

на основании двух полевых этапов исследования были разработаны по 4 сценария развития каждого из индустриальных сегментов, а также индустрии в целом. Сценарные разработки по сегментам приведены в Приложении 2. Здесь же мы остановимся на результатах количественного опроса, после чего приведем 4 сценария развития онлайн-индустрии, опубликованные в 2013 г.

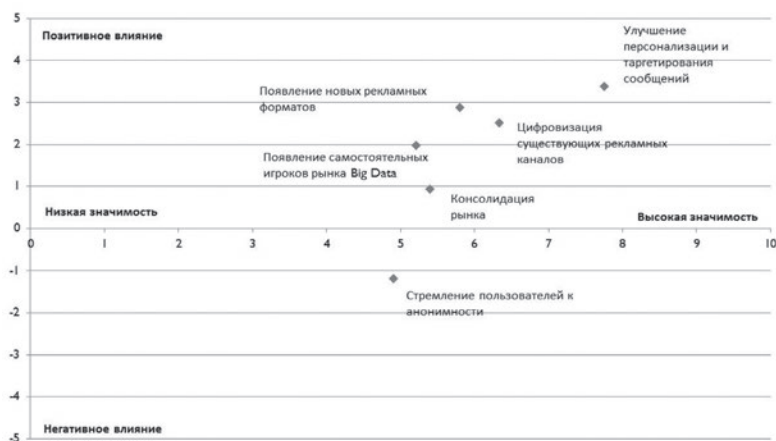
Итак, данные, полученные по сегменту «Электронная коммерция», свидетельствуют о том, что на первом месте (как по значимости, так и по уровню положительного влияния) находится создание новых логистических цепочек. Перечислим остальные тенденции в порядке убывания их значимости: приход новых игроков из смежных рынков и «офлайна», новые каналы интернет-маркетинга, появление компаний в новых рыночных нишах, либерализация законодательства (в том числе параллельный импорт, электронный документооборот), удешевление эквайринга, реформирование Почты России, появление новых способов безналичной оплаты. Единственной негативной тенденцией в сегменте «Коммерция и платежи» в период до 2018 года экспертами было признано сохранение высокой доли наличных. Впрочем, она также наименее значима из 9 названных.

Рисунок 1. Тенденции, ожидаемые в сегменте «Электронная коммерция» (2013–2018 гг.)



Среди названных экспертами шести тенденций, которые будут актуальны в сегменте «Маркетинг и реклама» в течение 2013–2018 гг., пять оцениваются в качестве значимых и позитивных. При этом наиболее заметна тенденция улучшения персонализации и таргетирования сообщений. Обратная тенденция — стремление пользователей к анонимности — рассматривается в качестве существенно менее сильной. Кроме того, в числе других положительных тенденций были названы цифровизация существующих рекламных каналов, появление новых рекламных форматов и самостоятельных игроков на рынке Big Data, а также консолидация рынка.

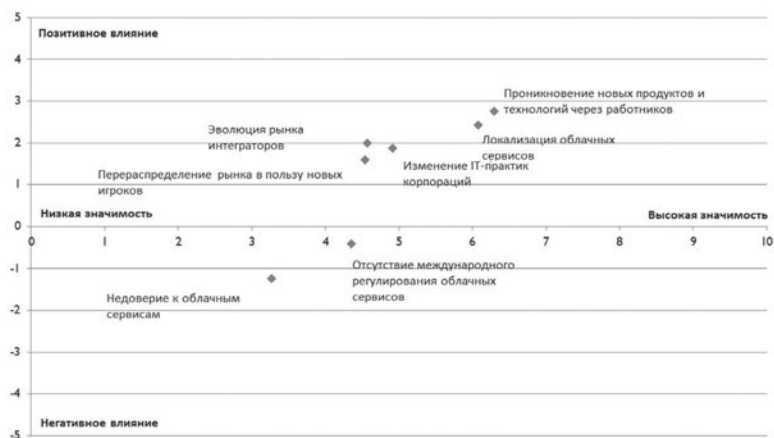
Рисунок 2. Тенденции, ожидаемые в сегменте «Маркетинг и реклама» (2013–2018 гг.)



Наиболее важными тенденциями российского инфраструктурного сегмента в среднесрочной перспективе стало проникновение новых продуктов и технологий через работников, а также локализация облачных сервисов. Данные тенденции представляются не только наиболее актуальными для периода анализа, но и наиболее позитивными. В числе положительных тенденций со средней значимостью также изменение ИТ-практик корпораций, эволюция рынка интеграторов, перераспределение рынка в пользу новых игроков; скорее негативной опрошенным видится проблема отсутствия меж-

дународного регулирования облачных сервисов. Тенденция сохранения недоверия к облачным сервисам оценена в качестве наиболее негативной, однако сила ее влияния — самая низкая.

Рисунок 3. Тенденции, ожидаемые в сегменте «Инфраструктура» (2013–2018 гг.)

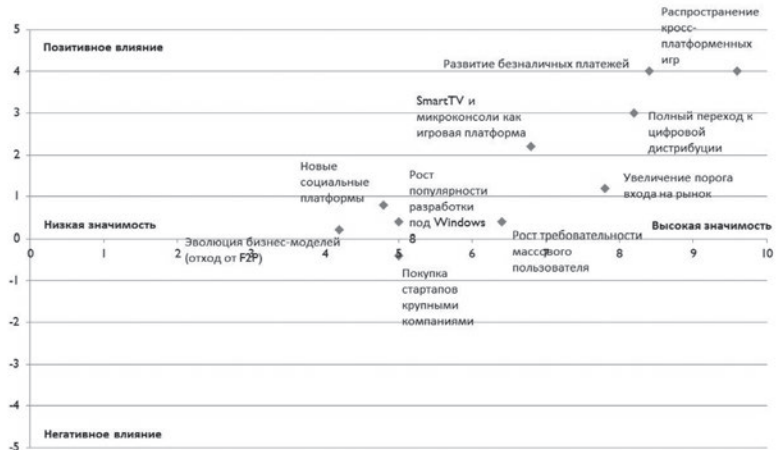


Всего участниками исследования сегмента «Игры» были названы 10 тенденций, которые будут определять развитие рынка в ближайшие пять лет. При этом 9 из них рассматриваются в качестве позитивных. Единственной негативной тенденцией названа покупка стартапов крупными компаниями; значимость ее оценивается как средняя, а отрицательный характер выражен весьма слабо.

Наиболее позитивным и одновременно значимым названо расширение кросс-платформенных игр. Столь же позитивным, однако немного менее значимым эксперты считают развитие безналичных платежей. На третьем месте по обоим параметрам — полный переход к цифровой дистрибуции. В числе других позитивных тенденций с высокой значимостью — Smart TV и микроконсоли как игровая платформа, увеличение порога входа на рынок, рост требовательности массового пользователя и рост популярности разработки под Windows 8. Позитивного влияния опрошенные представители рынка ждут также от новых социальных платформ

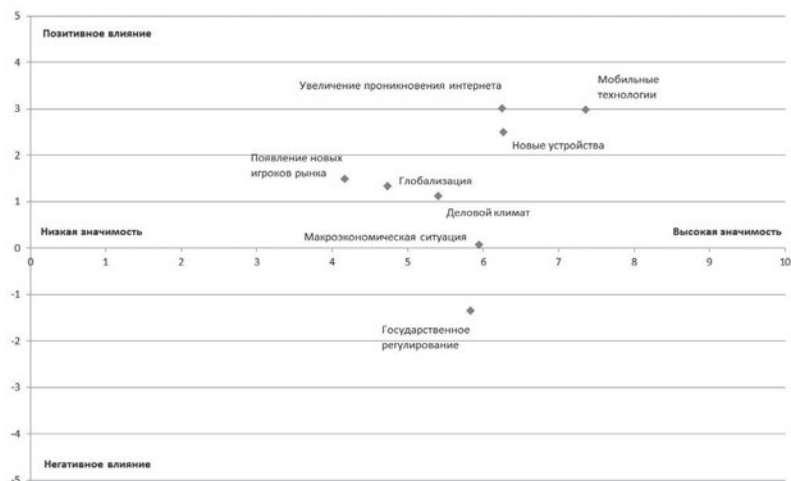
и эволюции бизнес-моделей (отход от F2P), однако их значимость будет сравнительно невысокой.

Рисунок 4. Тенденции, ожидаемые в сегменте «Игры» (2013–2018 гг.)



Среди ожидаемых факторов влияния на онлайн-рынках в целом наиболее значимым эксперты признали развитие мобильных технологий, что потенциально приведет к массовым изменениям и перераспределению ресурсов. Влияние этого фактора расценивается как весьма позитивное. Еще два фактора оцениваются как достаточно позитивные — это новые устройства и увеличение проникновения интернета, однако их значимость немного ниже. Другие положительные и значимые факторы — появление новых игроков рынка и глобализация. В группу позитивных факторов с низкой значимостью отнесены деловой климат и макроэкономическая ситуация. Наконец, государственное регулирование, по мнению экспертов, будет иметь слабое негативное влияние, его значимость оценивается как среднее.

Рисунок 5. Ожидаемые факторы влияния на экономику Рунета (2013–2018 гг.)



Помимо тенденций и факторов на онлайн-рынках экспертам было предложено дать оценку предложенным прогнозам развития сегментов, сформулированным на форсайт-сессиях. В каждом случае были приведены оптимистический и пессимистический прогнозы. Тексты прогнозов и оценка вероятности реализации каждого из них в пятилетней перспективе представлены в Таблице 3. Развернутые прогнозы, включающие оптимистический, пессимистический и два нейтральных сценария развития по каждому сегменту, приводятся в Приложении 2.

Таблица 3. Краткие экспертные прогнозы развития сегментов экономики Рунета в пятилетней перспективе и вероятность их реализации по оценкам экспертов (в%)

Сегмент	Оптимистический прогноз	Пессимистический прогноз
1. Электронная коммерция	Рост выше 30% на фоне благоприятной экономической обстановки, введение мер по стимулированию безналичных платежей и защите, развитие внутреннего рынка. Вероятность реализации — 53%.	На фоне экономического спада обещания государства по стимулированию сегмента рынка не будут выполнены. Рост около 10%. Вероятность реализации — 47%.
2. Маркетинг и реклама	Сохранятся текущие темпы роста выше 20%. Проникновение интернета к 2018 году достигнет 90%; произойдет перераспределение рекламных бюджетов от телевидения к интернету. Вероятность реализации — 71%.	На фоне экономического спада и жесткого регулирования интернета популярность традиционных каналов медиа сохранится. Темпы роста — менее 10. Вероятность реализации — 29%.
3. Инфраструктура	На фоне улучшения делового климата будет быстро повышаться доля использования интернета и облачных сервисов компаниями. Темпы роста превысят 50%. Вероятность реализации — 48%.	На фоне экономического спада, усиления регулирования интернета и монополизации инфраструктуры Рунета будет происходить отток клиентов к иностранным компаниями. Темпы роста ниже 20%. Вероятность реализации — 52%.
4. Игры	Рост выше 20% при введении мер по стимулированию безналичных платежей и невмешательства государства в регулирование интернета. Вероятность реализации — 56%.	Рост около 10% при активном регулировании интернета государством и отсутствии мер по стимулированию делового климата. Вероятность реализации — 44%.

Наконец, сценарные разработки по каждому из сегментов были сведены в четыре вероятных сценария развития российской онлайн-индустрии в целом. Реализация того или иного сценария могла произойти в зависимости от внутренней экономической политической конъюнктуры, а также от динамики и вектора глобальных экономических процессов, в которые вовлечена Российская Федерация. Приведем описание этих сценариев.

1. Консервативный сценарий

В случае реализации данного пути развития неуклонно будет расти количество законопроектов, касающихся регулирования интернета, а также повышаться негативная динамика экспертной оценки регулирующего воздействия примерно 40–50% этих законодотворческих инициатив.

Эксперты отмечают, что при реализации «консервативного» сценария развития Рунета экономическая целесообразность регулирования и динамика роста рынка станут менее значимыми, пренебрежимыми на фоне общемировых изоляционистских тенденций и политических реалий.

Меры по стимулированию делового климата не будут предприниматься или станут неэффективными. В условиях общемирового экономического спада произойдет постепенная монополизация инфраструктуры, экспансия телекомов на рынок контента и сервисов.

Популярность традиционных каналов маркетинга сохранится.

Государство будет присутствовать на рынке в качестве регулятора; роль государства как игрока минимальна, в госполитике будут преобладать долгосрочное планирование, поддержка крупного бизнеса и ситуативное реагирование.

Кроме того, произойдет заметный отток клиентов к иностранным компаниями и в «офшорные» юрисдикции.

2. Инновационный сценарий

При реализации данного сценария в мире в целом и в России в частности, будут реализованы программы, направленные на повышение конкурентоспособности, открытости. В мире и особенно в развивающихся странах будет быстро повышаться доля использования интернета компаниями, в том числе в области маркетинга. Миллионы людей получают доступ к образованию, базовым услугам посредством интернета. Произойдет бурное развитие контентных сервисов, интернет станет главным источником доходов для индустрии развлечений.

Безналичные, мобильные платежи будут в дальнейшем стимулировать развитие бизнеса, особенно в развивающихся странах, где

доступ к финансовым услугам затруднен. В России на фоне улучшения делового климата и благоприятной экономической обстановки, будет быстро повышаться доля использования интернета компаниями, в том числе в области маркетинга.

Государство проведет ряд мер по стимулированию безналичных платежей и развитию внутреннего рынка и вернется к политике невмешательства в регулирование интернета.

В полной мере реализуются федеральные и региональные программы, направленные на повышение конкурентоспособности и экономической эффективности российского сегмента интернета с привлечением экспертов, научных деятелей по интернет-проблематике, международных аналитических агентств.

Таким образом, с одной стороны, государство активно присутствует на рынке в качестве одного из крупных игроков. С другой стороны, стимулируя развитие и диверсификацию сервисов с помощью регуляторных послаблений, оно предоставляет на рыночной основе доступ к собственной экономической инфраструктуре и повышает прозрачность условий ее использования.

В госполитике прогнозируется акцент на среднесрочные прогнозы, поддержка компаний малого и среднего форматов, стимулирование конкуренции и прозрачная система распределения и использования результатов госзаказов. Среднегодовой рост рынка до 2018 года составит 35–40%.

3. Стабилизационный сценарий

Данный сценарий предполагает, что международное регулирование интернет-индустрии по сравнению с 2012 годом не меняется. Наиболее привлекательными с инвестиционной точки зрения становятся интернет-индустрия развивающихся стран (в основном азиатских).

Усилятся тенденции к консолидации игроков рынка и влияние судебной системы на развитие инноваций и бизнес-моделей. На фоне невысокого роста экономики основной рост будет наблюдаться в сегментах электронной коммерции и рекламы.

Развитие «облачного» сегмента замедлится, в области использования персональных данных продолжится «дикий Запад», но будет

отложено внедрение проектов в области индустриального интернета, медицины.

Политика государства в регулировании интернет-индустрии по сравнению с текущей не поменяется; законотворчество в интернет-сфере замораживается.

Динамика основных показателей проникновения интернета остается на прежнем уровне. Усилятся тенденции к консолидации игроков рынка.

Государство будет выступать в роли крупного заказчика на рынке, однако механизмы и пути использования результатов госзаказов останутся непрозрачными и не будут иметь рыночного веса и экономического эффекта.

В описанных условиях будет преобладать ситуативное реагирование и метапозиционирование в проектах («информационная поддержка»).

Среднегодовой рост рынка до 2018 года составит 15–20%.

4. Кризисный сценарий

Наступление данного сценария возможно при резком ускорении нескольких вялотекущих процессов, прямо влияющих на интернет-индустрию и регулирование.

На наш взгляд, «триггером» подобных событий может стать новый экономический кризис, когда ряд более традиционных отраслей одновременно попытаются противостоять силам «креативного разрушения», которые несут в себе интернет-технологии.

Признаками наступления такого «кризисного» сценария развития Рунета могут стать:

1. стремление сразу нескольких крупных страновых интернет-сегментов ослабить влияние на интернет США и американских интернет-компаний при неспособности США адекватно ответить на вызов в связи с экономическим кризисом;
2. усиление общемировых тенденций к контролю над национальными зонами, в том числе через принятие новых законов и на уровне международных организаций (МСЭ и т.п.);

3. резкое обострение конфликта между телекоммуникационными и интернет-компаниями (сервисными и контентными), связанного с пирингом трафика, ОТА-сервисами и сетевой нейтральностью;
4. снижение темпов роста мобильного и обычного ШПД, удорожание доступа в сеть;
5. принятие под влиянием медийных корпораций ряда международных соглашений, закрепляющих реакционный режим регулирования интеллектуальной собственности в интернете (в варианте соглашений АСТА-ТРП-ТАФТА); телекомы станут крупнейшими игроками на рынке контента;
6. введение новых мер по защите конфиденциальности и персональных данных, что резко снизит инвестиционную привлекательность и эффективность новых рекламных моделей в Интернете, трансграничных облачных сервисов;
7. вытеснение интернет-компаний из сферы платежей и финансов. При этом рост интернет-экономики будет наблюдаться только в отдельных сегментах, произойдет резкое снижение уровня инноваций.

Среднегодовой рост рынка до 2018 года составит 3–6%, но будет наблюдаться только в отдельных сегментах (например, электронная коммерция). Ситуация кризиса 2008–2009 гг., когда интернет-индустрия Рунета росла на фоне общего снижения, не повторится.

Кризисный барометр (2015 г.)

В четвертой волне «Экономики Рунета» отдельный блок занимал «кризисный барометр». Данная группа вопросов была посвящена отношению экспертов к кризису в российской экономике в целом и онлайн-экономике в частности.

Оценки степени влияния экономического кризиса на различные сегменты экономики в 2014 и 2015 гг. приведены в Таблице 4. Данные последней позволяют сделать следующие выводы. Во-первых, в целом влияние кризиса на экономику оценивается как «выше среднего» (оценки 7–8 по шкале от 1 до 10). При этом эксперты

считают, что на экономику страны в целом кризис оказал более существенное влияние, чем на ее онлайн-сегмент. Если сила кризиса во всей российской экономике в 2014 и 2015 гг. ощущалась примерно в равной степени, то в интернет-сегмент кризис пришел позже. Онлайн-рынки ощутили его влияние в 2015 г. Среди различных сегментов наиболее заметен кризис в сфере электронной торговли.

Таблица 4. Экспертные оценки влияния экономического кризиса на различные сегменты экономики в 2014 и 2015 гг. (%)

«1» — влияние отсутствует, «10» — влияние крайне сильное

№	Сегмент	Оценки 1-3	Оценки 4-7	Оценки 8-10	Меди- ана
1	Российская экономика 2014	2,7	35,1	62,2	8
2	Российская экономика 2015	4,1	30,1	65,8	8
3	Интернет-рынки 2014	8,1	51,4	40,5	7
4	Интернет-рынки 2015	4,1	45,2	50,7	8
5	Сегмент «Маркетинг и реклама» 2014	4,7	47,1	48,2	7
6	Сегмент «Маркетинг и реклама» 2015	10,8	39,8	49,4	7
7	Сегмент «Инфраструктура» 2014	3,8	42,3	53,8	8
8	Сегмент «Инфраструктура» 2015	7,7	46,2	46,2	7
9	Сегмент «Электронная коммерция» 2014	8,0	44,0	48,0	8
10	Сегмент «Электронная коммерция» 2015	4,0	32,0	64,0	9
11	Сегмент «Электронный контент» 2014	10,0	60,0	30,0	7
12	Сегмент «Электронный контент» 2015	10,0	40,0	50,0	7

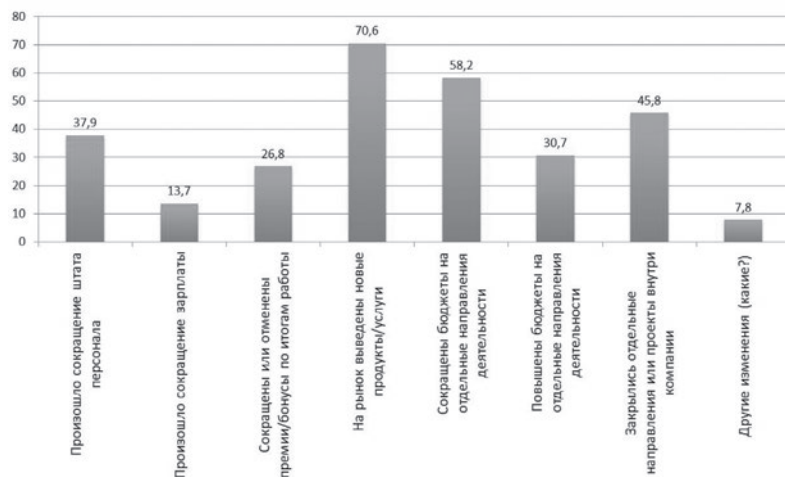
Согласно информации, полученной от экспертов, кризис значительно повлиял на политику подавляющего большинства компаний. Как правило, в предложенном списке произошедших перемен они выбирали несколько вариантов. По распределению ответов видно, что на фоне усиления конкуренции происходит оптимизация и структуризация процессов внутри компаний. Это вынуждает бизнес выводить на рынок новые продукты или услуги, данный вариант ответа выбрали около 70% опрошенных. Также происходит сокращение отдельных направлений, наименее доходных; часть проектов закрываются, данную тенденцию отметил каждый второй эксперт. Кроме того, примерно тридцать процентов экспертов указали на повышение бюджетов отдельных направлений, наиболее доходных, наиболее понятных заказчикам, наиболее востребованных.

Сокращение штата персонала — еще одна заметная тенденция, связана она с оптимизацией работы компании и сокращением бюджетов. Около 40% компаний вынуждены были оптимизировать персонал, то есть сократить своих сотрудников. Другим вариантом решения стало сокращение зарплаты, однако эта мера оказалась не слишком популярной: на нее пошли около 14% предприятий. Каждая четвертая компания сократила бюджет на персонал за счет сокращения или отмены его бонусной составляющей (26,8%), что позволило удержать часть людей, не увеличив или даже уменьшив их годовой доход.

Среди экспертов нет единых представлений о хронологии кризиса, причем это касается как его начала, так и предполагаемого завершения.

24,6% участников исследования полагают, что кризис начался в 4 квартале 2014 г., немногим меньше (19%) указывают на первый квартал 2015 года. Другими словами, наиболее распространенной является точка зрения, что кризис в интернет-индустрии начался на рубеже 2014–2015 гг. Вместе с тем четвертая часть экспертного сообщества отмечает кризисные тенденции уже во втором или третьем квартале 2014 г. Обратим внимание также на тот факт, что 11,1% опрошенных утверждают, что кризиса в онлайн-индустрии нет, тогда как 6,5% затруднились ответить на вопрос о его начале.

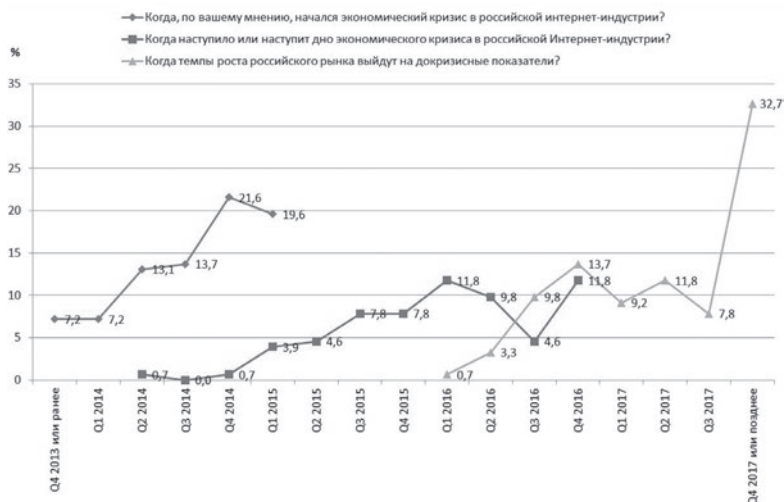
Рисунок 6. Распределение ответов на вопрос: «В приведенном списке укажите такие изменения, произошедшие в вашей компании за последний год, которые напрямую связаны с экономическим кризисом» (%)



Оценки экспертов свидетельствуют о том, что кризис воспринимается как затяжной. Достижение его дна прогнозируется в 2016 году: равное количество опрошенных (11,8%) прогнозируют, что это произойдет в 1 или 4 квартале этого года. Еще порядка 15% указывают на 2 или 3 квартал. Таким образом, падение онлайн-экономики будет происходить около полутора-двух лет. Стоит отметить, что содержательный ответ на вопрос дали 64,3% опрошенных, тогда как у 25,5% данный вопрос вызвал затруднение.

Восстановление онлайн-экономики, то есть возвращение динамики ее развития к докризисным показателям примерно треть экспертного сообщества (32,7%) прогнозирует к четвертому кварталу 2017 года и позднее. Примерно пятая часть аудитории настроена более оптимистично, допуская завершение кризиса к концу 2016 года. Но в целом возвращение экономики Рунета к кризисным показателям, согласно участникам опроса, займет около года-полутора при условии благоприятной экономической обстановки.

Рисунок 7. Мнение экспертов о временных рамках экономического кризиса



Выводы

Проведенное исследование подтвердило тот факт, что на основании экспертных оценок возможно не только получение консолидированных оценок и краткосрочных (годовых) индустриальных прогнозов в сфере онлайн-экономики, но и подготовка среднесрочных (пятилетних) сценариев развития индустрии. Безусловно, окончательно подвести итоги проделанной работы можно будет только по завершении периода прогнозирования, то есть в 2019 г. Однако и сейчас, когда прошло 2 года, очевидно, что происходящее развитие весьма близко к одному из сценариев — стабилизационному. И это при том, что два года, прошедшие после исследования, были весьма насыщены событиями, повлиявшими как на экономическое, так и на политическое развитие России.

По состоянию на 4 квартал 2015 г. рост онлайн-экономики продолжается. Исключение составляют два рекламных рынка, демонстрирующих отрицательную динамику, а также рынка доменов,

на котором отмечена нулевая динамика. Конечно, речь идет о рублевых оценках; учет произошедшего в исследуемый период резкого падения курса рубля относительно основных мировых валют (прежде всего, доллара и евро), а также инфляции позволил бы интерпретировать ситуацию иначе. Как бы то ни было, эксперты в основном согласны с наличием кризиса в онлайн-индустрии, хотя и соглашаются с тем, что он гораздо мягче по сравнению с другими сферами отечественной экономики. Хотя у экспертов нет консолидированного видения хода кризиса, большинство из них говорят об изменениях, произошедших в их компаниях в связи с изменившейся рыночной ситуацией. Наиболее часто при этом говорят о выводе на рынок новых товаров и/или услуг, а также сокращении корпоративных бюджетов на отдельные направления работы.

Приложение 1. Дефиниции исследуемых рынков (из анкеты 2015 г.)

I. Сегмент «Реклама и маркетинг»

1. Веб-разработка и мобильная разработка. Услуги в области разработки интернет-сайтов и мобильных приложений (включая дизайн и стиль, тестирование и проектирование их «юзабилити» и т.д.).
2. Контекстная реклама/ Performance. Услуги по размещению рекламы с оплатой за результат (поисковая реклама, таргетированная реклама, лидогенерация)
3. Медийная реклама / Display. Услуги по размещению рекламы внутри медиаконтента.
4. Видеореклама. Услуги по размещению видеорекламы внутри размещаемого в интернете видеоконтента (на хостингах видео, в онлайн-кинотеатрах и т.д.).
5. Маркетинг в социальных медиа (SMM). Услуги по продвижению товаров и услуг, а также связям с общественностью в социальных медиа, а также услуги аналитических сервисов.
6. Поисковая оптимизация. Услуги в области улучшения «видимости» сайтов в результатах поиска через поисковые системы.

II. Инфраструктура

7. Программное обеспечение как услуга (SaaS). Модель продажи и использования программного обеспечения, при которой поставщик разрабатывает веб-приложение и самостоятельно управляет им, предоставляя заказчику доступ к программному обеспечению через интернет.
8. Хостинг. Услуги по предоставлению вычислительных мощностей для размещения и/или хранения информации на серверах, постоянно находящихся в интернете.
9. Домены. Предоставление доменных имен в интернете.

III. Электронная коммерция

10. Онлайн-ритейл. Продажа физических товаров и офлайн-услуг (билеты на мероприятия, купоны и др.) через интернет.
11. Электронные платежи. Рынок онлайн-расчетов между компаниями и пользователями за товары и услуги в интернете.
12. Онлайн-туризм. Рынок продажи авиа и ж/д билетов, туров, бронирования гостиниц и доп. услуг (прямые продажи и агентские продажи).
13. Другие услуги.

IV. Электронный контент

14. Игры. Особый вид услуг, который предполагает доставку игр через интернет, включая игры в социальных сетях и на мобильных платформах, а также продажу виртуальных товаров и сервисов внутри игровых приложений.
15. Видео. Продажа, прокат и подписка на доступ к потоковому вещанию легального видеоконтента через интернет (не включая видеорекламу).
16. Музыка. Продажа, прокат и подписка на доступ к потоковому вещанию легального аудиоконтента через интернет.
17. Книги и СМИ. Продажа и подписка на электронные книги, периодические издания и СМИ.

Приложение 2. Описания сценариев развития сегментов онлайн-индустрии в пятилетней перспективе (из отчета 2013 г.)

1. Сценарии для сегмента «Электронная коммерция».

В сегменте «Электронная коммерция» наиболее сильной положительной тенденцией эксперты считают «создание новых логистических цепочек», а наиболее негативным влияющим фактором — «государственное регулирование».

Набор сценариев в сегменте «Электронная коммерция», таким образом, связан с ключевым драйвером, которым является усовершенствование механизмов доставки товаров и услуг, приобретаемых через интернет. На рынке контента это касается постепенного спада объемов продаж на физических носителях и вытеснение их онлайн-овым скачиванием на устройства, что сопряжено с подготовкой цифровых копий не на уровне дистрибуции, а на уровне издателя. В сегменте розничного интернет-ритейла речь идет о развитии в первую очередь офлайновой инфраструктуры поставок, то есть о создании отдельных супермаркетов, пунктов выдачи товаров, заказанных через интернет.

Что касается негативных факторов в сфере государственного регулирования, которые выступают тормозом в данном сегменте, то к таковым в сфере контента эксперты относят борьбу с пиратством, которая подрывает работу многих онлайн-овых поставщиков контента.

А на рынке ритейла, по мнению экспертов, наоборот, невмешательство государства не дает развиваться рынку. В первую очередь речь идет о том, что государство определенными действиями (ограничение оборота наличных денег, обязательство торговых точек принимать пластиковые карты, создание общенациональных Wi-Fi-сетей) могло бы стимулировать рост рынка интернет-коммерции.

Положительный сценарий. Таким образом, положительный сценарий развития сегмента «Электронной коммерции» будет предполагать стимулирование государством безналичных платежей, развитие логистики, в том числе за счет реформирования Почты

России и в то же время гибкое отношение к интернет-площадкам по продаже контента (в частности, чтобы исключить экономические потери легального рынка в связи с борьбой с пиратством).

Отрицательный сценарий. Данный сценарий предполагает, что доля наличных денег в экономике и отсутствие стимулирующих безналичный расчет действий со стороны государства приведут к замедлению темпов роста онлайнowych платежей и интернет-торговли, а ограничения естественного характера (транспортные, климатические, организационные) не дадут развиваться новым механизмам логистики и развитию цепей поставок. Все это будет сопровождаться драконовскими мерами по борьбе с пиратством, что негативно будет сказываться на рынке легального контента в интернете. В результате доля продаж на физических носителях перестанет сокращаться.

Нейтральный сценарий 1. В соответствии с данным сценарием государство будет прилагать усилия для ограничения оборота наличных денег и стимулирования безналичных расчетов. При этом регулятор будет гибко относиться к ограничению пиратства, чтобы не затронуть легальный рынок интернет-контента. Вместе с тем естественные факторы и конфигурация сетей поставок не дадут совершенствоваться сегменту и развивать новые модели и механизмы поставок, что также приведет к тому, что доля продаж контента на физических носителях перестанет сокращаться.

Нейтральный сценарий 2. В рамках данного сценария развитие цепочки поставок, логистики и совершенствование технологий почтовой доставки (а также развитие альтернативных крупных курьерских служб) дадут определенный импульс развитию отрасли. Вместе с тем такое развитие будет тормозиться высокой долей наличных денег в экономике (в связи с отсутствием принимаемых государством мер по стимулированию безналичных платежей) и жесткими мерами по борьбе с пиратством.

2. Сценарии для сегмента «Маркетинг и реклама».

Если рассматривать сегмент «Маркетинг и реклама», то в нем эксперты в качестве наиболее значимой положительной тенденции

(движущей силы) определили «улучшение персонализации и таргетирования сообщений». При этом наиболее негативным и значимым фактором также было названо «государственное регулирование»

Таким образом, набор сценариев в сегменте «Маркетинг и реклама» связан с ключевым драйвером, которым является разработка все более совершенных механизмов таргетирования аудитории, что, в свою очередь, стимулирует рост рекламного и маркетингового сегмента в интернете. Эксперты относят сюда слияние баз данных и сбор информации с разных видов сервисов (в том числе все большее использование данных самих рекламодателей, таргетирующих аудиторию через электронные каналы коммуникации), по которым перемещается пользователь. Кроме того, сюда же относится рост такого тренда, как учет геолокации потребителя в социальном маркетинге.

Что касается основных факторов в сфере государственного регулирования, тормозящих данный сегмент, эксперты называют среди них в первую очередь ужесточение законодательства в области персональных данных, что будет всячески препятствовать возможностям таргетирования, а это, в свою очередь, будет тормозить рост рекламного рынка во всем интернете. Кроме того, в качестве негативного фактора эксперты называют угрозу регулирования выдач поисковых систем (с целью борьбы с пиратством), что отрицательно скажется на рынке поисковой рекламы.

Положительный сценарий. В данном сегменте положительный сценарий предполагает, что все более и более концентрированные компании смогут собирать подробные и максимально таргетированные данные о потребителях, что будет способствовать существенному росту всего сегмента. При этом государство никак не будет ограничивать данную деятельность: ни через введение регулирования на выдачу результатов поиска в поисковых системах, ни через запрет на использование в коммерческих целях персональных данных.

Отрицательный сценарий. Данный сценарий, напротив, предполагает, что государственное регулирование в области персональных данных — как на национальном, так и на глобальном уровне — станет существенным фактором, тормозящим развитие отрасли, а дополнительное регулирование результатов выдачи в поисковых

системах приведет также к оттоку рекламодателей из поискового сегмента рекламы. В этом случае стоит ожидать перетекания бюджетов в медийную рекламу в интернете (включая видеорекламу).

Нейтральный сценарий 1. Исходя из этого сценария можно предположить, что государство не будет существенным образом ужесточать законодательство в области защиты персональных данных, а также вводить ограничения в сфере выдачи в поисковой рекламе, однако новые механизмы таргетирования постепенно перестанут появляться, за счет чего рост данной отрасли замедлится, и интернету, вполне вероятно, не удастся оттянуть большую часть бюджетов из традиционной медийной рекламы.

Нейтральный сценарий 2. В данном сценарии, по всей вероятности, рост способностей крупных интернет-агрегаторов, поисковиков и социальных сетей по таргетированию аудитории будет компенсироваться драконовским регулированием в сфере защиты персональных данных, запрета на их сбор в отдельных сегментах и регулированием поисковых машин. Это приведет к тому, что интернет-индустрия будет постоянно искать новые механизмы таргетирования аудитории, а государство постоянно будет изобретать новые механизмы регулирования этой сферы. В этом смысле рост отрасли замедлится, а сама она будет в значительной степени прирастать за счет медийной рекламы.

3. Сценарии для сегмента «Инфраструктура».

В сегменте «Инфраструктура» в качестве наиболее сильной и положительной тенденции эксперты выделили «Проникновение новых продуктов и технологий через работников», наиболее негативным фактором здесь оказалось «государственное регулирование».

Набор сценариев в сегменте «Инфраструктура», таким образом, связан с ключевым драйвером, которым является наиболее положительная тенденция, то есть технологический прогресс, в результате которого новые технологии — преимущественно в рабочей сфере (в бизнесе) — постепенно проникают в социальную жизнь. При этом эксперты отмечают, что данное проникновение будет в значительной степени опираться на спрос, формируемый

не технической и ИТ-политикой компаний, а конечными потребителями облачных сервисов.

Ключевым тормозом или, по крайней мере, негативным фактором этой модели выступают государственное регулирование и нормативная база, которая может создавать препятствия для развития данного сегмента. Среди элементов такой негативной нормативной базы участники экспертной панели называли регулирование области обязательств размещения серверов на территории России и монополизации каналов связи с зарубежными провайдерами.

Положительный сценарий. Если описывать таким образом положительный сценарий, то он будет предполагать максимальную вовлеченность пользователей на рабочих местах в использование облачных сервисов, что, в свою очередь, будет стимулировать бизнес к массовому корпоративному спросу на SaaS-сервисы. Это станет сопровождаться не только существенным ростом рынка и отсутствием мер прямого регулирования со стороны государства, но и наличием четких инструментов для гарантирования конфиденциальности хранящихся в облачных сервисах данных.

Отрицательный сценарий. Данный сценарий предполагает, что крупные ИТ-компании, интеграторы и в целом крупные производители ПО не дадут реализовать тотальное проникновение SaaS в бизнес-среду. При этом государственное регулирование будет делать упор на чрезмерные запретительные меры, такие как ограничения территориального размещения серверов облачных сервисов и хостинга на территории страны, а также в области регулирования связей между локальной и глобальной инфраструктурами.

Нейтральный сценарий 1. Если исходить из данного сценария, вмешательство государства будет касаться только гарантий основных прав (конфиденциальности в первую очередь) пользователей облачными службами, однако при этом степень распространенности SaaS будет небольшой в связи с лобби крупных компаний, предлагающих коробочные решения и ПО.

Нейтральный сценарий 2. Данный сценарий предполагает, что при положительной динамике роста и распространения облачных сервисов (в том числе на корпоративном уровне, в бизнесе) государственное регулирование будет в значительной степени сдерживать

данный рост за счет введения определенных технических регламентов, касающихся размещения серверов и сайтов, контролировать контент и т.д.

4. Сценарии для сегмента «Игры».

В данном сегменте ключевым трендом (движущей силой), по мнению представителей отрасли, является распространение кросс-платформенных игр (позволяющих играть на различных устройствах). Что же касается негативных факторов влияния, то также наиболее значимым из них участники отрасли определили государственное регулирование отрасли.

Таким образом, по мнению экспертов, ключевым драйвером в данном сегменте является рост «кросс-платформенности», что способствует росту рынка в целом и позволяет вовлекать клиента в более активное потребление игр и дополнительных сервисов к ним посредством их присутствия на разных медийных платформах. К этим платформам и к данной тенденции в сфере кросс-платформенности относятся игры на мобильных терминалах (с двумя основными платформами: Android и iOS) и в социальных сетях (соответственно, переносимые на любые устройства), а также консоли и набирающее вес интерактивное ТВ (Smart TV). При этом эксперты прогнозируют, в первую очередь, движение от игр в социальных сетях в сторону игр на мобильных терминалах, а также постепенную возможность играть в одну и ту же игру, интегрируя и соединяя разные устройства (например, телевизор и планшетный терминал). Помимо этого, к проблеме кросс-платформенности имеет в некотором роде отношение переход пользователей с одной игры на другую, что позволяет повысить шансы монетизации.

Как и во всех других сегментах, основным элементом риска участники отрасли считают политику государства в этой сфере. Сюда относится в первую очередь законодательное регулирование проблемы пиратства и нелегального копирования, которое, по мнению участников отрасли, приводит к избыточному регулированию сферы контента (и игр в частности), а также построенная волюнтаристским образом система маркировки контента.

Положительный сценарий. В данном сегменте положительный сценарий предполагает, что кросс-платформенность будет усиливаться: игры будут интегрироваться на различных устройствах (как персональных, так и коллективных типа телевизора), что, безусловно, приведет к росту спроса на игровой контент и платные его версии. При этом государство будет более активно защищать права правообладателей от пиратства, не регулируя рынок контента игр и не вводя запретов на отдельные их категории.

Отрицательный сценарий. Данный сценарий предполагает, что технологический прогресс, дрейф в сторону «закрытых платформ» и т.д. приведет к ограничениям дальнейшего развития кросс-платформенности и расширению игр на разных носителях. Это уменьшит шансы на дополнительные доходы большинства разработчиков. При этом ситуация будет усугубляться растущими ограничениями на определенного типа контент в играх, что само по себе приведет, помимо всего прочего, к сжатию определенных сегментов рынка, падению уровня креативности разработчиков и т.д.

Нейтральный сценарий 1. Исходя из данного сценария государство не будет существенным образом регулировать контент игровой индустрии и будет проводить эффективную политику в сфере защиты прав разработчиков. Вместе с тем рынок слегка затормозится в интенсивном развитии в связи с исчерпанием возможностей для дальнейшего расширения кросс-платформенного взаимодействия (за счет закрытых платформ и т.д.). Это приведет в значительной степени к монополизации игровой отрасли и ее зависимости от технологических платформ.

Нейтральный сценарий 2. Этот сценарий исходит из предположения, что рост кросс-платформенности и взаимодействия различных устройств и разработчиков будет компенсироваться относительно высоким уровнем пиратства и значительными мерами государства по вмешательству в контент игровой отрасли.

Раздел 2

Методология онлайн- исследований: дизайн инструментария

Разработка опросов в стиле бонсай^{©1}

Пулестон Джон
(*Global Market Insite, Inc.*)

В данной статье рассматриваются техники, необходимые для разработки опроса в стиле бонсай — короткого и четко сформулированного опроса.

В разработке дизайна опросов мы подошли к решающему моменту. Все понимают, что исследования должны меняться не для того, чтобы привлечь постоянно растущее количество людей, желающих использовать мобильные телефоны и планшеты, а для того, чтобы конкурировать с другими онлайн-развлечениями; опросы должны стать более приятным занятием для людей, которые все меньше хотят концентрировать внимание на чем-либо в течение продолжительного времени. Чтобы справиться с этими вызовами, должны появиться новый тип и новый язык дизайна опросов.

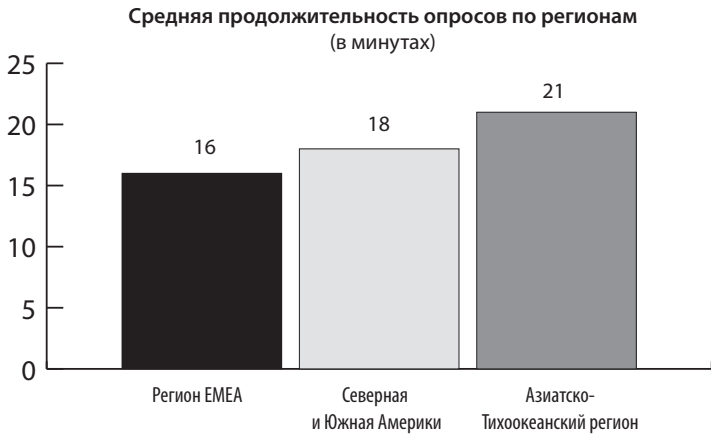
Приведенный ниже график служит иллюстрацией существующей проблемы. Данные предоставили коллеги из Lightspeed Research в 2013 году, они показывают, что только ¼ пользователей смартфонов хотят тратить больше 5 минут на заполнение опросов, и меньше трети готовы тратить больше 15 минут. Как адаптировать опросы, которые традиционно длятся около 20 минут или больше, под мобильные платформы?

¹ “© Copyright (2015) by ESOMAR® — The World Association of Research Professionals. This paper first appeared in Asia Pacific 2014 published by ESOMAR (www.esomar.org)

МАКСИМАЛЬНОЕ ВРЕМЯ ПРОХОЖДЕНИЯ ОПРОСА	КОМПЬЮТЕР	ПЛАНШЕТ	СМАРТФОН
Не более 5 минут	2%	9%	27%
Не более 10 минут	9%	24%	45%
Не более 15 минут	19%	42%	65%
Не более 20 минут	34%	65%	73%
Не более 25 минут	42%	71%	77%
Не более 30 минут	65%	81%	85%

Как отметил Берни Манилофф в своем докладе, посвященном дизайну опросов, на конференции ESOMAR 3D, даже небольшие изменения в исследовании могут вести к изменению его результатов, и даже незначительное изменение шкалы Лайкерта меняет ответы. Многие исследователи, вспоминая пословицу «погибать, так с музыкой», используют эту возможность, чтобы полностью пересмотреть свой подход к онлайн-исследованиям в надежде интегрировать в свою работу последние достижения в области повышения эффективности структуры и дизайна опросов.

Я представляю эту статью на конференции ESOMAR Asia по двум причинам. Во-первых, Азия является лидером по использованию мобильных исследований. В Азии число респондентов, участвующих в опросах с мобильных телефонов или планшетов, на треть выше, чем в Европе и Северной Америке, и составляет около 20%, а на некоторых рынках Азии достигает 25%*. Во-вторых, азиатские опросы обычно значительно длиннее, чем средние европейские онлайн-исследования, и занимают в среднем на 4 минуты больше времени, а кроме того, у них на 10% выше показатель числа покинувших опрос респондентов, что означает более низкое качество полученных данных по сравнению с другими рынками.



*Источник: данные исследования Lightspeed 2014 г.

О содержании работы:

Эта статья во многом является практическим руководством, основанным на нескольких годах работы с клиентами по всему миру и попыток улучшить дизайн-опросы. Я постарался выявить общие проблемы, которые воздействуют на длину опроса и степень вовлеченности в него респондентов, и предложить некоторые идеи о том, как их исправить.

Я понимаю, что большая часть статьи не является чем-то революционным. Многие методики, которые я выделяю, достаточно линейны и могут быть хорошо известны некоторым из вас, однако даже с учетом этих факторов я надеюсь, что данная работа может быть использована как своеобразный «чек-лист» возможных подходов.

Эта статья разбита на 2 части — можно назвать ее «двойным альбомом». В первой части будут представлены некоторые практические шаги для уменьшения длины опроса, а во второй части будут рассмотрены подходы к разработке опросов для мобильных телефонов и планшетов с применением различных способов повышения их эргономичности и увлекательности.

Часть 1: Пять ключевых этапов процесса сокращения опроса

Первая часть статьи определяет ключевые этапы процесса оптимизации длины опроса и максимизации его эффективности.

Этап 1: Приобретите «очки респондента».

Этап 2: Используйте итеративный метод разработки опроса.

Этап 3: Найдите и исключите избыточные вопросы.

Этап 4: Установите квоты на количество ответов и вопросов.

Этап 5: Примените более продвинутые методы приписывания и масштабирования.

Этап 1: Приобретите «очки респондента»

Если вам необходимо перестроить и сократить готовый опрос, я бы рекомендовал приобрести «очки респондента». К примеру, вот такие можно найти на сайте Ebay за вполне приемлемую сумму в 19,99 долларов США и с доставкой.



Надев такие «очки респондента», вы сами немедленно превратитесь в одного из ваших респондентов. После этого вам следует прочитать свой опрос и отметить каждый вопрос, который, как вы думаете, может показаться довольно скучным.

Затем передайте эти очки своим коллегам, друзьям, членам семьи и попросите их сделать то же самое. Я неоднократно замечал, что мои дети-тинэйджеры дают при этом честные и очень ценные ответы типа: «Никто не ответит на этот вопрос и за миллион лет, пап».

Сравнение всех полученных отзывов поможет вам составить четкое представление о том, над какими вопросами еще требуется поработать.

Это может показаться вам шутивым советом, но попытка взглянуть на ситуацию с точки зрения респондента — лучший способ понять, как можно улучшить свой опрос.

Выявление статистических индикаторов проблематичных вопросов

Если вы хотите провести более серьезную работу, и у вас есть данные прошлых опросов, вы можете проанализировать общие индикаторы, указывающие на то, что респонденту становится скучно при прохождении опроса:

1. Доля покинувших опрос респондентов
2. Слепое заполнение
3. Уменьшение затраченного на ответ времени

Выход респондента из опроса, особенно в середине исследования, это то, на что следует обратить особое внимание. Обычно этот показатель не превышает 1% для любого из вопросов исследования, но если он выше 3% — это уже является серьезным индикатором проблематичного вопроса (часто проблема заключается в том, что респондент не может перейти к следующему вопросу или не знает, что ответить).

Слепое заполнение, вероятно, является самым известным индикатором скуки респондента, главным признаком более общей проблемы шаблонных ответов. На каждого участника, чья скука проявляется в бездумном выборе одного и того же варианта ответа снова и снова, приходится довольно внушительное число (по нашим оценкам, как минимум треть) тех, кто просто выбирает ответ случайным образом. При этом подобные случайным образом выбранные варианты ответа обычно составляют некую систему (отсюда и термин — шаблонные ответы). Показатель слепого заполнения на уровне 8% сам по себе кажется не слишком высоким, однако он может указывать на то, что до 40% респондентов не отвечают на вопросы должным образом.

Мы считаем, что анализ среднего количества времени, затраченного на ответ на вопрос, является самым удачным способом выявле-

ния факторов, вызывающих скуку у респондентов. Это достаточно простой показатель, и — как показывает следующий график — вполне понятный для стороннего наблюдателя.

Пример снижения среднего времени ответа на вопросы



Этот пример взят из исследования, в котором респондентов просили оценить 44 отдельных продукта. К концу задания респонденты тратили на каждый вопрос менее двух секунд, что просто недостаточно для формирования суждения, особенно принимая во внимание тот факт, что только на внимательное прочтение задания требуется около двух секунд.

Вот некоторые практические правила:

- необходимо стремиться к тому, чтобы показатель числа покинувших опрос респондентов не превышал 1% для любого вопроса;
- любой вопрос с таблицей вариантов ответа, в котором респондент тратит менее двух секунд на каждую строку, служит индикатором снижения его внимания;
- то же относится и к любому вопросу с множественным выбором, в котором респондент должен отмечать варианты ответа галочкой, в том случае, если на выбор каждого из вариантов респондент тратит менее 0,5 секунды;
- степлер-тест: если вы распечатали опросную анкету, пытаетесь скрепить ее степлером и он не может пробить все листы анкеты — ОНА СЛИШКОМ ДЛИННАЯ!

Этап 2: Используйте итеративный метод разработки опроса

Итак, вы выявили проблематичные вопросы и поняли, что анкета слишком длинная. Как лучше всего сократить ее?

Почти все методы сокращения анкеты предполагают наличие у вас статистических данных, которые можно проанализировать и понять, какие вопросы убрать. И для многих исследователей это становится первой проблемой.

Если вы анализируете существующее трекинговое исследование или опрос, который уже проводили до этого, тогда, скорее всего, у вас есть все необходимые данные. Но если вы создаете опрос с чистого листа, то никаких данных у вас нет. Как же можно решить эту проблему?

Чтобы ее решить, необходимо применить итеративный подход к разработке дизайна опроса и включить в процесс фазу пилотного тестирования анкеты. Это, пожалуй, самый важный совет, который я могу вам предложить в этой статье.

Вместо того чтобы разрабатывать весь опрос целиком, сначала выделите основные части исследования, в которых вы не уверены, и проведите пилотное тестирование. Так вы получите необходимые данные, чтобы определить избыточные вопросы и установить требования к выборке.

Я понимаю, что для многих компаний это существенное изменение подхода, которое создает дополнительные временные и денежные затраты. Мы настоятельно рекомендуем проведение пилотного тестирования опросов, если затраты на отбор участников превышают 10 тысяч долларов, так как практически наверняка это в итоге экономит вам деньги.

Этап 3: Найдите и исключите избыточные вопросы

Первое, на что надо обратить внимание в опросе,— это избыточные вопросы, на которые даются похожие или такие же ответы, как на другие вопросы анкеты. Они обычно встречаются в списках утверждений для оценки и слов для ассоциаций.

Существуют довольно сложные статистические методы поиска таких вопросов, но самый простой и не менее действенный — это использование коэффициента корреляции Пирсона, которая присутствует в стандартных функциях программы Excel. Необходимо сравнить ответы на все вопросы друг с другом и отыскать среди них пары с наиболее высокими показателями корреляции.

В ситуации, когда у двух похожих вопросов корреляция составляет $R=\pm 0,65$, мы рекомендуем либо убрать один из вопросов, либо чередовать их в анкете так, чтобы каждый был виден только половине участников выборки.

А теперь мне бы хотелось вас предостеречь: Вы можете подумать, что не обязательно проводить статистический анализ, чтобы найти похожие вопросы. Неужели их невозможно увидеть на глаз? Иногда это действительно невозможно. Взгляните на представленный ниже пример — вы сможете указать вопросы с высоким показателем корреляции?

Как бы вы оценили компанию по следующим факторам:

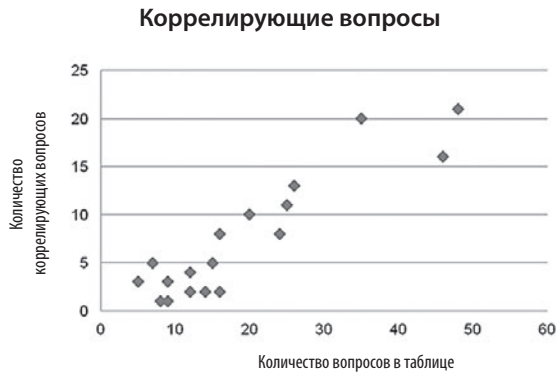
1. обеспечивает высокое качество по разумной цене;
2. превосходит мои потребности;
3. предоставляет креативные решения;
4. отвечает быстро;
5. предоставляет конкурентные цены;
6. это компания, которую я порекомендую.

Очевидная корреляция: «Обеспечивает высокое качество по разумной цене» и «Предоставляет конкурентные цены», которую вы скорее всего нашли. В данном случае корреляция действительно находится на уровне 0,68, таким образом, в анкете достаточно указать только один пункт из двух. Но если провести корреляционный анализ, можно также выяснить, что варианты «Превосходит мои потребности» и «Предоставляет креативные решения» коррелируют даже сильнее (на уровне 0,69). Мало кто сможет определить это на глаз.

Возможно, причина такой высокой корреляции состоит в абстрактности обоих утверждений. Респондентам часто бывает сложно разграничить нюансы оценки, которые мы им предлагаем, поэтому они отвечают, основываясь на своем общем представлении о предмете или испытываемых по отношению к нему чувствах.

Многие базовые оценки отношения тесно связаны с долей рынка бренда и не предоставляют практически никаких других сведений, так что мой совет: вопросы такого типа следует сводить к самым основам.

В поисках наиболее ярких примеров широкой распространенности избыточных вопросов, сотрудники нашей команды по обработке данных проанализировали 10 случайно выбранных опросов, проведенных в Азиатско-Тихоокеанском регионе. В 18 рассмотренных табличных вопросах они нашли практически линейную зависимость между количеством высказываний в таблице и уровнем избыточности вопросов. В среднем 38% вопросов в этих 10 опросах показали корреляцию более 0,65 и, таким образом, могли стать кандидатами на удаление.



Мы также выяснили, что в таблице с 10 утверждениями велика вероятность того, что хотя бы два утверждения частично совпадают по смыслу. В списке из 40 утверждений могут пересекаться до 20 из них. Поэтому при наличии в вашем опросе больших таблиц такого типа анализ, скорее всего, поможет выявить возможности для сокращения анкеты.

Такие простые тесты могут оказаться очень эффективными, но более сложные методики представят еще более важные результаты. Иногда при наличии нелинейного соотношения между двумя

вопросами коэффициент корреляции Пирсона не позволяет их выявить, в то время как кросс-корреляционный анализ трех-четырех вопросов покажет даже более сильную корреляцию.

Этот метод потребует привлечения более опытного маркетолога или статистика, а также дополнительных затрат времени и средств, но окажется гораздо более полезным при проведении масштабного трекингового исследования.

Сконцентрируйтесь на измерении статистически важных параметров

Зачастую нам приходится работать с клиентами, стремящимися включить в опрос длинные списки брендов, многими из которых пользуется только очень небольшая часть аудитории, а также длинные списки возможных причин для покупки товара, которые включают в себя множество аспектов, упоминаемых совсем небольшим числом респондентов.

Клиенты могут воспринимать такой подход как более тщательный анализ, но именно такая тщательность является причиной перегруженности анкет.

Я предлагаю использовать следующее правило: проанализируйте размер выборки и проведите статистический тест для определения уровня значимости в 95%. Любое утверждение, частота выбора которого ниже уровня статистической погрешности для показателя 95%, должно быть исключено из опроса.

Вот примерные границы возможной погрешности:

Выборка 100= +6%

Выборка 200= +5%

Выборка 400= +4%

Выборка 1000= +2%

Необходимо подумать о том, чтобы убрать из опроса любые отдельные ответы, частота выбора которых ниже этих уровней. Помните, что пороговый уровень начинается с 2%, вне зависимости от того, насколько большая у вас выборка, поскольку в любом вопросе случайно выбранные ответы могут с легкостью достичь уровня в 1%, таким образом, нет никакой необходимости фиксировать ответы с частотой появления менее 2% так как их нельзя считать достоверными. (Например, если вы опросите 500 человек на предмет

наличия у них машины марки Jaguar, которой в действительности владеют менее 1% населения, то около пяти опрошенных, у которых она есть, ответят «Да», в то время как еще пять человек отметят этот вариант «по ошибке»).

Мой совет — будьте максимально прагматичны: составьте минимально возможный список своих основных конкурентов и добавьте вариант «Другое», чтобы охватить остальные возможные ответы (и, пожалуйста, не заставляйте респондентов вписывать свои варианты ответа, поскольку это пустая трата времени для всех, и я никогда не видел, чтобы кто-то смог применить полученные при этом данные с реальной пользой).

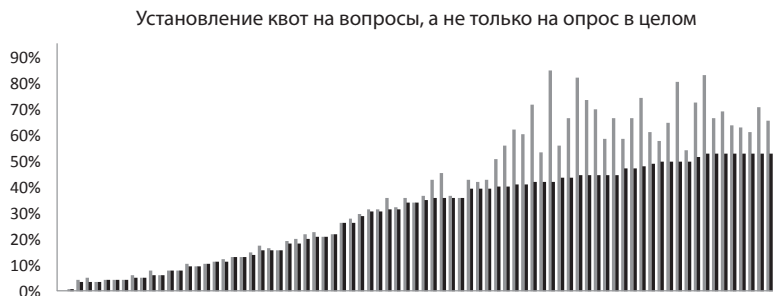
Этап 4: Установите квоты на количество ответов и вопросов

С другой стороны, некоторые варианты ответов на вопросы могут предоставить вам даже больше данных, чем вам необходимо. Простое решение, предлагаемое многими программами для проведения опросов, заключается в том, чтобы установить необходимое количество ответов по каждой теме, а затем использовать метод выбора «наименее заполненных» квот, чтобы следующие респонденты отвечали на вопросы, по которым пока меньше данных.

Установление квот: пример

Этот пример взят из реального исследования покупательского поведения. Оранжевые линии показывают, что некоторые поведенческие модели встречаются намного чаще других. Некоторые респонденты, чаще совершающие покупки, были опрошены по большому числу категорий и, таким образом, заполняли опрос практически вдвое дольше. Поэтому в следующей волне исследования (отмечена голубым) мы решили ограничить количество категорий, по которым может быть опрошен респондент. Ни одну категорию не должны были оценивать более, чем 50% выборки, при этом использование метода выбора наименее заполненной категории для ответа обеспечило равномерное распределение данных по всем категориям. Это помогло сократить длину 20-минутного опроса до 16 минут — довольно существенная экономия времени. И что, пожалуй, еще более

важно: максимальное время заполнения опроса для каждого респондента снизилось с 30 минут до 20.

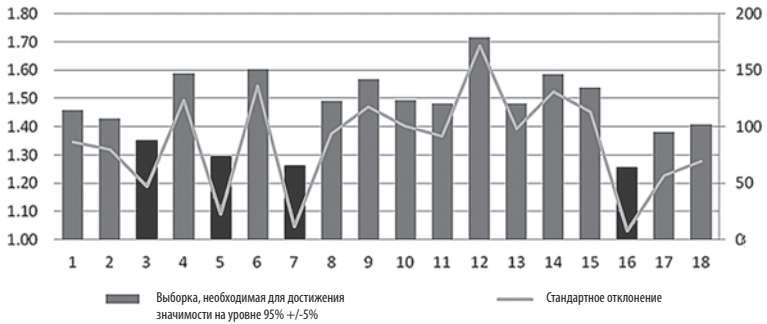


В некоторых вопросах для достижения статистической значимости необходимо меньшее количество респондентов, чем для других.

Например, бинарный вопрос, такой как: «Какая из двух реклам вам нравится больше» технически может достичь 95% уровня значимости с выборкой из пяти человек, и лишь очень редко для получения достоверного ответа вам придется опросить больше 40 человек. С другой стороны, если вы спрашиваете о привычках в получении информации из СМИ и проводите демографический анализ данных, то вам, возможно, придется опросить тысячи человек.

Практически в каждом исследовании наблюдается смешение этих двух типов вопросов: для одних требуется мало респондентов, для других — много. Более того, если вы задаете табличные вопросы, то можете обнаружить, что необходимое количество ответов для каждого также может сильно отличаться (пример приведен ниже). Все зависит от степени возможной вариации данных. Размер необходимой для каждого вопроса выборки определяется стандартным отклонением в ответах, и если вы его знаете для каждого вопроса, то можете перейти от определения размера выборки для всего исследования целиком к определению размера выборки для каждого вопроса и таким образом сэкономить.

Оценка размера выборки на основе стандартного отклонения



Установление квот по отдельным вопросам лучше всего проводить путем рандомизации (таким образом, 1 из n респондентов ответит на каждый вопрос), так как это позволит получить равномерное распределение ответов.

Используя только эту технику, нам удавалось сократить время прохождения опроса для каждого отдельного респондента более чем на 30%.

Проведите пилотное тестирование для определения размера выборки

Подсказка: Не устанавливайте окончательные параметры выборки до проведения пилотного тестирования, так как до этого вы не сможете четко определить, сколько человек вам на самом деле следует опросить для проведения исследования. Это одно из реальных преимуществ пилотного тестирования.

Пример: исследовательская компания изначально планировала опросить 800 человек, но, выяснив минимально необходимый размер выборки для получения статистически значимого результата по каждому из вопросов для тех демографических групп, которые ей было необходимо проанализировать, она пришла к выводу, что в действительности для опроса требуется опросить только 600 человек, а для 80% вопросов — всего 400 или даже меньше. В действительности после проведения пилотного тестирования на 15% вопросов уже набралось необходимое количество ответов, и потому

больше не было необходимости включать их в основную волну. В результате компания смогла уменьшить длительность опроса на 30%, а размер выборки — на 25%.

Требования к размеру выборки по вопросам



Выявление стабильных и нестабильных показателей

В трекинговых исследованиях также необходимо отслеживать, как варьируются ответы на каждый вопрос от волны к волне. Вы увидите, что некоторые из них более стабильны, чем другие, и, следовательно, можно задуматься о том, чтобы убрать их из некоторых волн, или опять же опрашивать по ним меньшее количество респондентов.

Например, в трекинговом исследовании потребления алкоголя выбор конкретного бренда напитка может сильно варьироваться от волны к волне, но ответ на вопрос о том, добавляет ли респондент в напиток лед, варьируется лишь на уровне долей процента. Таким образом, этот вопрос, вероятно, следует задавать раз в два года, чтобы отслеживать общие изменения в потреблении льда.

Опять же вы можете это сделать с помощью простого сравнения статистической разницы между двумя группами ответов — случайной выборкой из прошлой волны и пилотной выборкой. Разница между ними и будет определять количество человек, которых необходимо опросить в следующей волне. В случае с использованием льда, например, вам потребуется опросить только 10% выборки, но в вопросе на знание брендов, скорее всего, необходимо будет получить ответы всех респондентов.

Итеративный подход

Если у вас нет никаких предыдущих данных, то самый прагматичный подход заключается в установлении квот на отдельные вопросы. Просто изначально разделите выборку на группы по 100 человек, затем проанализируйте статистическую вариацию и только тогда начинайте устанавливать квоты по вопросам.

Этап 5: Примените более продвинутые методы приписывания и масштабирования

Чтобы еще больше сократить опрос, можно рассмотреть возможность использования более сложных методов.

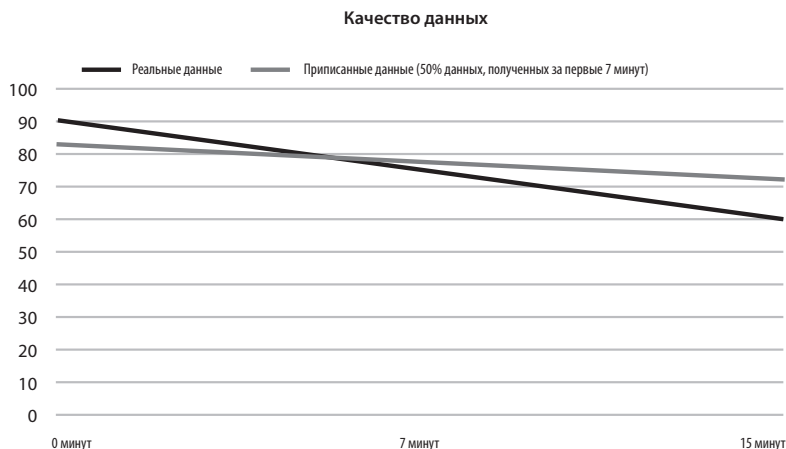
Используйте более масштабные методы приписывания / моделирования

Моделирование ответов часто воспринимается как обман, и многие клиенты весьма нервно воспринимают приписывание, считая его чем-то вроде «черной магии» или искусства создания ответов из ничего. Но это ни то, ни другое. Мой совет — не бояться этих методов: это может быть действительно эффективным способом оптимизации исследования, но, чтобы его использовать, необходимо быть действительно хорошим специалистом по маркетингу. В данной работе я не буду вдаваться в детали этого метода, просто объясню, почему вам стоит его рассматривать.

Приписывание позволяет моделировать ответы с точностью до 85%, что, как я объяснял ранее, практически настолько же хороший результат, как тот, что вы получите, опрашивая реальных респондентов.

Чтобы проиллюстрировать ценность приписывания, давайте возьмем воображаемое 15-минутное исследование и представим, что в начале опроса респонденты дают ответы по шкале Лайкерта с точностью 90%, при этом точность ответов при использовании приписывания составляет 85%. Затем точность ответов респондентов начинает снижаться. Через 7 минут после начала опроса точность падает до 80%, а в самом конце — до 60%, в то время как точность приписывания сохраняется на прежнем уровне (эти цифры, боюсь, типичны для тех моделей поведения респондентов, с которыми мы

сталкиваемся в реальных исследованиях). Таким образом, приблизительно на 6 минуте опроса точность данных, полученных с помощью приписывания, потенциально становится выше, чем точность реальных данных.



Представьте, что вы наполовину сократили опрос — до 7 минут — при этом лишь половина респондентов выборки отвечает на каждую половину опроса. Вы можете приписать ответы одной группы другой с точностью на уровне 85% первых 7 минут. Произведя несложные подсчеты, мы выясним, что 7-минутное исследование предоставляет на 4% более точные данные, чем 15-минутное полное исследование.

Данный подсчет является упрощением, поскольку точность ответов варьируется в зависимости от формата и темы. Но поскольку вы вряд ли захотите использовать метод приписывания для целого опроса, то в ситуации наличия 30 оценочных вопросов в конце опроса мой вам совет — разделить их на две группы, чтобы каждый человек отвечал только на 15. Я могу практически гарантировать, что, если вы это сделаете, полученные в итоге результаты будут не хуже — а, возможно, даже лучше — тех, что можно получить при ответе всех респондентов на все 30 вопросов.

Масштабирование опроса

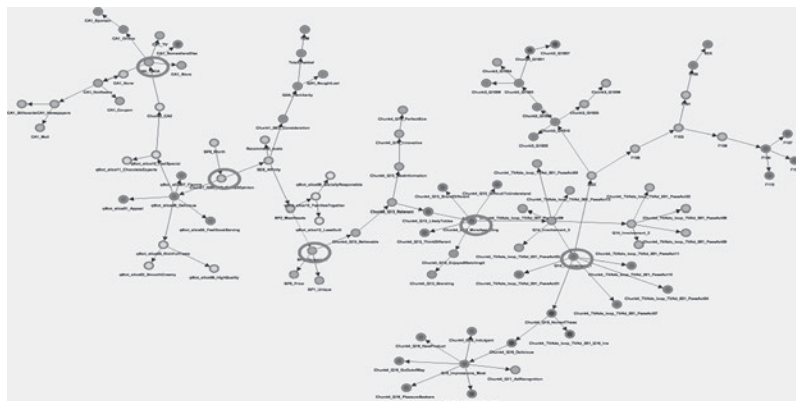
Один из методов, активно изучаемых исследователями, стремящимися сократить свои опросы, состоит в разделении опроса на блоки с тем, чтобы каждая потенциально отличная от другой группа заполняла свою часть. Это позволяет респондентам заполнять опрос и покидать его, выбирая, какую его часть они хотят заполнить, без потери вознаграждения.

Чтобы эффективно применять эту методику, необходимо понять, какие части исследования обязательно должны идти вместе, а какие можно разделить.

Например, спонтанное знание брендов и знание брендов с подсказкой разделить нельзя — эти вопросы обязательно должны идти вместе.

Чтобы все сделать правильно, необходимо использовать продвинутые статистические методы. Так же, как приписывание, это не самое простое решение, требующее помощи опытного специалиста.

Стоит признать, что я сам не до конца разобрался с тем, как применять эти методы (интересно, можно ли признаваться в таких вещах в статье для ESOMAR), поэтому я не считаю себя достаточно квалифицированным специалистом, чтобы предоставить их детальное описание. Основной принцип состоит в том, что находится неявная взаимосвязь между вопросами или группами вопросов, которые объединяются в сеть, как в предложенном ниже примере, приведенном экспертом из команды научного маркетинга компании TNS. После этого вы принимаете решение по разделению опроса на части. Как и при разделке мясной туши, «разрезы» необходимо делать по наиболее слабым соединениям, между узлами, которые четко разграничивают отдельные части тела нашего «животного».



Если вы хотите узнать об этом методе более подробно, я бы порекомендовал вам прочитать «Modular survey design: Bite Sized Chunks 2» — замечательную статью, опубликованную моим коллегой Фрэнком Келли вместе с Сэмом Стивенсом из Millward Brown и Алексом Джонсоном из Kantar в рамках конференции CASRO.

Их ранние эксперименты с этим методом вполне оптимистичны. Они обнаружили, что, предоставляя респондентам возможность самим решать, какие части опроса заполнять, исследователи мотивируют их заполнить его до конца, при этом уровень удовлетворенности от участия в подобном опросе оказывается выше. Они также выяснили, что могут использовать данные частичных заполнений анкет, которые обычно отбрасываются, и таким образом расширять объем значимой выборки до 20%.

Фрэнк Келли и его команда сейчас занимается изучением нового подхода к масштабированию, который называется «модульные взаимосвязи» и заключается в процессе рандомизации и взвешивания данных для определения комбинации двух из трех модулей в сгруппированных блоках опроса и включении в выборку ответов покинувших опрос респондентов, которые обычно не учитываются, таким образом, предлагая возможность их эффективного использования. Так как обычно вопросы сгруппированы в логические блоки или модули, каждому из них можно назначить определенный размер выборки, используя методы, описанные мной выше. Таким образом,

в рамках одного опроса для одного модуля может потребоваться опросить 1000 человек, для другого — 750, а для третьего — 500. Следовательно, можно сократить, например, 24-минутную анкету с показателем покинувших опрос респондентов на уровне 40% до 15-минутной и получить 2250 полностью заполненных модулей вместо 1000 полностью заполненных опросов, включающих в себя все три модуля. Респонденты, покинувшие опрос раньше времени, также получают вознаграждение за те блоки, которые они заполнили. Его статья на эту тему недавно была опубликована в рамках Конференции цифровых исследований Castro (11–12 марта 2014 г.).

ЧАСТЬ 2: Разработка и подготовка опроса в стиле бонсай

Сократив опрос статистическими методами, можно перейти к следующей стадии процесса его разработки, в рамках которого также можно добиться необходимого сокращения затраченного времени и усилий респондентов.

Снова выделим пять основных этапов:

1. консолидируйте и сгруппируйте вопросы, используя гнездовые методы суб-вопросов и «панели»;
2. меняйте вопросы;
3. задавайте более умные и интересные вопросы;
4. выберите подход к разработке дизайна исследования;
5. эргономичность дизайна опроса.

Шаг 1: Консолидирование и группировка вопросов

Первое, что следует проанализировать,— это возможность консолидирования и группировки вопросов.

Один из наиболее разочаровывающих респондента моментов в опросе с плохо организованной структурой — это ситуация, когда его просят поразмыслить над какой-либо темой или проблемой, затем переходят к другой, а после этого снова возвращаются к первой. Это заставляет их вспоминать и отставлять в сторону полученную информацию — словно мысленно открывать и закрывать разные ящики стола снова и снова — что требует времени и ментальных

усилий. Необходимо спроектировать опрос таким образом, чтобы каждый такой «ящик» можно было открыть только единожды.

Например, в нашем исследовании мы хотим узнать, какие бренды респонденты знают, любят и используют.

В стандартном исследовании сначала выясняют степень знакомства респондента с каждым из брендов по очереди, затем спрашивают, нравится ли ему каждый из них и использует ли он их продукцию, также представляя их по очереди.

Теоретически более эффективным является комбинирование этих трех вопросов в один простой мыслительный процесс. Возьмем пример из опроса, посвященного книгам. Изначально анкета включала в себя 3 разных вопроса:

V1. Про какие из этих книг Вы слышали?

Не слышал (а)

Слышал (а)

Читал (а)

V2. (Спрашивать про те книги, о которых респондент слышал)

Какие из этих книг Вы хотели бы прочитать?

Хотел (а) бы прочитать

Не хотел (а) бы прочитать

V3. (Спрашивать про те книги, которые респондент читал)

Насколько Вам понравилась эта книга?

Не понравилась

Понравилась

Очень понравилась

Используя комбинирование, все эти три вопроса можно задать за один раз:

Не слышал (-а) об этой книге	Слышал (-а), но не читал (-а) и не хоте- лось бы прочесть	Слышал (-а), но не читал (-а), хотя хотелось бы прочесть	Слышал (-а), читал (-а), но не понравилась	Слышал (-а), читал (-а) и понравилась	Слышал (-а), читал (-а) и очень по- нравилась
---------------------------------	---	--	---	---	--

Есть у этого подхода большая проблема. Респондентам не нравится отвечать на вопросы с такими длинными списками ответов, поскольку им либо приходится удерживать в голове большой объем информации, либо перечитывать вопрос несколько раз.

Решение этой проблемы состоит в том, чтобы показать варианты выбора с помощью эффектной иконографии. Мы часто видим, что иконки используются в исследованиях исключительно декоративно, но в этом типе вопросов они могут быть функциональными.

					
Не слышал (-а) об этой книге	Не хотелось бы прочесть	Хотелось бы прочесть	Читал (-а), но не понравилась	Понравилась	Очень понравилась

Хорошая иконка при взгляде на нее действует как быстрое визуальное напоминание, на обработку которого требуется лишь доля секунды, и респондентам не приходится запоминать все варианты ответа. Иконки снижают нагрузку на память и делают вопрос проще. Таким образом, получается вопрос, который содержит в себе данные предыдущих трех и требует лишь немного больших усилий со стороны респондента, чем каждый из них в отдельности (заметьте, что благодаря иконкам нет необходимости повторять фразы «слышал ли» и «читал ли»).

Разумно комбинируйте варианты ответа

Одна из вещей, которых следует избегать при разработке опроса для любой платформы, это длинные списки брендов, а по сути, любые длинные списки вообще. Если такой список встречается в вашем опросе, я бы порекомендовал использовать метод разумной комбинации, который означает разбивку длинных списков на вопросы с первичным выбором и последующие вопросы с дальнейшим выбором, где это возможно. Я советую делать это не только из-за требований к опросам в стиле бонсай, но также и потому, что это позволит улучшить полученные данные.

Вот пример из трекингового исследования по напиткам, который проиллюстрирует мою идею:

Изначально вопрос был задан так:

Какие типы напитков Вы употребляли за последние 24 часа?

Среднегазированные безалкогольные напитки
Слабогазированные безалкогольные напитки
Соки / Нектары /
Сокосодержащие напитки
Овощные соки / сокосодержащие напитки
Водопроводная вода
Бутилированная вода — негазированная, не ароматизированная
Бутилированная вода — газированная, не ароматизированная
Бутилированная вода — негазированная, ароматизированная
Бутилированная вода — газированная, ароматизированная
Питьевая вода в больших бутылках (5л и больше)
Улучшенная вода (например, витаминизированная вода)
Спортивные напитки
Энергетические напитки
Чай — черный
Чай — зеленый
Кофе
Обычное молоко
Ароматизированное молоко
Йогуртовые напитки
Смузи
Напитки на основе сои
Другие безалкогольные напитки

Для респондента работа с такими списками представляет определенные трудности. Кроме того, этот список слишком длинный, и на мобильных телефонах его приходится прокручивать, что также является проблемой, о которой я расскажу чуть позже. Такие списки обычно читают невнимательно, и в результате мы получаем некачественные данные.

Треть списка занимают различные виды бутилированной воды. В этом примере 10% респондентов отметили один из вариантов бутилированной воды, но меньше 1% отметили каждый из ее видов (то есть это меньше чем упомянутый выше процент случайного заполнения).

После изменения с добавлением подвопроса по видам бутилированной воды вопрос выглядит так:

Какие типы напитков Вы употребляли за последние 24 часа?

Среднегазированные безалкогольные напитки	
Слабогазированные безалкогольные напитки	
Соки / Нектары/	
Сокосодержащие напитки	
Овощные соки /	
сокосодержащие напитки	
Водопроводная вода	
Бутилированная вода	→ Подвопрос
Спортивные напитки	
Энергетические напитки	Негазированная
Чай — черный	Газированная
Чай — зеленый	Ароматизированная —
Кофе	негазированная
Обычное молоко	Ароматизированная —
Ароматизированное молоко	газированная
Йогуртовые напитки	Питьевая вода в больших
Смузи	бутылках (5л и больше)
Напитки на основе сои	Улучшенная вода (например,
Другие безалкогольные напитки	витаминизированная вода)

Этот вопрос позволил освободить 90% опрошенных — тех, кто не употреблял бутилированную воду — от необходимости прочтения длинного подробного списка и лучше вписался в формат мобильных опросов.

10% респондентов, которые выбрали вариант «бутилированная вода», можно задать дополнительный вопрос про ее вид и предоставить им выбор из подсписка.

Этап 2: Меняйте вопросы

Другие способы сделать длинные табличные вопросы короче

Кроме статистических методов существует множество способов сокращения усилий при ответе на вопросы путем простого изменения их формулировки.

Этот способ, вероятно, лучше всего продемонстрировать, представив пример реального вопроса из исследования и проанализировав разные способы его формулирования, отмечая при этом возможности для сокращения времени опроса и повышения его эффективности.

Вот пример довольно типичного опроса, который вам, скорее всего, захочется сократить:

Используя шкалу от 1 до 10 оцените, насколько каждый из следующих 15 факторов важен для Вас при выборе мобильного телефона?

Дизайн

Функциональность

Качество динамиков

Качество камеры

Разрешение экрана

Размер экрана

Вес

Время расходования заряда батареи

Операционная система

Поставщик мобильного телефона

Объем памяти

Количество бесплатных СМС

Количество бесплатных минут

Объем данных

Это реальный вопрос из реального исследования. Ответ на него занимает около 37 секунд. В данном случае проблема состояла в том, что полученные в итоге данные были недостаточно дифференцированы, поскольку большинство респондентов считают все эти элементы своих мобильных телефонов важными, и мы получаем ответы

по каждому из них на уровне 6–7 баллов. Мы подсчитали, что для определения пяти самых важных факторов нам необходимо опросить более 500 респондентов, чтобы выявить статистически значимые различия, а для оценки всех факторов нам необходимо более 2000 респондентов, что превышает размеры выборки большинства исследований. Можно ли найти более короткий и эффективный способ сбора той же информации?

Задавайте предварительные вопросы для сокращения списка: Хороший способ сократить объем работы для респондента — уменьшить список, задав предварительный вопрос-фильтр, в ходе которого будут отобраны только самые важные факторы, таким образом, респонденту не придется оценивать все 15. Количество выбранных ответов зависит от выборки. Я бы старался избегать вопросов с просьбой выделить самые важные ответы, поскольку респонденты обычно ограничиваются одним или двумя вариантами. Затем, если понадобится, вы сможете задать вопрос о том, насколько каждый из этих элементов важен. Так экономится около 10 секунд, и размер выборки требуется также меньший — около 400 респондентов будет достаточно, чтобы выделить топ-5 факторов.

Используйте выборочное ранжирование: Если вы хотите еще больше упростить вопрос, попросите не выбрать, а проранжировать пять самых важных факторов. Вам не придется спрашивать про важность каждого из них, и это потребует приблизительно столько же усилий, сколько и выбор. Экономия времени — 15 секунд, минимальный размер выборки — 300 человек.

Максимальная дифференциация на основе выбора: Другой способ дифференциации ответов состоит в том, чтобы показывать по пять факторов и просить из каждого набора выбрать наиболее и наименее важный. Эта техника предоставляет качественные данные, не зависящие от кросс-культурных особенностей, с которыми вы сталкиваетесь, анализируя данные оценок с использованием шкалы на разных рынках. Экономия времени — 15 секунд, минимальный размер выборки снова 300.

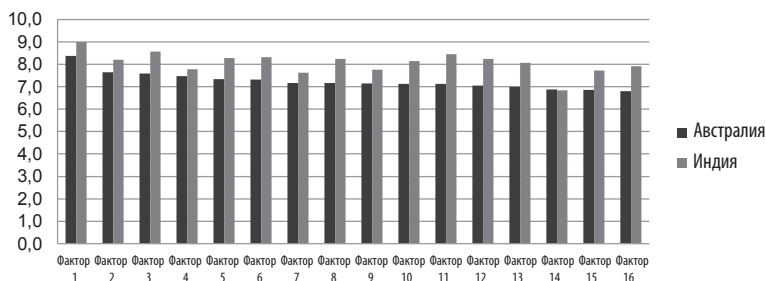
Используя все эти три метода вы, с одной стороны, да, теряете часть информации о каждом факторе, но я считаю, что эта информация на самом деле оказывается не слишком полезной в связи с тем,

что итоговые оценки могут отличаться минимально, но при этом требовать больших объемов выборки. Становясь более прагматичным в своем отношении к информации, которую вы собираете, вы можете сэкономить время и уменьшить размер выборки.

Чтобы акцентировать ваше внимание на различиях этих подходов, я представляю вам следующие графики, демонстрирующие довольно ощутимую разницу, которую вы можете иногда увидеть при переходе от метода выставления оценки к методу максимальной дифференциации на основе выбора. Данные были собраны во время реального эксперимента на двух разных рынках (к сожалению, я не могу сообщить более точные данные в связи с конфиденциальностью этой информации).

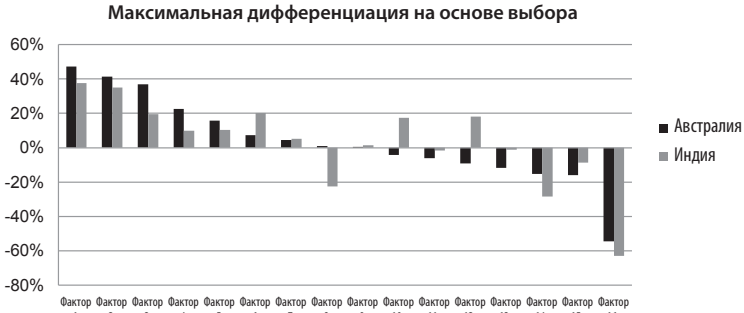
Первый график показывает результат применения традиционного подхода, в котором каждый фактор получил оценку от 7 до 7,5, а кроме того, проявилась классическая проблема интерпретации культурных различий, состоящих в том, как респонденты отвечают на вопросы (индийские респонденты обычно отвечают более позитивно), что усложняет сравнение результатов между странами. Чтобы эти различия были значимыми, нужна выборка более 1000 респондентов.

Традиционная оценка по 10-балльной шкале



График, приведенный ниже, показывает ответы на тот же вопрос, но с использованием метода максимальной дифференциации. Здесь вы сможете увидеть разграничения между факторами более наглядно и проследить четкие различия между двумя странами, ко-

торые не обусловлены культурой. Эта техника также требует от респондентов меньше времени, и размер выборки может быть менее 300 человек.



Примечание: Вы можете подумать, что так мы экономим лишь несколько секунд — в чем смысл? На самом деле, экономя 20 секунд здесь и 30 секунд где-то еще, а также используя другие методы выбора вопросов, можно сократить опрос на 3–4 минуты. А когда вы увидите, что показатель покинувших опрос респондентов, которые заполняют опрос с мобильного, увеличивается примерно на 2–3% в минуту, то поймете, что 4-минутная экономия позволяет повысить показатель полного завершения опроса на 10%, а этим «уже не следует пренебрегать», как говорят у нас в Великобритании.

Этап 3: Задавайте более умные и интересные вопросы

Разработка опроса в формате бонсай не ограничивается его сокращением, я бы хотел, чтобы вы также поразмыслили над тем, как задать более умные и интересные вопросы.

Существует большое количество инновационных и увлекательных способов формулировки вопросов, основанных на принципах поведенческих наук, геймификации и работы рынков прогнозов, благодаря которым прохождение опроса становится более интересным и приятным времяпровождением для респондента.

Поскольку я уже написал несколько работ по этой увлекательной теме, я не буду останавливаться на ней в данной статье. Впрочем, в описании примера SingTel в конце данной статьи я еще останавлиюсь на том, как можно применить эти методы в реальных исследованиях.

Этап 4: Выберите подход к разработке дизайна исследования

Теперь мы переходим непосредственно к процессу разработки дизайна исследования.

Первое решение, которое вам необходимо принять при проектировании исследования для нескольких различных платформ и устройств,— это выбор минимального стандарта. Ключевым в данном случае является вопрос о том, что делать с обычными мобильными телефонами?

На многих рынках обычные мобильные телефоны до сих пор являются доминирующим форматом устройств. Проблема заключается в том, что в 95% исследований нашим панелистам предлагается заполнить слишком длинный и сложный для простого мобильного телефона опрос. И тех, кто старается заполнить его до конца, очень мало, доля покинувших опрос респондентов составляет до 80%.



Мой совет заключается в том, что при исследовании рынка, аудиторию которого вы хотите охватить с помощью телефона, вам необходимо выделить его в отдельный проект и дополнительно сократить. Для успешного проведения опроса на мобильном телефоне необходимо уменьшить число вопросов примерно до 15. Конечно, невозможно проводить эффективные исследования объемом всего в 15 вопросов, и для их создания потребуется настоящий мастер бонсай-исследований. Если вы хотите поучиться у настоящих гуров бонсай-исследований, я вам рекомендую посмотреть на некоторые работы Дженни Хофмайер и команды TNS в Южной Африке. Они задают только самые необходимые и важные вопросы, сокращая огромные трекинговые исследования по брендам до 8 вопросов.

Для остальных мой совет: просто забыть об обычных телефонах как о платформе для стандартных опросов и сфокусироваться на смартфонах, поскольку на большинстве рынков именно они обычно используются для проведения исследований.

Сейчас смартфоны имеют разный размер и форму. Минимальный размер экрана, под который мы рекомендуем проектировать исследование, составляет 800 x 480 пикселей или более. В настоящее время это минимальный размер экрана 85% смартфонов, и этот показатель имеет тенденцию к увеличению.

Книжная или альбомная ориентация?

Следующее, на что следует обратить внимание: исследования для ПК обычно выполнены в альбомной ориентации, в то время как для мобильных телефонов лучше подходит книжный формат.

Респонденты предпочитают использовать мобильный телефон в книжном формате, и в некоторых вопросах, например, с длинными списками, на мобильных телефонах приходится прокручивать их намного меньше.



Проблема, однако, состоит в том, что большинство опросов, которые мы создаем, оказываются слишком широкими по формату, и многие вопросы просто не подходят для демонстрации на экране мобильного телефона в книжном режиме, как видно из примера ниже.



К тому же, с прагматической точки зрения, так как со смартфонов и планшетов заполняется меньшая часть анкет, то и адаптировать все под книжный формат кажется неразумно.



Два решения: адаптивный дизайн или фиксированный альбомный формат

Нам необходимо выбрать из двух вариантов: 1. адаптивный дизайн, то есть адаптация формата опроса под мобильные телефоны и использование разных версий опроса для ПК и планшетов; 2. подход «один размер для всех» с использованием фиксированного альбомного формата.

Выбор варианта зависит преимущественно от используемого метода проведения опроса, его типа и сложности. Кроме того, у каждого из них есть свои плюсы и минусы.

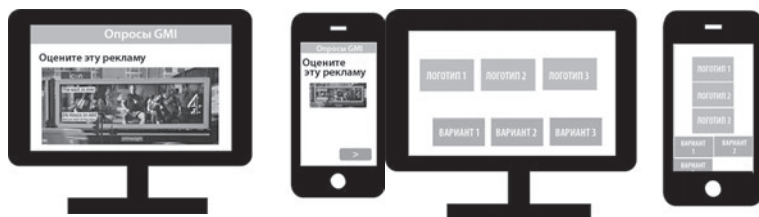
Адаптивный дизайн: за и против

Достоинством адаптивного дизайна является то, что респондентам не обязательно поворачивать телефон горизонтально, и они могут заполнять опрос одной рукой. В этом случае требуется меньше прокручивать списки вниз, и даже размещение кнопки «Далее» внизу страницы является более интуитивно понятным и эффективным. К тому же ПК-версия опроса никак не меняется в случае адаптации его к мобильному формату.

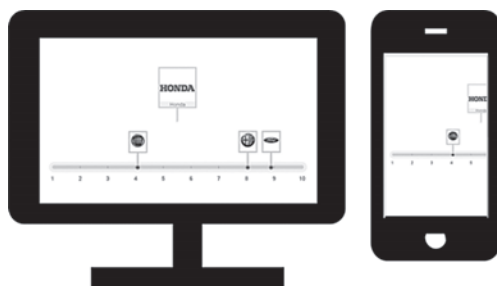
Для упрощения адаптации многие компании, создающие ПО, разработали сложные, кастомизированные решения для дизайна опросов.

Однако есть и недостатки. Некоторые вопросы невозможно полностью адаптировать под книжный формат, так как определенный контент, например, реклама, в ширину занимает больше места, чем

в высоту. Адаптивный дизайн также может оказаться неэффективным, когда дело доходит до форматирования: некоторые вопросы могут выглядеть неаккуратно на мобильном телефоне и не умещаться на экране. Это зачастую приводит к необходимости проведения двух опросов вместо одного с последующей проверкой.



Второе, что следует помнить: многие наиболее успешные сравнительные методики, используемые в маркетинговых исследованиях, предполагают проведение сравнения на одном экране по горизонтали, так как более естественным для человека является просмотр по горизонтали, чем сверху вниз, и, как обнаружили многие разработчики мобильных приложений, горизонтальное расположение является более эффективным.



Фиксированный альбомный формат: за и против

Итак, второй вариант — выбрать горизонтальный формат для всех платформ и просить респондентов повернуть телефоны.



Преимуществом этого подхода является то, что создавать такой дизайн намного проще — он может быть практически одинаковым для мобильных телефонов и компьютеров, и, кроме того, у вас больше возможностей использовать более креативные форматы вопросов.

Недостатком, помимо неудобства поворота телефона, является то, что все равно придется хотя бы минимально адаптировать весь дизайн опроса для мобильного телефона и, возможно, даже пересмотреть подход к разработке опроса в целом.

Итак, что вы выберете?

Нам, как стороне, предоставляющей услуги по разработке опросов и обладающей незначительным контролем над контентом исследования, достаточно сложно использовать адаптивный дизайн. Обычно в исследовании, где нас просят адаптировать дизайн, встречается один-два вопроса, которые не будут работать в книжном режиме, и мы останавливаемся на фиксированном альбомном варианте.

Однако если у вас есть возможность повлиять на содержание исследования и его дизайн, чтобы убедиться, что все вопросы можно адаптировать и перенести на другую платформу, и при этом вы используете правильное программное обеспечение, я бы порекомендовал выбрать адаптивный вариант, который оставит лучшее впечатление у респондентов.

Разработка опросов единого формата, без различий между платформами

Если вы размышляете над дизайном опросов, подходящим для всех платформ, вот некоторые советы для вас:

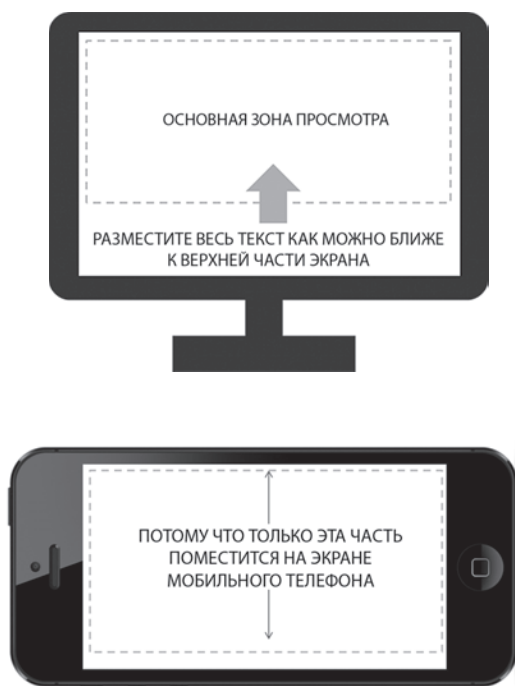
1. Помните, что важен каждый пиксель!

2. Уменьшите общий масштаб и обратите внимание на соотношение сторон, а не пикселей;
3. помните об «основной зоне просмотра»;
4. любыми способами избегайте необходимости прокручивания списков;
5. сделайте вопросы как можно более интерактивными, если вам позволяют это имеющиеся технологии.

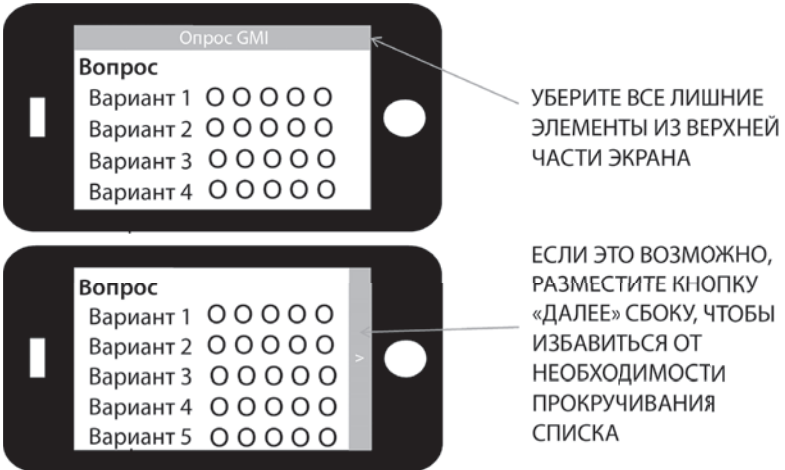
Основная зона просмотра

Самое важное, что надо учитывать при разработке дизайна, подходящего для всех платформ,— это существование «основной зоны просмотра», расположенной в верхней части экрана.

Постарайтесь расположить в этой зоне как можно больше контента, поскольку именно ее будет видно на экране мобильного телефона.



Это может привести к изменению шаблона страниц опроса. Многие традиционные онлайн-опросы предусматривают размещение логотипа компании вверху страницы, а это уменьшает ценное пространство основной зоны просмотра.



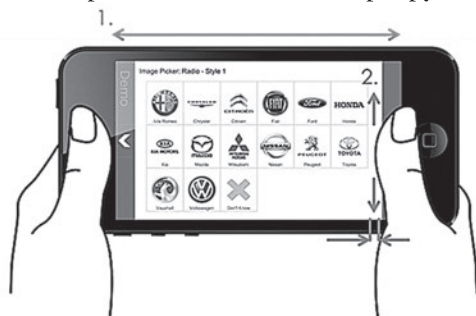
Помните, что важен каждый пиксель

Следующий совет заключается в минимизации расстояния между ответами, в этом случае в видимую зону можно будет уместить максимальное количество вариантов.

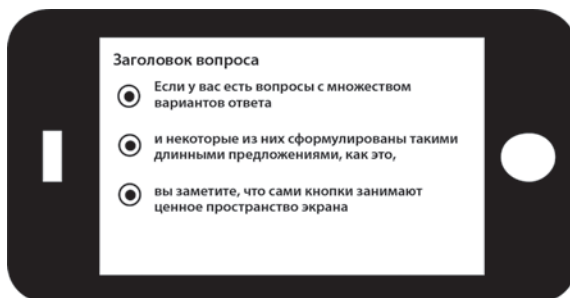


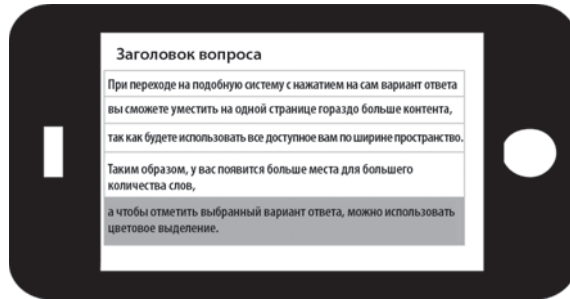


Максимизируйте используемую площадь экрана, но если вопрос предполагает необходимость прокручивания списка, оставьте немного места с правой стороны, так как иначе респондент может случайно нажать на вариант ответа, пытаясь прокрутить список вниз.



Если технологические возможности позволяют, я бы порекомендовал перейти от так называемого «традиционного» вида кнопок выбора ответа, которые занимают дополнительное место, к варианту с возможностью нажатия на сам ответ.





Используйте автоматический переход, где это возможно

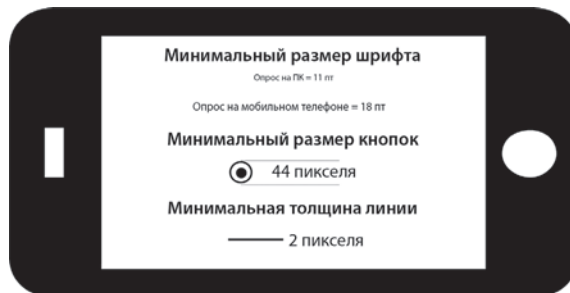
Лучше избегать ситуаций, в которых респонденты вынуждены прокручивать страницу вниз для поиска кнопки «Далее», поскольку в некоторых приложениях появляется панель управления, что еще сильнее уменьшает имеющееся пространство.

Увеличенный размер шрифта и минимальный размер кнопок

Мы рекомендуем увеличить размер шрифта в тексте вопросов до 18 пт и даже больше, если позволяет пространство.

Все опции и кнопки ответа также должны быть увеличены до высоты минимум 44 пикселей (это стандарт компании Apple).

Линии также должны быть толще, если вы планируете их использовать. 2 пикселя — это минимум, поскольку в современных смартфонах плотность пикселей очень высокая и иногда однопиксельные элементы просто исчезают.



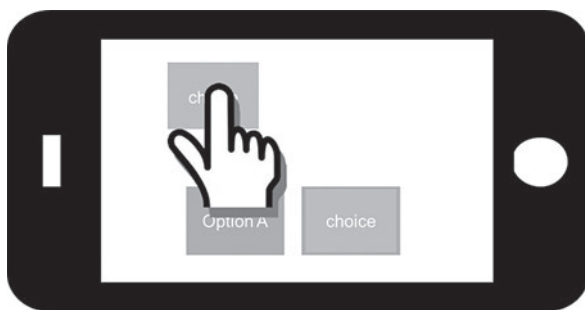
Используйте интерактивный формат вопросов

Интерактивные вопросы могут рассматриваться как более интересные в обычном формате опроса и намного более ценные в опросах, подходящих для всех платформ.

Динамичные таблицы, в которых после того, как вы выбираете ответ для одного высказывания, оно динамично изменяется на другое высказывание, оказываются замечательным вариантом оформления вопроса для мобильного формата, в то время как традиционные таблицы просто не помещаются на экран.



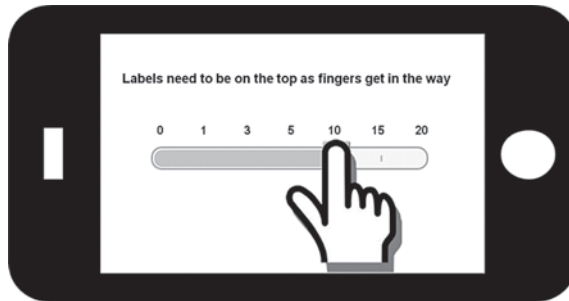
Формат «потяни и отпусти» намного более интуитивно понятен на сенсорном экране, поэтому такой вариант представления вопроса оказывается очень полезным.



Так как все большее число этих форматов становятся доступны в HTML, я настоятельно рекомендую включать их в опросы.

Измените схему расположения элементов

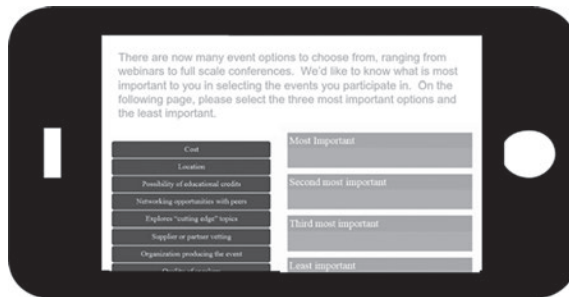
Существует еще множество более мелких деталей, которые также необходимо продумать. Например, следует убедиться, что вы поместили открытый вопрос настолько высоко на экране, насколько это возможно, оставляя место для вывода клавиатуры. Обозначения должны располагаться над шкалой, а не под ней, чтобы респондент не закрывал их пальцем.



Обозначения должны располагаться над шкалой, чтобы респондент не закрывал их пальцем

Длинные инструкции и тексты вопросов иногда лучше размещать на отдельном экране, чтобы они не занимали ценное место.

Сравните:



И

There are now many event options to choose from, ranging from webinars to full scale conferences.

We'd like to know what is most important to you in selecting the events you participate in.

On the following page, please select the three most important options and the least important.

Инструкции

Please select the three most important options and the least important for your conference

Cost
Location
Possibility of educational credits
Networking opportunities with peers
Expenses "cutting edge" topics
Supplier or partner setting
Organization producing the event
Quality of speakers
Interest in speakers
Focuses on issues important to my business
Other options

Most important

Second most important

Third most important

Least important

Описание задания

Кроме того, постарайтесь сформулировать каждый вопрос максимально кратко — как в Твиттере. Таким образом можно сэкономить много места.

Не стесняйтесь использовать иконки и визуализацию там, где есть для этого место. Они не только могут оживить опрос, но и играют важную роль в объяснении различных вариантов ответов, как уже было показано ранее.

What kinds of research does your organization provide to clients or purchase from others?
You may tick more than one

Quantitative data gathering and analysis	Qualitative data gathering and analysis	Social media monitoring and analysis	Secondary research or meta-analysis	Measurement: e.g. Media, Retail

Лучше уменьшайте масштаб, а не увеличивайте

Помните, что респонденты проходят опросы и на планшетах, у которых экран может быть больше. Когда вы программируете опрос, лучше предусмотреть возможность уменьшения, а не увеличения масштаба.



Текущий стандарт дизайна в размере 1024 x 600

Наша рекомендация на сегодняшний день состоит в том, что основная ширина опросного экрана должна составлять около 1024 пикселей, при этом в длину хорошо видимая зона занимает около 600 пикселей. Однако с увеличением экрана этот показатель меняется.

Когда вы программируете на HTML, заложите возможность масштабирования по ширине, так, чтобы она соответствовала ширине экрана мобильного телефона — и это максимизирует использование доступной площади. На планшетах и компьютерах этот формат не будет полностью соответствовать размерам экрана, однако размер опросной страницы окажется вполне приемлемого размера (см. фото ниже).



Найдите программу для тестирования дизайна опроса

Вам также понадобится программа для тестирования дизайна опроса, подобная той, речь о которой пойдет ниже. Я боюсь, что это немного дороже, чем очки респондента, которые я также советовал вам приобрести, но это необходимо при проектировании дизайна опроса, подходящего для любого устройства.



Анализ примера: Создание опроса в стиле бонсай для исследования привычек и предпочтений потребителей для Singtel

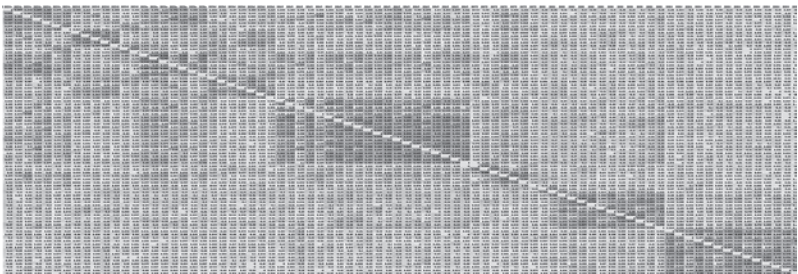
Сингапурская телекоммуникационная компания Singtel решила нам проанализировать методы создания бонсай-опроса на примере своих исследовательских проектов.

Исследование, которое они попросили нас провести, является типичным 25-минутным опросом на тему знания и отношения к мобильным телефонам. Исследование включало в себя множество вопросов о конкурентных сервисах и тарифах. Большой блок вопросов был посвящен выявлению используемых функций мобильных телефонов, включая различные типы приложений, и сравнению вариантов использования телефона по сравнению с компьютерами и планшетами. Также он включал в себя вопросы, посвященные определению поведенческих моделей.

Первым шагом было создание рабочей версии опроса и проведение пилотного исследования с участием 100 респондентов. Мы не задавали им демографических вопросов и несколько основных вопросов, которые, как мы знали, не изменятся.

Уже на стадии пилотного тестирования мы выявили несколько возможностей для сокращения опроса.

После проведения корреляционного анализа Пирсона мы обнаружили 24 пересекающихся вопроса (со значением более 0,65), которые мы сгруппировали. Таким образом мы сэкономили 1 минуту. Схема показывает, как выглядит результат теста: красным выделены заметно коррелирующие вопросы.



Мы провели анализ объема выборки для отдельных вопросов и обнаружили, что максимальный объем выборки составляет 1000 респондентов, если мы хотим разделить данные на четыре части для демографического анализа, но, так как только для 10% вопросов требовалась такая большая выборка, мы решили, что можно сократить ее размер до 800 человек. Как вы видите, оставшиеся вопросы требовали размера выборки в 400 человек. Таким образом, меняя порядок предъявления секций A, C, E, G & I и используя квоты на вопросы, мы сможем сократить исследование в среднем, на 8 минут.



В целом в этом 25-минутном исследовании, по нашим оценкам, объем выборки можно сократить на 25% и уменьшить длину до 15 минут без значительной потери данных.

Дизайн опроса

Как в большинстве исследований привычек и предпочтений потребителей, одна из основных проблем состояла в повторяющемся характере вопросов. И это стало очевидно, когда мы стали изучать данные и выявили снижение внимания респондентов во многих важных вопросах. На графике внизу приведены данные по вопросам о том, как используется мобильный телефон, компьютер и планшет.



Итак, что мы можем сделать, чтобы исправить эту ситуацию? Наш подход в подобной ситуации состоит в том, чтобы сделать участие в исследовании для респондента более интересным и увлекательным.

Применение методики актуализации вопроса

В предыдущих статьях я говорил о концепции актуализации вопроса, состоящей в том, что респонденты готовы приложить больше усилий, если они чувствуют, что существует реальная ясная цель у того, что они делают.

Наш подход в данном исследовании состоял в том, чтобы попросить респондентов принять участие в специальном опросе, чтобы помочь предугадать, как изменятся привычки использования мобильных телефонов в будущем. Мы сочли, что это более привлекательная для вовлечения респондентов тема, и построили вопросы на ее основе.

Which type of service do you think will grow most next year?

Pre-paid Post-paid

The market agrees with you.
You win a point!

How often do you use your mobile phone for... I predict usage will...

Activity	Frequency	Prediction
Taking photos / videos	2 or 3 times a day	1 2 3 4 5
Using GPS / navigation / location-based services	Once a month	1 2 3 4 5
Playing games	Once a day	1 2 3 4 5
Social networking	Once a month	1 2 3 4 5
Listening to music	More than once an hour	1 2 3 4 5
Sending / receiving SMS	Once a week	1 2 3 4 5

Well done! You have got 1 out of 5 predictions correct.
You win a star!

В исследовании мы попросили респондентов спрогнозировать, как изменятся привычки использования мобильных телефонов в будущем и сравнили их ответы с ответами остальных респондентов. Мы сделали это в формате игры: если они угадывали то, что остальные респонденты называли возможным, то получали баллы по завершению каждой секции.

Элементы дизайна

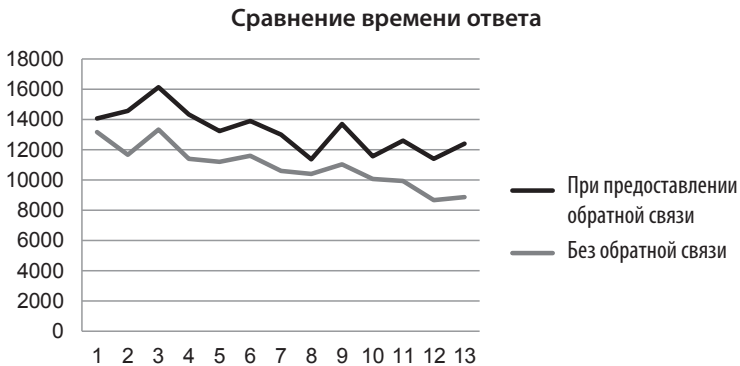
Что касается дизайна, то мы выбрали фиксированный альбомный вариант и максимально использовали все креативные форматы вопросов, которые были нам доступны. Мы отредактировали все вопросы, убедились, что они помещаются в зону видимости и что ни в каких вопросах респондентам не придется прокручивать страницы вниз на экране мобильного телефона.

В ходе пилотного тестирования мы заметили, что значительная часть респондентов прекращает заполнять опрос уже в самом начале, поэтому мы изменили приветствие, сделав его более радужным. Взгляните на пример, приведенный ниже: первый вопрос был сформулирован, чтобы привлечь внимание и заинтересовать.



Чтобы протестировать эффективность этой концепции, мы провели второй этап пилотного тестирования, используя разные выборки, и одну группу респондентов попросили сделать предсказание без предоставления информации о его правильности, а другим рассказывали о полученных ранее результатах и награждали звездочками за правильный выбор варианта.

Этот способ предоставления обратной связи действительно помог заинтересовать респондентов темой опроса, время ответа на вопросы увеличилось в среднем на 18% (что может показаться не особо заслуживающим внимания, однако следует учитывать, что 2/3 времени уходит только на то, чтобы прочесть вопрос, и, таким образом, увеличение на 18% говорит о том, что практически вдвое увеличилось время размышления над каждым ответом).



Также мы наблюдали статистически значимое улучшение качества ответов и практически 50%-ное снижение частоты шаблонных ответов по сравнению со стандартным отклонением на уровне респондентов.

Теперь время заполнения увеличилось на 3 минуты, что в некоторых случаях может восприниматься негативно, но не в данном случае, потому что на самом деле больше респондентов завершили опрос полностью (на 5%), а их удовлетворенность повысилась на 20%. Лингвистический анализ отзывов об опросе выявил большой

энтузиазм по отношению к механизмам вознаграждения, и респонденты посвящали опросу больше времени, потому что они хотели этого и были мотивированы.

Содержание

Надеюсь, эта статья помогла вам познакомиться с некоторыми базовыми методиками, которые позволят вам проектировать исследования в стиле бонсай.

Я понимаю, что статья оказалась длинной, поэтому вот ее краткое содержание:

1. думайте как респондент;
2. применяйте итеративный подход;
3. убирайте лишние вопросы;
4. применяйте квоты на вопросы и блоки вопросов;
5. не бойтесь пользоваться методиками моделирования и приписывания;
6. комбинируйте вопросы, используя иконографику;
7. формулируйте более короткие, но эффективные вопросы;
8. относитесь к респондентам как к думающим машинам и задавайте более умные и интересные вопросы;
9. используйте открытые вопросы;
10. когда программируете опрос, не забывайте о зонах видимости и важности каждого пикселя.

Я понимаю, что этот процесс не всегда оказывается простым, и одна из основных сложностей, с которой вы при этом столкнетесь, будет состоять в том, чтобы убедить клиента или собственный маркетинговый отдел, что такого рода изменения необходимы. Удачи вам в этом!

Не всегда короче значит лучше

*Инна Бурдейн
(The NPD Group)*

Поиски рекомендаций для подготовки онлайн-исследований на различных веб-сайтах неизбежно приводят к рассуждениям о длительности опроса. Наиболее распространенное мнение — опрос должен быть коротким. Определение «короткого» опроса может варьироваться, но в целом считается, что чем длиннее опрос, тем заметнее снижается качество данных, показатели числа завершенных опросов и удовлетворенности респондентов. В последнее время из-за мобильной революции специалисты уделяют больше внимания, чем когда-либо прежде, преобразованию длинных опросов в короткие и простые для того, чтобы респонденты не устали. Это стремление к более коротким опросам должно основываться не только на предположениях, но и на результатах реальных исследований. Существует общее мнение, что чем короче опрос, тем лучше, т.к. время у людей ограничено, а заполнение длинных анкет утомительно.

Некоторые исследования, безусловно, подтверждают этот факт. Исследование 2004 года, проведенное Ратодом и Бруной, а затем продублированное в 2009 году компанией SSI, подтвердило, что качество данных ухудшается при увеличении длительности опроса. Фреде (2010) обнаружил линейную зависимость между временем заполнения опроса и показателем числа респондентов, покинувших опрос до его завершения. Галесик и Босняк (2003, 2007) выявили аналогичную линейную зависимость, добавив, что разнообразие ответов в таблицах тем ниже, чем дальше расположен вопрос.

Интересно отметить, что исследования по определению оптимальной длины интервью дают довольно большой разброс вариантов. Компания Gallup предлагает не превышать 5 минут. В Survey Monkey добавляют, что опросы длительностью от 6 до 10 минут

вполне приемлемы, но в анкете длиной более 11 минут будет выявлено много искажений данных. Чадоба (2011) обнаружил, что отказы от дальнейшего заполнения опроса начинаются на отметке 7–8 минут. Фреде (2010) полагает, что опросы длительностью от 8 до 10 минут заполнять интересно, а опросы длительностью от 17 до 24 минут становятся скучными, более сложными и менее полезными. Наиболее распространенным мнением с момента доклада SSI в 2009 году считается порог в 20 минут, больше уже слишком много.

Тем не менее в результате некоторых исследований были выявлены факторы, которые влияют на заполняемость анкет больше, чем длительность опросов. Компания MarketTools опубликовала несколько статей, в которых пришла к выводу, что «простого дизайна опросника и простых правил заполнения недостаточно для того, чтобы добиться желаемого уровня вовлеченности респондента». По результатам их исследования были определены другие факторы (такие, как время, затраченное на заполнение одной страницы), которые гораздо более существенно влияют на оценку опроса, чем его длительность. Они утверждают, что «не существует правила, которое может быть применимо во всех случаях», и искажения, которые возникают в длинных опросах, могут быть скорректированы с помощью изменения других параметров исследования.

Помимо разногласий по дизайну опросов, существует также несоответствие между советами по их длительности и привлечению профессиональных респондентов. Существует группа участников панелей, которые регулярно проводят часы за заполнением различных опросов на разные темы. Эта группа респондентов, которая составляет большинство участников панелей, многократно заполняет утомительные опросы охотно и добросовестно. Кажется, что как минимум для этой группы людей длина опроса не так важна, как для остальных, будь то новички или случайные панелисты, заполнившие только одну анкету.

Цель этой статьи состоит в том, чтобы обогатить исследования по длительности опросов за счет более глубокого изучения самой концепции длительности и ее составляющих, а также описания взаимодействия, возникающего между структурой анкеты и человеком, ее заполняющим.

Рабочие гипотезы:

1. Длина опроса не так важна, как его сложность. Длинный и простой опрос предпочтительнее, чем короткий, но сложный.
2. Субъективное восприятие длительности опроса зависит скорее от чувства удовлетворенности в процессе заполнения анкеты и в момент его завершения, чем от его фактической длины.
3. Некоторые панелисты, особенно опытные, более терпимы к длинным и сложным опросам.
4. Сокращения длительности опроса недостаточно для того, чтобы привлечь «случайных» панелистов, зато это уменьшит объем информации, которую можно получить от тех, кто готов ее предоставить.

Структура опроса

Всего в опросе было пять блоков, различающихся между собой по количеству и сложности вопросов¹. Одна пятая участников отвечали на 12 стартовых вопросов, которые не варьировались по сложности. Три блока вопросов включали в себя: (1) вопросы на недобросовестность заполнения анкеты, которые позволяют выявить склонность респондентов к обману в ответах, (2) опыт участия в работе панели, (3) личные качества, включая стремление к познанию и «большую пятерку» характеристик личности.

Вопросы, которые следовали далее, были посвящены бережливости и поделены на две группы в зависимости от сложности. Панелисты отвечали либо на 24, либо на 36 вопросов, и эти вопросы были либо «простыми», либо «сложными». Сложность определялась по тому, требовалась ли дополнительная умственная работа (например, необходимость более тщательно продумать ответ; ранжирование вместо простой оценки; выбор единственного варианта вместо

¹ Под «вопросом» подразумевается один вопрос, группа вопросов или таблица размером от 10 до 20 строк.

множества подходящих и т.д.). Под сложностью не подразумевалась трудность понимания — оба опроса должны были быть понятными и логичными. Целью было сохранить тематику опроса, но варьировать при этом умственные затраты, требуемые для его прохождения. В некотором смысле это должно было показать, как можно упростить опрос, если необходимо уменьшить нагрузку на респондента. Демографически репрезентативная выборка в 5182 респондента была взята из панели компании NPD и случайным образом распределена по пяти блокам, указанным в Таблице 1.

Таблица 1. Структура экспериментального опроса

	Простой	Сложный
12 вопросов		
24 вопроса		
36 вопросов		

В конце опроса было добавлено несколько вопросов, по которым определялся опыт респондента в качестве участника панелей. Ответы на них, наряду с показателем доли респондентов, завершивших опрос, оценкой качества (например, «слепое» заполнение, логические соответствия между ответами) и уровнем отклика на приглашение к участию в следующем исследовании послужили нам в качестве основных зависимых переменных. При этом фактическое и субъективно воспринимаемое время заполнения опроса также были учтены.

Результаты

Неудивительно, что более сложные версии опроса требовали больше времени на заполнение, чем простые. Интересно отметить, что заполнение 24-вопросной сложной анкеты заняло на 2 минуты меньше времени, чем заполнение 36-вопросной, более длинной, но более простой версии, что позволяет напрямую сопоставить время заполнения и сложность опроса (см. Рисунок 1).

Рисунок 1. Время в минутах, затраченное на прохождение опроса, с разбивкой по версиям опроса

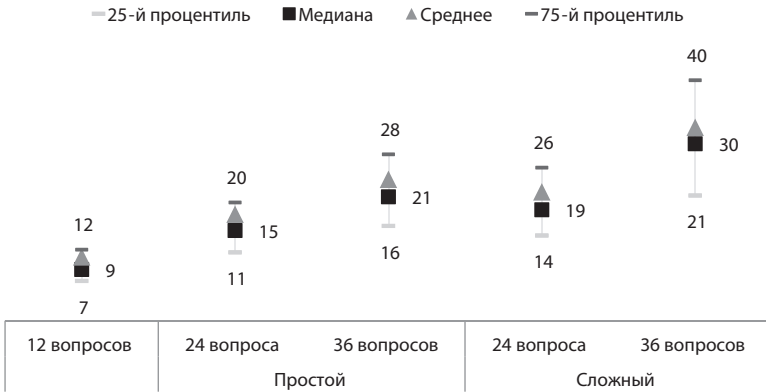
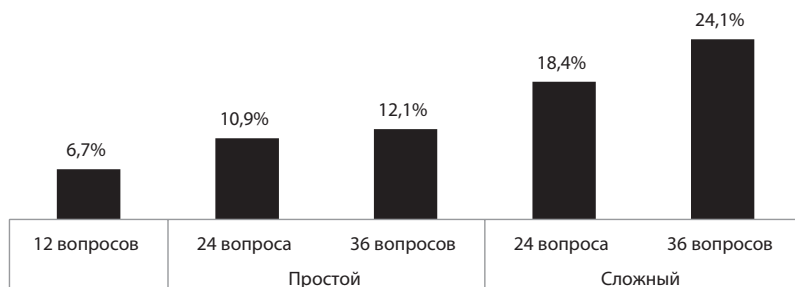
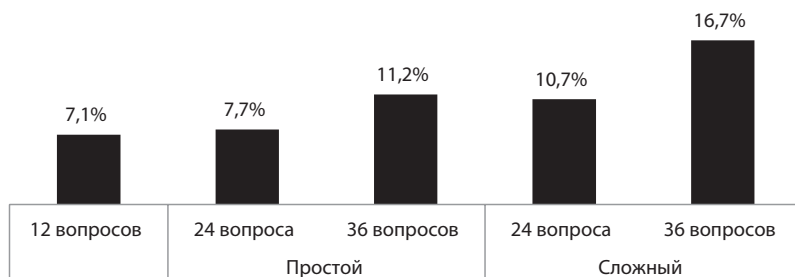


Рисунок 2. Доля покинувших опрос респондентов с разбивкой по версиям опроса



Среди тех респондентов, кто прошел опрос до конца, не было отмечено значимых отличий в удовлетворенности опросом между теми, кто заполнял короткую и среднюю версии. Однако добавление 12 вопросов как в простой, так и в сложный опрос привело к снижению уровня удовлетворенности респондентов, см. Рисунок 3.

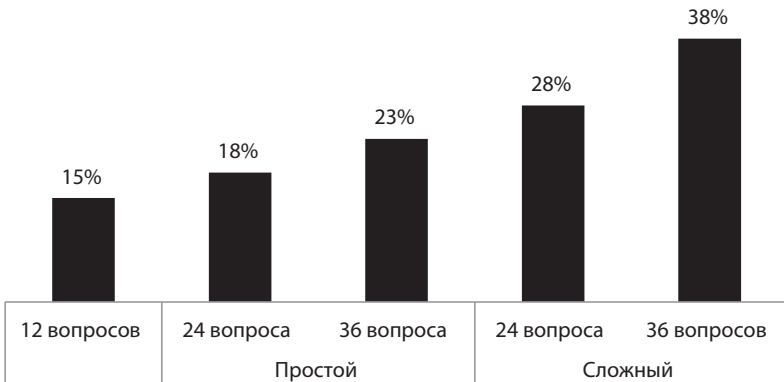
Рисунок 3. Доля ответивших, что заполнение опроса не стоило затраченных усилий, с разбивкой по версиям опроса



На Рисунке 4 приведены данные двух предыдущих рисунков, обобщенные на основе предположения, что те, кто покинул опрос, скорее всего, сделали это потому, что сочли, что опрос не стоит их усилий. На графике ниже показано, как все респонденты — не только ответившие на заключительные вопросы с предоставлением об-

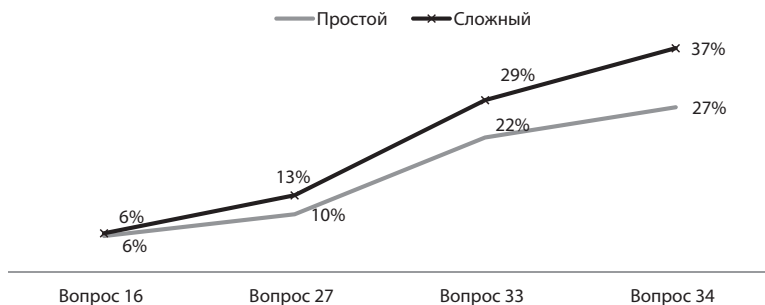
ратной связи — отнеслись к каждой из пяти его версий. Следует отметить, что если бы у нас не было четвертого блока (24-вопросная сложная анкета), то мы могли бы прийти к заключению, что удовлетворенность респондентов падает с увеличением длины опроса. Действительно, в четвертой группе неудовлетворенных меньше, чем в третьей, однако не длина опроса приводит к такому результату, а, скорее, наличие дополнительных вопросов и уровень их сложности.

Рисунок 4. Доля ответивших, что заполнение опроса не стоило затраченных усилий, включая покинувших опрос респондентов, с разбивкой по версиям опроса



На следующем графике показано влияние длительности опроса на вероятность слепого заполнения. Увеличение сложности опроса чаще всего бывает связано с ростом количества вопросов, но здесь оно осталось прежним. На приведенном ниже графике видно, что более сложный опрос приводит к увеличению показателя слепого заполнения. Что касается времени заполнения, то вопрос № 27 в сложной версии задавался примерно в то же время, что и вопрос № 33 в простой версии — примерно на 20-й минуте. Конечно, это не точное время заполнения таблицы, а лишь время от начала заполнения опроса. На вопросе № 33 показатель слепого заполнения возрастает потому, что этому вопросу предшествовало несколько таблиц, а не потому, что панелист уже 20 минут заполняет опрос.

Рисунок 5. Показатель слепого заполнения с разбивкой по версиям опроса



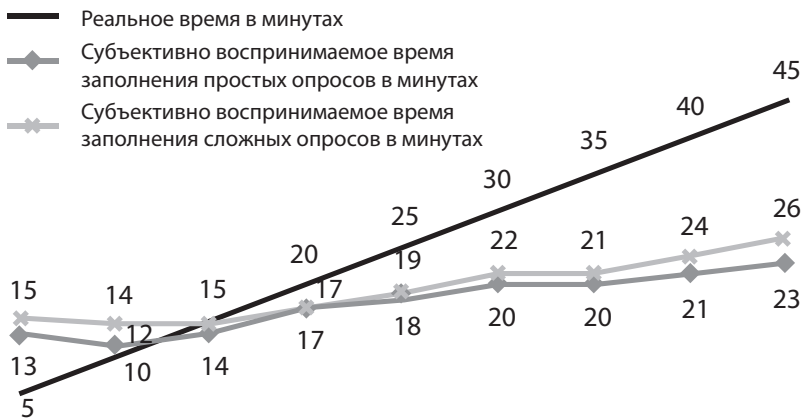
Если говорить кратко, то данные показывают, что длина опроса не так предсказуемо влияет на качество полученных данных, показатель числа покинувших опрос респондентов и их удовлетворенность, как его сложность. Гипотеза 1, согласно которой длинные и простые опросы предпочтительнее сложных, но коротких, подтверждается.

Гипотеза 2: Субъективное восприятие длительности опроса зависит, скорее, от чувства удовлетворенности в процессе заполнения анкеты и в момент его завершения, чем от его фактической длины.

Прежде чем перейти непосредственно ко второй гипотезе, рассмотрим показанное на Рисунке 6 сравнение субъективной оценки времени, потраченного на опрос, с фактическим. Для ясности следует отметить, что респондентов просто спрашивали, сколько минут, по их мнению, они заполняли опрос. Этот показатель, очевидно, смещен в сторону фактического времени, так как панелисты могли посмотреть на часы и узнать, сколько именно времени они потратили на опрос. Тем не менее, большинство панелистов, кажется, этого не сделали, т.к. корреляция между воспринимаемым и реальным временем заполнения составила 0,5. Рисунок 6 еще раз обращает наше внимание на то, о чем говорилось ранее: более простые опросы субъективно воспринимаются как более короткие по сравнению

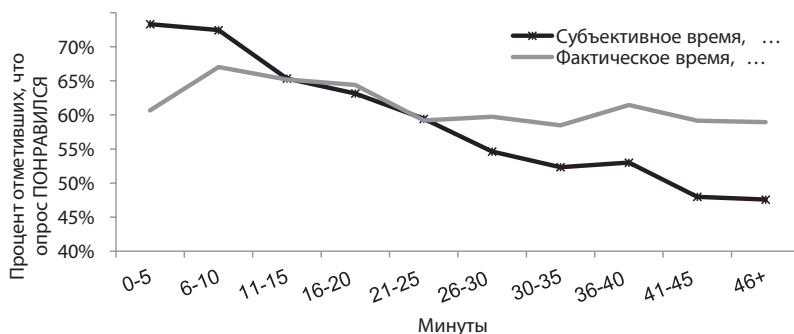
со сложными опросами. Еще один вывод заключается в том, что субъективно воспринимаемое время заполнения оказывается меньше фактического в тех случаях, когда длительность опроса превышает 15 минут, в то время как опросы длительностью до 10 минут воспринимаются как более длинные, чем они есть на самом деле.

Рисунок 6. Субъективно воспринимаемое и фактическое время заполнения с разбивкой по версиям опроса



На Рисунке 7 отмечены процентные доли респондентов, ответивших, что опрос им понравился, в соотношении с показателями фактического и субъективно воспринимаемого времени его заполнения. Если не учитывать субъективную оценку времени, то можно отметить проявление — хоть и незначительное — регрессионной модели с использованием показателя фактического времени заполнения в минутах для предсказания уровня удовлетворенности. Однако добавление показателя субъективно воспринимаемого времени заполнения лишает значимости регрессионную модель, основанную на фактическом времени заполнения. Кроме того, на графике видно, что если воспринимаемая продолжительность опроса составляет менее 11 минут, то это приносит большую удовлетворенность, чем в случае фактически потраченных на опрос 10 минут времени.

Рисунок 7. Доля респондентов, отметивших, что опрос им понравился, в соотношении с показателями фактического и субъективно воспринимаемого времени его заполнения.



Таким образом, вторая гипотеза также подтверждается на практике. Субъективно воспринимаемое время заполнения более тесно связано с удовлетворенностью опросом, чем фактически затраченное время. Возможно, более удачной формулировкой сделанного вывода будет не «короткие опросы лучше», а «опросы, которые воспринимаются как более короткие, лучше». Это как 10 минут разговора, которые могут пролететь незаметно или еле тянуться — ощущение времени основывается скорее на субъективном восприятии, чем на реальной длительности того или иного процесса.

Гипотеза 3: Некоторые панелисты, особенно опытные, более терпимы к длинным и сложным опросам.

Как уже говорилось, первые 12 вопросов помогали определить различные личностные характеристики респондентов и способствовали выявлению склонности к недобросовестному заполнению опросов. Склонность к недобросовестности определялась с помощью пяти вопросов, которые были использованы для разделения респондентов на четыре группы с помощью дискриминантного ана-

лиза². Дискриминантная функция была построена и протестирована в более ранней работе, в которой был проанализирован широкий диапазон видов недобросовестного поведения. Четыре кластера, описанные в рамках дискриминантного анализа:

- Группа А/В: добросовестные респонденты, любят проходить опросы
- Группа С: любят опросы, но возникают трудности со сложными вопросами
- Группа D: рассеянные респонденты, не любят проходить опросы
- Группа F: случайные и непостоянные респонденты, высокий уровень недобросовестности

Несколько многомерных регрессий было построено по данным в зависимости от личностных черт респондентов, группы недобросовестности и опыта участия в панели. Наибольшую прогностическую силу имеют склонность к недобросовестности и опыт. Более того, было выявлено несколько примечательных взаимосвязей между недобросовестностью и опытом. Для дальнейшего анализа респонденты были разделены на четыре группы:

1. Опытные³, группы А/В — 36% выборки
2. Опытные, группа С — 27% выборки
3. Новички⁴, группы А/В/С — 16% выборки
4. Группы D и F, как опытные, так и новички — 21% выборки

Согласно теории, на которой основана третья гипотеза, опытные респонденты из групп А/В и, возможно, также из группы С представляют собой «профессиональных респондентов», или панелистов, которые подходят к опросам серьезно и демонстрируют высокое качество заполнения длинных и/или утомительных опросов. При-

² Дискриминантная функция была описана в нашей прошлой работе, представленной на конференции CASRO 2010, в которой мы проводили кластерный анализ по более чем 60 факторам определения недобросовестности респондентов для создания кластеров от А до F.

³ Опытные респонденты — это те, кто принимал участие в опросах NDP 10 и более раз.

⁴ Новички — респонденты, принявшие участие менее чем в 10 опросах NDP.

мечательно, что эти два кластера составляют большую часть выборки. Между тем, новички из групп А/В/С являются случайными панелистами, и у них другой подход к заполнению опросов. Наконец, к группам D и F относятся недобросовестные респонденты, поэтому длительность и дизайн опроса для этой группы не имеют большого значения. Они отказываются от ответов или выбирают ответ случайным нажатием кнопки как в пятиминутных, так и в десятиминутных опросах. Их данные должны быть отбракованы.

На Рисунке 8 показана доля покинувших обзор респондентов по каждой из групп с разбивкой по версиям опроса. Заметим, что те, кто покинул опрос до того, как были заданы вопросы, определяющие степень добросовестности респондента, сюда не включены. Неудивительно, что опытные респонденты из групп А/В с меньшей вероятностью бросали опрос, даже если он был длинным и сложным. Так, добавление 12 вопросов в легкую версию исследования не увеличило процент бросивших заполнение опроса. Иными словами, переход от 15-минутного опроса к 21-минутному опросу не повлиял на заполнение опроса опытными панелистами групп А/В. С другой стороны, при добавлении 12 дополнительных вопросов в легкую версию, уровень отказа от дальнейшего прохождения опроса для новичков возрос с 3% до 12%. В случае с 24-вопросной сложной анкетой доля покинувших опрос подскочила до 25% для новичков, хотя сам опрос был немного короче, чем в случае с 36-вопросной простой версией.

Несмотря на то, что опытные панелисты групп А/В не прерывали заполнение длинной версии простого опроса, они все же оценивали ее как не стоившую затраченных усилий, см. Рисунок 9. Между тем, новички из групп А/В/С с большей вероятностью отмечали, что опрос не стоил затраченных усилий при любых условиях сложности и длительности. Это говорит о том, что новички имеют завышенные ожидания от опроса по сравнению с тем, что в итоге получают, независимо от длины анкеты.

Рисунки 8 и 9 были объединены на Рисунке 10 для оценки суммарной удовлетворенности респондентов с включением данных по предоставленной ими обратной связи и по числу покинувших опрос участников. Вполне в соответствии с выдвинутой теорией,

Рисунок 8. Доля покинувших опрос респондентов с разбивкой по группам и версиям опроса

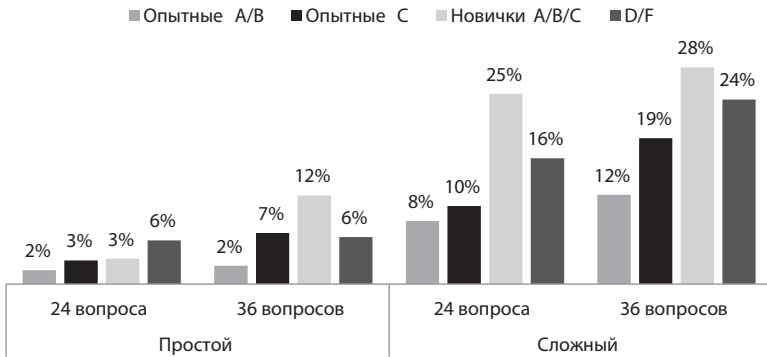
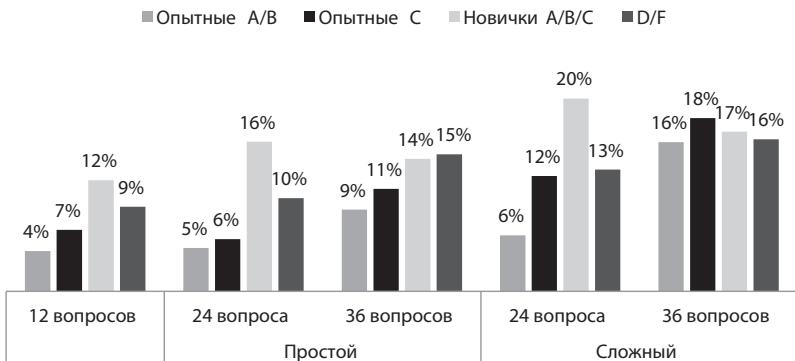
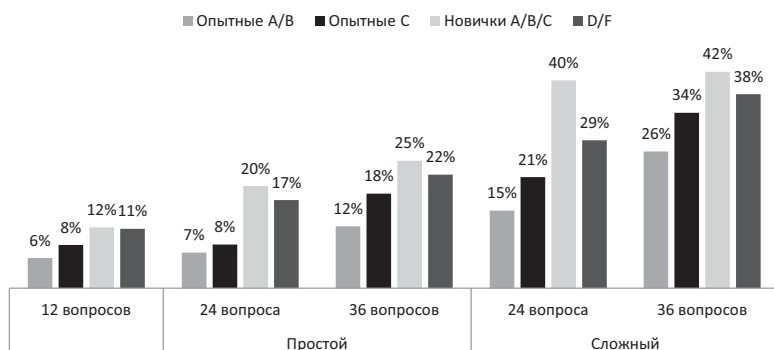


Рисунок 9. Доля респондентов, отметивших, что опрос не стоил затраченных на него усилий с разбивкой по группам и версиям опроса



опытные панелисты групп А/В проявили терпимость к дополнительным вопросам, особенно в более простой версии опроса. Между тем, новички не одобрили появление добавочных вопросов, особенно в сложной версии. Однако переход от 24 к 36 вопросам в сложной версии несущественно повлиял на показатель удовлетворенности в группах новичков А/В/С.

Рисунок 10. Доля ответивших, что заполнение опроса не стоило затраченных усилий, или покинувших опрос респондентов с разбивкой по группам и версиям опроса



На Рисунках 11 и 12 приведены данные по слепому заполнению опроса для каждой группы респондентов. Неудивительно, что слепое заполнение довольно распространено среди респондентов групп D/F. Другой важный вывод, сделанный по результатам тестирования, состоит в том, что сложность опроса заметно усиливает тенденцию к слепому заполнению. Примечательно, что опытные панелисты групп А/В не лишены склонности к слепому заполнению, хотя и делают это значительно реже, чем другие респонденты. Тем не менее, с каждой дополнительной таблицей возрастает усталость даже у самых лучших респондентов. Для сведения: вопрос № 16 — это пятый табличный вопрос, который видят респонденты. Таблицы варьируются между собой в зависимости от того, насколько трудоемко их заполнение, сколько времени требуется на прочтение и насколько большой

список вариантов в них представлен, но пятая таблица как в легкой, так и в сложной версии не привела к существенному увеличению частоты слепого заполнения у «хороших» респондентов. Вопрос № 27 был 13-м по счету табличным вопросом, так что хотя на графике кажется, что это только второй вопрос, в действительности, панелисты заполнили уже около 13 таблиц, прежде чем дошли до него.

Рисунок 11. Показатели слепого заполнения с разбивкой по группам респондентов, легкая версия опроса

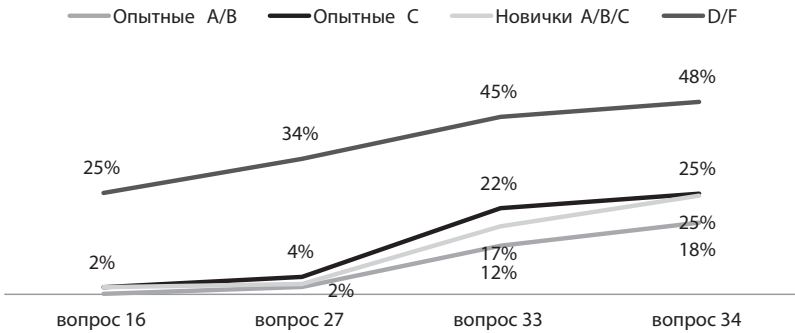
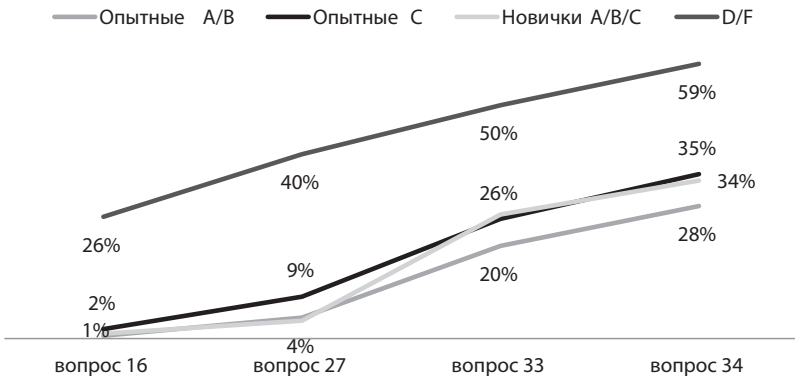
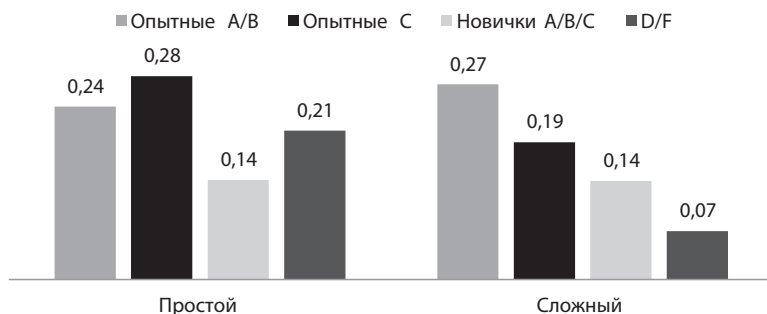


Рисунок 12. Показатели слепого заполнения с разбивкой по группам респондентов, сложная версия опроса



На рисунке 13 представлен еще один показатель качества данных — соотношение между уровнем личного дохода и обеспокоенностью респондента по поводу своего финансового положения. Этот вопрос был задан между вопросами № 27 и № 33, указанными в таблицах выше, т.е. в момент, когда хорошие панелисты начинают некачественно заполнять таблицы. Однако опытные панелисты групп А/В отвечают одинаково как в сложной, так и в простой версии — это говорит о том, что усталость от заполнения табличных вопросов не распространяется на остальные вопросы. Опытные панелисты группы С и панелисты групп D/F демонстрируют более низкие показатели соотношения в сложной версии опроса.

Рисунок 13. Соотношение между уровнем личного дохода и обеспокоенностью своим финансовым положением



В целом гипотеза 3 была подтверждена, хотя и с некоторыми оговорками. Хорошие / опытные панелисты демонстрируют значительно большую терпимость по отношению к более длинным и сложным опросам. Это не означает, что они не могут уставать при заполнении многочисленных и длинных таблиц. Тем не менее, если опрос достаточно простой, они легко заполняют 10–13 таблиц с минимальным уровнем слепого заполнения. Тенденция к слепому заполнению начинает проявляться у них фактически на 14-й по счету таблице. Таким образом, можно сказать, что переломный момент действительно существует, однако он не является критическим.

Гипотеза 4: Сокращения длительности опроса недостаточно для того, чтобы привлечь «случайных» панелистов, но это уменьшит объем информации, которую можно получить от тех, кто готов ее предоставить.

Последняя гипотеза касается анализа стоимости и эффективности сокращения длины опроса. Выход респондентов из опроса, несомненно, приводит к искажению результатов в том случае, если покинувшие опрос участники отличаются от тех, кто прошел его до конца. Однако в случае с панелистами групп D/F — даже если они отличаются от прочих — их ответы настолько неточны, что не имеет смысла оставлять их данные в общей базе. Даже если они представляют собой определенный тип личности или покупателя, они не достаточно честны, чтобы их ответы представляли собой ценность. При анализе данных ответы респондентов, демонстрирующих высокую склонность к недобросовестности, должны быть удалены.

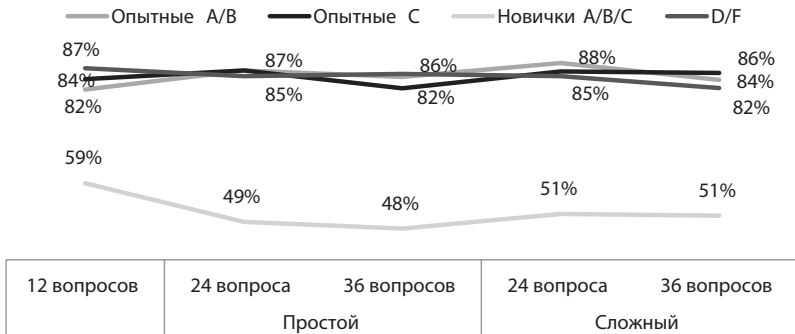
Прочие важные рекомендации — помимо исключения ответов недобросовестных панелистов — можно подразделить на группы по типам респондентов, которых мы хотим привлечь с их помощью. Опытные панелисты групп A/B, которые составляют 1/3 выборки, похоже, готовы потратить 15–30 минут на опрос с 13-ю таблицами и заполнить его, отвечая честно и с удовольствием. Существует и определенный критический момент: если попросить их заполнить еще одну таблицу с длинным списком вариантов ответа, это приведет к тому, что даже хорошие панелисты прибегнут к слепому заполнению, однако достаточно «простые», но длинные опросы в этой группе вполне оправданы. Опытные панелисты группы C менее успешно справляются с дополнительными вопросами и еще менее успешно — со сложными. Для них вполне приемлемыми можно называть достаточно простые опросы длительностью от 11 до 20 минут, даже при наличии множества таблиц. Новички, в свою очередь, приступают к опросу с совершенно иными ожиданиями. Они, похоже, абсолютно нетерпимы к увеличению длительности или сложности опроса, и для них чем короче, тем, определенно, лучше.

Вопрос остается прежним: стоит ли сокращать длину опроса для всех участников? На Рисунке 14 показана доля откликнувшихся

на приглашение принять участие в следующем опросе после первого экспериментального. Наш вопрос звучал так: каким образом опыт, полученный при участии в данном опросе, повлияет на интерес респондента к будущим опросам? В конце концов, если заполнение длинных опросов приводит к отказу респондентов от подобного участия в будущем, то издержки выбранной методики оказываются гораздо выше, чем в случае простого выхода из текущего опроса. Судя по данным, представленным на Рисунке 14, для большинства панелистов полученный опыт не сказывается на их желании участвовать в будущих опросах независимо от версии текущего. Многие новички не соглашались на участие в будущих опросах независимо от опыта, полученного в рамках текущего. Было отмечено лишь одно исключение — отклик на новый опрос был значительно выше у тех новичков, кто в этот раз заполнял самую короткую версию. Это говорит о том, что, если мы хотим, чтобы новые респонденты чаще заполняли опросы, они должны получать более короткие версии длинных опросов. Однако если длина опроса превышает 10 минут, причин для беспокойства нет. Никаких различий в отношении к будущему участию в опросах не было отмечено при прохождении как 15-минутного простого опроса, так и 30-минутного сложного. Это говорит о том, что для новых панелистов действует принцип «все или ничего», т.е. либо они получают 10-минутные интересные анкеты, либо длинную версию, как и все остальные. Другими словами, сокращения 40-минутной анкеты до 20 минут не достаточно, чтобы «привлечь» эту группу респондентов.

Что касается опытных панелистов, нет никаких причин для сокращения длительность опроса. Заполнение длинных или более сложных опросов не имеет негативных последствий для их поведения в будущем. Таким образом, представляется нецелесообразным отказываться от дополнительной информации, которую мы можем получить от большинства панелистов, только потому, что 16% участников не хотят тратить свое время на прохождение опроса. Особенно если учесть, что почти половина этих 16% больше не будет принимать участия в подобных исследованиях, даже если мы изменим структуру опроса, чтобы им угодить.

Рисунок 14. Готовность принять участие в следующем опросе с разбивкой по группам респондентов и версиям опроса



Выводы

Проведенный эксперимент затронул лишь небольшую часть рекомендаций по разработке опросных анкет и привлечению респондентов. Требуется провести множество других подобных экспериментов для анализа других элементов опросов с целью выявления наиболее неприятных для определенных групп участников. Однако собранные нами данные смогли подтвердить тот факт, что время, затраченное на прохождение опроса, не имеет такого значения, как впечатления панелиста и субъективное ощущение времени, потраченного на опрос. Один из наиболее высоко оцениваемых опросов в рейтинге NPD по времени превышает 30 минут, но речь в нем идет о домашних питомцах, и его проходят только те, у кого они есть. Неудивительно, что участники согласны потратить свое время, отвечая на вопросы по теме, которая им интересна. В мире маркетинговых исследований подгонка тем опроса под предпочтения респондентов не считается хорошей практикой, однако внесение небольших изменений в его структуру и стремление не раздражать участников сложными вопросами или многочисленными таблицами приносят большую пользу, чем простое сокращение длины анкеты на половину.

Другой немаловажный вклад этого исследования заключается в предложении более сложного подхода к группировке участников выборки. Если панелисты-новички демонстрируют низкую продолжительность концентрации внимания, хотя отвечают на интересные вопросы и не хотят выполнять скучных и утомительных заданий, возможно, следует направить усилия на то, чтобы выяснить, что может их удерживать в панели для участия хотя бы в одном-двух последующих опросах. Единый подход ко всем участникам может навредить, и вполне вероятно, что попытки удерживать новичков отпугнут опытных панелистов, на которых, по сути, и держится вся система подобных исследований. Кроме того, нет никакой необходимости «чинить то, что не сломано», так как постоянные добросовестные панелисты вполне готовы к участию в длинных опросах, которые так не любят новички или даже сами специалисты по маркетингу. В мире мобильных исследований, поточных выборок и профессиональных респондентов не существует причин, по которым опросы нельзя подбирать под определенные группы респондентов для выработки беспристрастных решений.

Новое поколение онлайн-анкет? Использование преимуществ интерактивности интернета

*Мелани Ревилла
(Исследовательский и экспертный центр
методологии опросов (RECSM) —
Университет Помпеу);
Карлос Очоа (Netquest);
Альберт Турбина (Netquest)*

*Мы хотели бы поблагодарить
Германа Лёве и Виллема Сэриса
за поддержку и ценные комментарии.*

1. Введение

С одной стороны, сбор высококачественных данных с помощью опросных анкет становится все более сложной задачей из-за проблем неполного охвата и низкого отклика респондентов, сложностей с достижимостью респондентов, а также из-за их склонности к разумной достаточности, т.е. к поиску удовлетворительных решений вместо оптимальных в целях минимизации психологических издержек [18]. В рамках опросов разумная достаточность означает, что респонденты выбирают первый приемлемый ответ вместо поиска наиболее подходящего из всех предлагаемых ответов, стремясь затратить меньше усилий [12, 11]. С другой стороны, новые методики, позволяющие собирать данные оригинальными, надежными и инновационными способами (например, пассивный сбор данных, большие данные или исследование методом наблюдения), оказывают давление на традиционные методы проведения опросов с помощью анкет. Мы считаем, что получить от респондентов данные высокого качества можно в том случае, если предложить для запол-

нения анкеты нового поколения, позволяющие совместить приятные впечатления от участия с хорошим качеством собираемой информации. Цель данной работы — проанализировать стратегию усовершенствования опросных анкет для более полного использования богатейших возможностей интернета.

Действительно, интернет предлагает массу возможностей. Он позволяет реализовать сложную систему переходов, проводить проверку ответов в режиме реального времени или использовать дополнительные материалы (изображения, видео, музыку, справочную информацию, определения терминов и т.д.). Также есть возможность записывать параданные опроса, чтобы потом использовать их для описания и оценки процесса его проведения [3]. Ряд исследований были посвящены вопросам использования возможностей интернета. К примеру, Конрад и Шобер [2] работали с аватарами, которые можно было персонализировать в рамках онлайн-опроса, выбрав для них внешность «интервьюеров». Используя панель LISS, Кьеруж и его соавторы [10] провели эксперимент, посвященный изучению эффекта от сопровождения онлайн-анкет технологией автоматического звукового воспроизведения текста — в данном случае респонденты не только видят вопросы на экране, но и слышат зачитывающий их голос.

Тем не менее, в большинстве онлайн-опросов все эти доступные возможности по-прежнему не используются. Вместо этого анкеты преимущественно создаются путем простого переноса вопросов с бумаги на интернет-страницу. Мы уверены, что онлайн-опросы должны развиваться по меньшей мере в двух важных направлениях. Во-первых, они должны предлагать более естественный способ сбора данных, лучше соответствующий имеющимся у респондента навыкам и предпочтениям. Пользователи все быстрее начинают испытывать скуку, переходя с одного сайта на другой, используя свои смартфоны, в течение кратких, но довольно частых временных периодов. В этих условиях привлечь их внимание для заполнения анкеты можно только в том случае, если по-новому подойти к этой задаче. Во-вторых, собираемые данные должны быть высокого качества. Поскольку в настоящее время онлайн-опросы, как правило, проводятся путем рассылки приглашений к участию среди онлайн-панелистов, которым выплачивается вознаграждение, последние

стремятся ответить на все вопросы как можно быстрее и получить вознаграждение. Такой подход можно назвать «исключительно рациональным» [13, р. 915].

Чтобы решить поставленную выше двойную задачу, мы предлагаем комбинировать ограничения на скорость заполнения анкет с динамичной, интерактивной демонстрацией вопросов, при которой информация появляется на экране последовательно, а не целиком и сразу. Принцип, лежащий в основе предлагаемых ограничений по скорости заполнения анкетных форм, состоит в следующем. Контролируя процесс заполнения анкеты путем установки минимально допустимого времени ответа на вопрос, мы можем в первую очередь отстранить от участия недобросовестных респондентов, а во вторую побудить тех, кто обычно действует по правилу разумной достаточности, приложить больше усилий, что в итоге позволит повысить качество получаемых данных. В целом мы предполагаем, что в ситуации, когда респонденты не могут быстро перейти к следующему вопросу, они будут вынуждены задуматься о том, стоит ли при заполнении руководствоваться принципом разумной достаточности.

Связь между временем, потраченным на ответ, и показателями его качества была зафиксирована в рамках целого ряда исследований. В работе Малхотры [13, р. 915] подтверждена связь между повышением скорости ответа и усилением эффекта преимущественного выбора, который определяется как тенденция чаще выбирать первый из предложенных вариантов [17]. Жэнг [20] и Грецки, Мейер и Шоэн [8] обнаружили, что повышение скорости ответов связано с учащением случаев слепого заполнения таблиц анкеты, представляющих собой систематический выбор одного и того же варианта ответа для всех строк в таблице. Жэнг и Конрад [21] также заметили, что это соответствие ярче проявляется среди респондентов с невысоким уровнем образования. Проявления эффектов преимущественного выбора и слепого заполнения часто используются как индикаторы следования принципу разумной достаточности [1] и, следовательно, как показатели низкого качества предоставленных ответов. Эти выводы также подтверждены результатами Ревилла и Очоа [16], полученными с использованием метода моделирования структурными уравнениями и анализа различных показателей качества.

Поскольку время, потраченное на ответ, и его качество взаимосвязаны, мы можем предположить, что качество собираемых данных можно повысить, влияя на скорость ответа. Исследователи не могут заставить людей думать. Однако в рамках онлайн-опросов они могут заставить их потратить на каждый вопрос не меньше определенного времени. В данном подходе есть два потенциальных ограничения. Во-первых, предшествующие исследования подтвердили только наличие связи. Если она не является каузальной, ограничение минимального времени ответа никак не скажется на его качестве. И, во-вторых, даже в том случае, если респонденты будут вынуждены потратить определенное минимальное время на каждый вопрос, это не гарантирует, что они будут дольше думать над своим ответом. Тем не менее это приводит к снижению стимула действовать по принципу разумной достаточности. Даже если респондент быстро выберет ответ, ему придется подождать, прежде чем он сможет продолжить заполнение анкеты. Таким образом, применяя принцип разумной достаточности, он по-прежнему сможет прикладывать меньше когнитивных усилий для ответа, но сэкономить время уже не сможет. Следовательно, мы можем выдвинуть гипотезу о том, что ограничение минимального времени ответа уменьшает «добавленную стоимость» разумной достаточности и в результате повышает качество ответов.

Кэпелнер и Чэндлер [16, р.3] провели эксперименты в этой области, используя два разных способа вынудить респондентов медленнее отвечать на вопросы анкеты: «Функция *контроля времени* (*Timing control*) идентична функции *контроля* (*Control*) за тем исключением, что кнопка «Продолжить» блокируется вращающейся иконкой, оставаясь неактивной в течение периода ожидания, по окончании которого вновь активизируется. Функция «*Капча*» (*Captcha*) не только замедляет процесс заполнения анкеты на определенный период, как и в случае с *контролем времени*, но и привлекает дополнительное внимание респондента к представленным инструкциям и возможным вариантам ответа, выводя текст вопроса постепенно, со скоростью 250 слов в минуту».

Принцип, лежащий в основе динамичной и интерактивной демонстрации текста, согласуется с аргументами Кэпелнера и Чэндлера в поддержку функции «Капча»: представление вопросов

подобным образом (т.е. постепенно выводя их на экран) вынудит респондентов сосредоточиться на своем задании и обдумать всю информацию вместо того, чтобы уделить внимание только какой-то части задания, к примеру, шкале ответов, на основе которой сделать предположение о сути вопроса. Поскольку информация выводится последовательно, респонденты поэтапно узнают о своем задании: сначала читают вводный текст, затем вопрос и, наконец, варианты ответов. Действительно, онлайн-опрос с этой функцией можно рассматривать в качестве приближенного к личному интервью, в рамках которого респонденты слушают говорящего и получают всю информацию постепенно, а не целиком и сразу.

Кэпелнер и Чэндлер [9] обнаружили некоторое положительное влияние использования перечисленных функций (особенно функции «Капча») на качество полученных данных. Однако уровень усталости респондентов в процессе заполнения анкеты повышался. При этом исследователи анализировали очень специфический тип онлайн-опросов, в котором принимали участие работники сервиса Amazon Mechanical Turk (MTurk). Кроме того, рассматривалось ограниченное количество функций, эффективность влияния которых на качество данных которых оценивалась всего по нескольким показателям.

В данной статье мы более глубоко анализируем эффективность применения обеих функций — ограничения минимального времени ответа и динамичного и интерактивного представления вопросов. Мы протестировали несколько новых идей: во-первых, мы рассмотрели больше функций, чем Кэпелнер и Чэндлер [9]; во-вторых, мы подстраивали скорость опроса под скорость чтения каждого респондента, и, в-третьих, мы учитывали больше показателей качества данных.

В Разделе 2 представлен весь тот долгий процесс, через который мы прошли, чтобы протестировать опросные анкеты нового поколения, включая предварительное исследование, лабораторный эксперимент, полевые работы исследования, к участию в котором приглашались добровольцы из онлайн-панели, а также разработку программного обеспечения. В Разделе 3 мы развиваем свою гипотезу об ожидаемом эффекте введения ограничений на время ответа,

а также динамичного и интерактивного представления вопросов. В Разделе 4 представлены результаты основного эксперимента. Наконец, в Разделе 5 представлены выводы по итогам нашей работы.

2. Предварительное исследование для разработки структуры экспериментального проекта

Выбранный подход был протестирован в рамках нескольких этапов, начиная с предварительного тестового исследования (далее «предварительное тестирование»), позволившего выявить несколько проблем. Затем мы провели лабораторный эксперимент и получили более глубокое понимание происходящего. Основываясь на всей полученной информации, мы провели полномасштабный эксперимент (далее «основной эксперимент»).

Важно отметить, что техническая сторона нашей работы, благодаря которой стали возможны динамические опросные анкеты, была выполнена с учетом того факта, что реальные опросы зачастую приходится программировать всего за несколько дней, и у людей, занимающихся этим, не всегда есть необходимое образование в области компьютерного программирования. Решение, использованное для программирования интерактивного опроса, протестированного нами в ходе исследования, способно легко трансформировать вопросы почти всех стандартных типов в вопросы с поэтапным выводом текста. Таким образом, его действительно можно применять в практической исследовательской деятельности на регулярной основе.

2.1 Предварительное исследование

Предварительное исследование было проведено в Мексике (декабрь 2013 г.) компанией Netquest (www.netquest.com), специализирующейся на онлайн-опросах и сертифицированной по стандарту качества ISO 26362. У компании есть собственные онлайн-панели в Испании, Португалии, Центральной и Латинской Америке. Netquest набирает панелистов «только по приглашению», используя базу адресов электронной почты людей, согласившихся предоставить свои контактные данные и получать электронные письма по-

сле участия в небольшом опросе удовлетворенности, размещенном на различных интернет-сайтах.

С полной версией опросной анкеты можно ознакомиться на странице: <http://test.nicequest.com/surveys/esnpa/force/>.

2.1.1 Функции

Функция № 1 идентична функции «контроля времени» в исследовании Кэпелнера и Чэндлера [9]. Автоматическая система контроля запрограммирована таким образом, что позволяет респонденту перейти к следующему вопросу только по прошествии определенного времени. Если участник проводит на странице больше минимально установленного времени до нажатия кнопки «Далее», то процесс перехода не отличается от процесса перехода в контрольной группе, где опрос проводится обычным образом, т.е. на каждой странице текст представлен целиком с самого начала, и респондент может перейти к следующему вопросу после ввода ответа, когда пожелает. Но если участник нажимает на кнопку «Далее» раньше установленного минимального срока, на странице откроется окно с предупреждающей надписью: «Мы считаем, что вы не уделите достаточно времени данному вопросу. Пожалуйста, прочтите его более внимательно, прежде чем перейти к следующему». После того как респондент закрывает это окно и возвращается к вопросу, вместо кнопки «Далее» появляется иконка с часами, исчезающая после завершения минимально установленного срока.

Мы также применили функцию «Капча» (Функция № 3), при которой вопрос появляется на странице посимвольно. В этом случае респондент не может сразу перейти к ответам, а должен постепенно ознакомиться с появляющимся на странице текстом. Это может стать для респондентов дополнительным стимулом к последовательному считыванию информации.

Кроме того, мы предложили новую функцию — нечто среднее между Функцией № 1 и Функцией № 3 — под названием «Ступенчатые вопросы» (Метод № 2). В данном случае информация появляется на странице блоками: сначала введение (если оно предусмотрено в вопросе), затем сам вопрос, после этого варианты ответов. Самой последней на экране возникает кнопка «Далее».

2.1.2 Определение минимального времени ответа

Функция «Капча» используется для установки минимального времени ответа для каждой страницы. Оно зависит от скорости, с которой символы появляются на странице. Полученный таким образом временной интервал затем используется и при использовании других функций.

Кэпелнер и Чэндлер [9] использовали фиксированную скорость на уровне 250 слов в минуту. Мы повторили это условие. Однако мы также вычисляли минимальное время ответа для каждого респондента в зависимости от его скорости чтения. Чтобы определить последнюю, перед началом опроса был добавлен соответствующий тест. Он позволяет изменить скорость в пределах $\pm\Delta T\%$ по отношению к установленному уровню в 250 слов в минуту. Эта «возможность подстройки» стала еще одним нововведением нашего исследования. Далее в работе мы будем использовать понятие «с тестом» для обозначения групп, прошедших такой тест скорости чтения перед началом опроса (условие настройки), и понятие «без теста» для всех остальных групп.

Опишем эту функцию более подробно. В процессе ее подстройки респонденты получают следующие инструкции: *«На этой странице вы увидите текст, который начнет постепенно проявляться нажатия кнопки «Начать чтение». После прочтения текста вам необходимо будет ответить, нормальной ли была для вас скорость появления текста. Уделите, пожалуйста, максимум вашего внимания этому заданию, так как по его завершении мы зададим вам несколько вопросов по прочитанному вами тексту»*. Если респондент сообщает, что скорость была хорошей, то он переходит к следующему этапу исследования. Если скорость показалась ему слишком низкой, то ему предлагается новый текст, появляющийся на экране быстрее (300 слов в минуту). И наоборот, если она оказалась слишком высокой для респондента, он переходит на страницу с новым текстом, появляющимся уже со скоростью 200 слов в минуту. Если после второй попытки респондент снова отмечает, что скорость была слишком высокой или слишком низкой, то он получает третью и последнюю попытку ее подбора.

Сразу после окончания теста для проверки соответствия выбранной скорости навыкам чтения респондента на экране на выбранной скорости появляется, а затем исчезает небольшой текст. Респондент должен ответить на вопрос по прочитанному. В шкале ответов представлен вариант «Я не успел прочесть текст». Если респондент выбирает этот вариант или не может правильно ответить на вопрос, мы делаем вывод, что нам не удалось подстроить скорость появления текста под его навыки чтения, и снижаем ее для последующего опроса.

2.1.3 Основные результаты

Тест скорости чтения

Система подстройки скорости появления текста на экране сработала не так, как ожидалось. Мы предполагали, что эта функция позволит респондентам повысить скорость появления текста на экране, и, таким образом, условия проведения опроса с возможностью подстройки окажутся менее строгими. Однако 72,8% респондентов в итоге выбрали более низкую скорость, чем в группах без подстройки (т.е. менее 250 слов в минуту). Еще 25,7% респондентов остановились на скорости 250 слов в минуту. Более высокую скорость выбрали немногие. Причиной полученного результата мог стать тот факт, что респондентов попросили уделить максимум внимания представленному тексту, так как затем по нему будет задан вопрос. Возможно, учитывая это условие, респонденты предпочли выбрать более низкую скорость, чтобы суметь запомнить весь текст. Более того, в процессе тестирования только 63,3% респондентов дали правильный ответ на вопрос; еще 18,4% сообщили, что не успели прочесть текст, а 18,2% дали неправильный ответ. Скорость, выбранная последними, была в итоге снижена нами на 50 слов в минуту.

Выбор контрольного вопроса на проверку внимательности (ИМС)

Проверка внимательности «представляет собой вопрос, вставленный в опросную анкету, который предлагает участникам исследования (...) предоставить подтверждение того, что они прочли все инструкции

в задании» [15, р. 867]. Если респондент правильно отвечает на вопрос, это свидетельствует о том, что он действительно прочел инструкции, а также служит индикатором качества полученных в итоге данных. В настоящем исследовании в рамках проверочного теста мы просили респондентов указать все варианты ответа, которые содержались в заданном ранее табличном вопросе, помимо тех пяти, которые они выбрали при ответе.

В контрольной группе проверку на внимательность прошли менее 30% респондентов. При этом в случае использования интерактивных анкет проверку на внимательность провалили 97% респондентов. Проанализировав способ представления ИМС в анкете, мы установили причину произошедшего. Указания находились в девятой строке таблицы, в которой респонденты должны были отметить по пять вариантов ответа в каждом из трех столбцов. При использовании функций № 2 и № 3 второй столбец появлялся после того, как респондент отмечал один вариант ответа в первом столбце. Проверка того факта, что он отметил все пять необходимых вариантов ответа, не проводилась. В этой и без того сложной таблице сам факт появления следующего столбца наводил респондентов на мысль, что достаточно было отметить только один вариант ответа. Вместо побуждения уделить больше внимания вариантам ответа, система, наоборот, побуждала их переходить дальше. Поэтому тест на внимательность необходимо было изменить.

Метод контроля времени ответа

Показатель числа респондентов, покинувших опрос до его завершения, оказался существенно выше в группах, где применялась функция контроля времени. Кроме того, эта функция также оказалась единственной, при применении которой значительно более высокая доля респондентов оценила вопросы анкеты как сложные. Также на анкету поступило больше жалоб.

Тестирование функции контроля времени в основном повторяло эксперимент Кэпелнера и Чэндлера [9]. В отличие от других функций здесь преимущества интерактивности интернета не используются, потому она была нам не так интересна, как другие. Поскольку результаты использования этой функции в рамках предварительно-

го исследования оказались значительно хуже остальных, мы решили не включать ее в наше исследование на последующих этапах.

2.2 Лабораторный эксперимент

Вследствие проблем, выявленных в процессе предварительного тестирования, мы пришли к необходимости получить более четкое представление о поведении респондентов, особенно в тех группах, в которых мы применяем исследуемые функции. Поэтому в сентябре-октябре 2014 г. мы провели лабораторное исследование в Испании. Респонденты заполняли онлайн-анкеты в помещении, где мы могли за ними наблюдать.

2.2.1 Методика исследования

В течение всего опроса мы вели видеосъемку экранов компьютеров и лиц участников. Анализируя полученные таким образом материалы, мы могли видеть реакцию респондентов, отследить движения мыши, невербальные сигналы и т.д. Ответив на все вопросы анкеты, респонденты предоставляли обратную связь, описывая свои впечатления от участия в опросе.

Восемнадцать участников исследования были случайным образом отобраны в контрольную группу, группу Функции № 2 («Ступенчатые вопросы») и группу Функции № 3 («Капча») с распределением по подгруппам тех, кто не проходил тест на скорость чтения, прошедших тест предварительного исследования (Тест № 1) и новый тест (Тест № 2). В рамках нового теста респондентов просили нажать на кнопку по мере их готовности начать чтение. После нажатия на кнопку на экране появлялся небольшой текст. По завершении чтения респонденты должны были снова нажать на кнопку. Это упражнение повторялось три раза подряд с разными текстами на общие темы (напитки), знакомые большинству респондентов. Средняя скорость второй и третьей попыток устанавливалась в качестве скорости всего последующего опроса в группах, где тестировались разработанные нами функции, однако она была ограничена пределами 200–350 слов в минуту. По окончании опроса участни-

ков просили пройти альтернативную версию теста на скорость чтения. Таким образом, в рамках обратной связи они могли сообщить нам, какой вариант понравился им больше.

Для проверки внимательности мы использовали два метода. Первый из них включал в себя ряд вопросов, в которых респондентов просили кликнуть мышкой в определенном месте в качестве подтверждения факта прочтения текста, вместо того чтобы выбрать вариант ответа из числа предложенных альтернатив. Второй представлял собой отдельный вопрос, в котором сначала проводилась инструкция (она выглядела как вводный текст), после чего был приведен «обычный» вопрос с кнопками выбора вариантов ответа. Респонденты должны были не отвечать на вопрос, а кликнуть мышкой по треугольнику на экране.

Для лабораторного исследования была разработана новая анкета, которую от анкеты предварительного тестирования отличал более привычный формат. Например, на первом этапе использовались изображения, которые надо было перетащить в нужное место, полосы прокрутки или вызов меню по нажатию кнопки мыши; на втором этапе анкета была построена на кнопках и окошках, позволяющих отметить все подходящие варианты ответа.

2.2.2 Основные результаты

Тест на скорость чтения

Во-первых, не было отмечено никаких проблем с пониманием двух предложенных тестов. Во-вторых, даже в тех случаях, когда по обоим тестам высказывались разные мнения, в целом респонденты предпочитали новый вариант. Он позволяет более четко подстроить скорость появления текста под скорость чтения каждого респондента, даже если, по мнению участников исследования, опрос проходит слишком медленно.

Два теста на внимание

Всего один человек не прошел тест на внимание, состоящий из блока вопросов, еще один — в случае, когда проверка проводилась отдельным вопросом. Оба респондента указали, что не видели

такого текста. Таким образом, проблема была не в непонимании задания, а в том, что опрашиваемые не уделили достаточного внимания написанному. Так как оба задания для проверки внимания были изложены достаточно понятно, мы оставили их для последующего опроса в реальных условиях, чтобы получить больше показателей качества последнего.

Прочие усовершенствования

Некоторые респонденты предложили нам отделить те вопросы, ответы на которые они хорошо знают и без подсказки, от прочих. Речь идет о личных сведениях. К примеру, некоторые жаловались, что им приходилось ждать, пока на экране появятся все варианты ответа на вопрос об их образовании, хотя подходящая для них альтернатива уже появилась. Основываясь на этом пожелании, мы изменили программу проведения опроса так, чтобы респонденты могли отметить верный вариант ответа сразу после его появления на экране. Кроме того, мы разделили вопросы на две группы. В случае вопросов, ответы на которые респонденты знали без подсказки и могли сообщить их без обязательного ознакомления со всеми существующими вариантами, сразу после выбора ответа можно было перейти на следующую страницу. В случае всех остальных вопросов участникам исследования приходилось ждать, пока не появятся все предлагаемые варианты ответа.

2.3 Окончательный вариант методики исследования

Основной опрос проводился в Испании с 20 ноября по 14 декабря 2014 г. с помощью онлайн-панели Netquest. С опросной анкетой, использованной в ходе эксперимента, можно ознакомиться по адресу: <http://test.nicequest.com/surveys/esnpa/force2>. Для получения репрезентативной выборки участников панели отбор проводился по двум пересекающимся квотам — по возрасту и полу. Количество наблюдений в каждой из шести случайным образом сформированных групп, различающихся по типам тестов, составило 249–250. В Таблице № 1 приведены сведения по шести условиям опроса.

Таблица № 1: Структура эксперимента

	С тестом	Без теста
Контрольная группа	Контрольная группа с Тестом № 2	Контрольная группа
Тестовая группа, Метод № 2: текст выводится по блокам	Тестовая группа, метод «Ступенчатые вопросы» с Тестом № 2	Тестовая группа, метод «Ступенчатые вопросы»
Тестовая группа, Метод № 3: текст выводится посимвольно	Тестовая группа, метод «Капча» с Тестом № 2	Тестовая группа, метод «Капча»

Респонденты были распределены по группам случайным образом. Таким образом, значимые различия между контрольной и тестовой группой должны быть следствием применения соответствующей функции проведения опроса. Во всех случаях мы сравнивали между собой контрольную и тестовую группы, в которых измерялась скорость чтения, а также контрольную и тестовую группы, в которых скорость чтения не измерялась. Поэтому в результате у нас получились четыре парных сравнения.

3. Чего мы ожидаем?

Мы хотим узнать, будут ли респонденты предоставлять данные более высокого качества в случае прохождения динамичных опросов с ограничениями на минимальное время ответа. Таким образом, основной ожидаемый нами положительный результат применения предложенных методов — это повышение качества ответов (H1). Мы также ожидаем, что благодаря последовательному представлению вопросов анкеты (посимвольно или по блокам) впечатления от участия в опросе будут оценены более позитивно (H2), особенно в случаях, когда респонденты могут изменить скорость представления вопросов в ходе теста на скорость чтения.

Однако, ограничивая свободу респондентов в плане выбора темпа прохождения опроса, мы можем столкнуться с проблемой другого рода. Участники исследования, желающие быстрее переходить от вопроса к вопросу, но не имеющие такой возможности, могут в результате фрустрации принять решение о прекращении участия в опросе. Поэтому мы ожидаем, что число респондентов, не завер-

шивших заполнение анкеты, будет выше в тех группах, где применяются разработанные нами функции (**Н3**), как уже ранее обнаружили Кэпелнер и Чэндлер [9]. Более того, оно будет выше в тех случаях, когда минимальное время для ответа на каждой странице анкеты устанавливается заранее на одном уровне для всех респондентов, чем в случаях, когда эта величина меняется в соответствии с результатами теста на скорость чтения (**Н4**). Действительно, тест должен адаптировать минимальное время ответа в соответствии со скоростью чтения респондента, снижая таким образом риск фрустрации. Если опрашиваемый читает быстро, тест это зафиксирует и позволит ему быстрее ответить на все вопросы анкеты. Следует также отметить, что случаи выхода респондентов из опроса рассматриваются обычно как отрицательный результат, однако если это касается тех, кто привык «пробегать» анкету как можно быстрее, не вчитываясь в вопросы и не обдумывая свои ответы, то данный эффект можно расценивать и с точки зрения положительного влияния на качество полученных в итоге данных.

4. Результаты основного опроса

4.1 Повышение доли отказов от участия?

Мы говорим об отказе от участия в тех случаях, когда респондент не завершает заполнение анкеты, даже если прошел все отборочные фильтры. Доля отказов AR высчитывается следующим образом: $AR = NbA * 100 / (NbA + NbF)$, где NbA — число покинувших опрос панелистов, а NbF — число панелистов, завершивших опрос. В нашем случае опрос покинули всего 124 панелиста, а завершили — 1497. Таким образом, доля отказов от участия составила 7,6%. В Таблице № 2 представлены показатели доли отказа от участия в различных экспериментальных группах.

Таблица № 2: Доля отказов от участия в разных группах (в %)

	Контрольная группа		Группа «Ступенчатые вопросы»		Группа «Капча»		Всего
	с тестом	без теста	с тестом	без теста	с тестом	без теста	
Доля отказов, AR	2,3	6,7	9,8***	9,1	9,5***	8,1	7,6

Примечание: знак *** указывает на $p < 0,01$; ** указывает на $p < 0,05$; * указывает на $p < 0,10$. В этой таблице и всех последующих звездочками отмечены результаты, отличающиеся от результатов контрольной группы. Тестирование значимых различий проводилось между контрольной группой с тестом и группой, где применялся метод «Ступенчатые вопросы» с тестом, контрольной группой с тестом и группой, где применялся метод «Капча» с тестом, контрольной группой без теста и группой, где применялся метод «Ступенчатые вопросы» без теста, контрольной группой без теста и группой, где применялся метод «Капча» без теста.

Показатели доли отказов от участия оказались несущественно выше в группах без теста на скорость чтения, в которых использовались разработанные нами функции. Однако в группах с тестом на скорость чтения, где применялись разработанные нами функции, число отказов от участия в двух случаях оказалось значимо выше по сравнению с контрольной группой. Этот результат стал для нас неожиданным, потому потребуется провести эксперимент повторно, чтобы убедиться, что это не случайность. Выявленные нами существенные различия обусловлены тем, что доля отказов от участия в контрольной группе с тестом оказалась крайне низкой (гораздо ниже, чем в контрольной группе без теста), и в меньшей степени тем, что доля отказов от участия в группах, где были использованы разработанные нами функции и проводился тест на скорость чтения, оказалась высокой.

В целом отмечено небольшое повышение доли незавершенных анкет, однако оно не настолько велико, чтобы отбросить динамические опросные анкеты как перспективный вариант. Фактически небольшое повышение доли незавершенных интервью среди респондентов, которые могли бы предоставить ответы очень низкого качества, в определенной степени можно назвать желательным. В данной ситуации необходимо проанализировать большее число параметров. На всех последующих этапах исследования мы принимали во внимание только 1497 респондентов, ответивших на все вопросы анкеты.

4.2 Тест на скорость чтения

В группе респондентов, прошедших наш опрос до конца, мы прежде всего проверяем, сработал ли наш новый тест на скорость чтения так, как ожидалось. В Таблице № 3 приведены средняя, медианная, минимальная и максимальная скорости чтения (количество слов в минуту) для трех испытаний в группах, в которых проводился этот тест (N=746). Первое испытание рассматривается в качестве обучающего, поэтому окончательная скорость вывода текста вычисляется в качестве среднего значения показателей второго и третьего испытаний. Кроме того, итоговые минимальное и максимальное значения ограничиваются показателями в 200 и 350 слов в минуту.

Таблица № 3: Показатели скорости чтения по отдельным испытаниям в группах с тестом

	Средняя	Медианная	Минимальная	Максимальная
Первое испытание	370	314	40	11,000
Второе испытание	586	356	31	42,000
Третье испытание	366	265	46	24,600
Итоговое значение	288	295	200	350

При использовании повторного теста респонденты демонстрировали в среднем гораздо более высокую скорость чтения, чем в ходе предварительного тестирования. Однако были выявлены огромные различия между минимальными и максимальными показателями. С одной стороны, некоторые респонденты показали крайне низкую скорость чтения, что, вероятно, было связано с параллельным решением ими других задач. С другой стороны, некоторые респонденты показали невероятно высокую скорость. Результаты на уровне 42 000 слов в минуту невозможны даже для самых быстрых читателей — такие респонденты, несомненно, не читали представленный на странице текст, а просто нажали на кнопку продолжения опроса. Лучшие из участников Мирового чемпионата по скоростному чтению демонстрируют среднюю скорость чтения на уровне 1000–2000 слов в минуту с усвоением 50% прочитанного текста. Результат чемпиона мира по скорочтению составляет 4700 слов в минуту с усвоением

67% прочитанного (Wikipedia). Чтобы привести крайние показатели к более реальным значениям, мы также учли медианные значения. Они неизменно оказывались выше уровня 250 слов в минуту, подтверждая тот факт, что повторные тесты позволяют подстраивать скорость вывода текста так, как мы этого ожидали.

Что касается вопроса, заданного сразу после теста с целью проверить, хорошо ли подобрана скорость вывода текста, правильный ответ на него дали только 24,6% респондентов. Даже с учетом того, что лишь 6,1% опрошенных отметили, что «не смогли прочесть текст», почти 70% участников не смогли выбрать правильный ответ, что наводит на мысль о неправильно выбранной скорости. Тем не менее в процессе опроса текст со страницы не исчезает. Респондент всегда может потратить на чтение больше времени. Напротив, если выбрана действительно низкая скорость, перейти от вопроса к вопросу быстрее возможности не будет. Кроме того, во многих вопросах нет никакой необходимости читать весь текст для выбора варианта ответа, так как используются классические формулировки (возраст, образование, профессия и т.д.) или же в группах вопросов используется общая формулировка, в которой меняются лишь отдельные слова. Таким образом, мы считаем оптимальным вариантом немного завышенную скорость вывода текста на экран.

4.3 Повышение качества?

4.3.1 Проверка внимания и совместный анализ

Сначала мы определяем долю респондентов, правильно выполнивших указания двух тестов на внимание, описанных в подразделе 2.2.1. Мы также включили в анкету «совместный» вопрос, при ответе на который респонденты должны выбрать наиболее предпочтительный вариант из четырех предложенных. В этом тесте были использованы вопросы, действительно используемые для проведения совместного анализа [6, 7, 5]. Респонденту демонстрируют несколько продуктов, в каждом из которых привлекательные для него характеристики сочетаются с нежелательными. Это заставляет участника исследования идти на компромисс, определяя, какие ха-

рактические характеристики оказывают более весомое влияние на его выбор. В нашем случае мы включили один вариант ответа, который, очевидно, был привлекательнее трех других (то же сочетание напитков, но по более низкой цене). Поэтому выбор других альтернатив служил для нас показателем того, что респондент не уделил анкете достаточно внимания. Таким образом, неудачное прохождение проверки на основе «совместного анализа» (т.е. выбор не лучшего варианта из предложенных) также является индикатором низкого качества ответов.

В Таблице № 4 указана доля респондентов (в процентах), правильно выполнивших указания двух проверок на внимание и выбравших лучший вариант в «совместном» вопросе.

Таблица № 4: Доля респондентов, успешно прошедших проверку на внимание и ответивших на вопрос по методу совместного анализа (в %)

	Контрольная группа		Группа «Ступенчатые вопросы»		Группа «Капча»		Всего
	с тестом	без теста	с тестом	без теста	с тестом	без теста	
ИМС в таблице	72,0	71,6	86,4***	84,7***	85,9***	82,8***	80,6
ИМС отдельно	55,2	48,8	73,9***	77,1***	87,6***	78,8***	70,2
Метод совместного анализа	88,8	90,8	87,6	92,4	90,0	89,2	89,8

Примечание: знак *** указывает на $p < 0,01$; ** указывает на $p < 0,05$; * указывает на $p < 0,10$.

Существенно больше респондентов прошли обе проверки на внимание во всех группах, где мы использовали разработанные нами опросные функции. Улучшения заметны сразу. В ситуации, когда информация выводится на экран постепенно, респонденты уделяют инструкциям больше внимания.

Что касается «совместного» вопроса, здесь, наоборот, существенных различий между разными группами респондентов зафиксировано не было. Вероятно, причина в том, что в данном случае ключевая информация, на основе которой можно дать правильный ответ, содержалась в предложенных вариантах ответа (цены были указаны на изображениях, одно из которых должны были выбрать респонденты). Большинство участников исследования, даже если они ру-

ководствуются принципом разумной достаточности, не выбирают ответ случайным образом. Мы полагаем, что они с большей вероятностью прочитают и обдумают варианты ответа, чем сведения, изложенные во введении или многословном вступлении к вопросу. Результаты, приведенные в Таблице № 4, подтверждают высказанные нами предположения.

4.3.2 Слепое заполнение

Слепое заполнение представляет собой выбор одного и того же варианта ответа для всех строк табличного вопроса. Так как слепое заполнение можно выявить автоматически во многих онлайн-опросах (большинство анкет содержат несколько подобных таблиц), это довольно распространенный показатель качества полученных данных, прежде всего, для коммерческих онлайн-панелей. Например, специализирующаяся на маркетинговых исследованиях компания GfK разработала программу для выявления таблиц, заполненных «вслепую» [14]. Однако выбор одного и того же варианта ответа для всех строк таблицы не обязательно означает низкое качество ответов. Если все пункты сходны по смыслу (например, все служат индикатором одной скрытой переменной и сформулированы утвердительно), то получение одинаковых ответов даже может считаться хорошим признаком. Купер и соавторы [4] предположили, что слепое заполнение чаще не может использоваться для поведенческих вопросов, чем для оценочных. Мы считаем, что слепое заполнение служит показателем следования принципу разумной достаточности в случае наличия большого списка вопросов, множества вариантов ответа и существенных различий в тематике пунктов списка. Если респонденты прилагают хотя бы малейшие усилия при работе с анкетой, одинаковые ответы становятся невозможными.

В нашем исследовании были две большие таблицы, исключающие возможность слепого заполнения. Первая включала 25 строк (плюс проверка на внимание) с оценочной одиннадцатибальной шкалой, вторая — 25 пунктов с пятибалльной шкалой оценки согласия-несогласия (плюс вариант ответа «не знаю»). Респонденты, отметившие один и тот же вариант ответа для всех N пунктов табли-

цы, определяются как «слепые респонденты». Мы также выделили тех, кто дал один и тот же ответ в N-5 пунктов, в отдельную категорию «почти слепых респондентов». В Таблице № 5 приведены данные по процентному соотношению выявленных «слепых» и «почти слепых» респондентов.

Таблица № 5: Доля «слепых» респондентов и вероятных «слепых» респондентов (в %)

		Контрольная группа		Группа «Ступенчатые вопросы»		Группа «Капча»		Всего
		с тестом	без теста	с тестом	без теста	с тестом	без теста	
Первый блок	N	3,2	2,8	2,4	1,2	0,4**	1,6	1,9
	N-5	6,8	5,2	3,2*	4,0	1,2***	4,8	4,2
Второй блок	N	4,8	6,0	4,0	1,2***	3,2	4,4	3,9
	N-5	12,8	14,8	8,4	7,2***	10,4	13,6	11,2

Примечание: знак*** указывает на $p < 0,01$; ** указывает на $p < 0,05$; * указывает на $p < 0,10$.

В целом, в группах, где были применены разработанные нами функции, случаев слепого заполнения было зафиксировано меньше, хотя разница не всегда статистически значимая. В первой таблице улучшения особенно заметны в группе с использованием функции «Капча» и тестом. Во второй таблице улучшения заметны преимущественно в случае использования метода «Ступенчатые вопросы» без теста.

4.3.3 Открытые нарративные вопросы

Что касается двух представленных в исследовании открытых нарративных вопросов¹, мы оценивали их по следующему принципу: чем больше знаков в ответе, тем он более точный и, следовательно,

¹ В анкете был также третий открытый нарративный вопрос, однако он касался впечатлений от участия в самом опросе. Вполне возможно, что на этот вопрос респонденты из групп, в которых мы тестировали разработанные нами функции, дали более развернутые ответы, так как у них были дополнительные основания дать комментарий. Поэтому здесь мы его не рассматриваем.

тем выше его качество. Кроме того, показателем качества служит процент содержательных ответов, полученных на эти открытые нарративные вопросы. Бессодержательные ответы — это бессмысленные фразы, формулировки, которые не являются ответом на поставленный вопрос, или «простые» ответы, такие как «не знаю» или «ничего».

Значимых различий ни в среднем количестве знаков в ответе, ни в доле значимых ответов выявлено не было. Если говорить о количестве знаков, медианные значения также оказались очень близкими. Подобный результат был получен и применительно к разным типами бессодержательных ответов. Это неудивительно, так как для того, чтобы дать содержательный ответ, респондентам необходима мотивация иного типа, которая, по нашим предположениям, не зависит от побуждения прочесть вопрос. По сути, использованные нами функции никак не мотивировали респондентов к написанию более развернутых ответов на вопросы.

Подводя итоги анализа качества полученных данных, отметим положительное влияние разработанных нами функций на оба теста на внимание и эффект слепого заполнения. В то же время не были зафиксированы различия в ответах на «совместный» вопрос, точности ответов и доле бессодержательных ответов на открытые нарративные вопросы. В целом эти функции повышают качество ответов на некоторые типы вопросов (в частности, на табличные вопросы, а также на вопросы, в которых важная информация приведена во вводной части), однако не на все. Исключение составляют, например, открытые вопросы или вопросы, в которых ключевая информация представлена среди вариантов ответа. Если вопросы представлены в понятном виде, на них легко дать ответ, то респонденты предоставят данные высокого качества. Наоборот, если для ответа на вопросы требуется приложить больше усилий (набор текста при ответе на открытый вопрос) или вопросы представлены таким образом, что приводят респондентов в уныние (огромные таблицы с множеством строк и вариантов ответа), следует ожидать снижения качества ответов. В отношении последнего интерактивность позволяет улучшить полученные результаты.

4.4 Более приятные впечатления от участия?

Мы выдвинули гипотезу, согласно которой, помимо качества улучшатся также и впечатления респондентов от участия в исследовании. Несмотря на то, что обычно онлайн-анкеты разрабатываются по принципу бумажных анкет для самостоятельного заполнения, существуют интернет-инструменты, использующие преимущества динамического интерактивного интерфейса. Все больше респондентов привыкают к опросам такого типа и предпочитают участвовать именно в них.

Для проверки данной гипотезы в конце нашей опросной анкеты мы попросили участников исследования оценить:

- насколько трудными или легкими им показались вопросы, в целом по шкале от «0 — Крайне трудные» до «10 — Крайне легкие»;
- как часто они старались ответить на вопрос наилучшим возможным образом — по шкале от «1 — Никогда» до «5 — Всегда»;
- не показался ли им этот опрос слишком длинным — по шкале: «Да, слишком длинный» (код 2), «Да, довольно длинный» (код 1), «Нет, вполне нормальный» (код 0) и «Нет, даже слишком короткий» (код –1).

Кроме того, мы попросили всех респондентов в группах, где использовали разработанные нами функции, ответить на вопросы об отдельных характеристиках нашей динамической анкеты:

- понравилась ли им интерактивная презентация вопросов — от «0 — Совсем не понравилась» до «10 — Очень понравилась»;
- была ли скорость вывода текста адекватной (символ с цифрой 1, если да, и с цифрой 0, если нет);
- было ли им гораздо сложнее (0) или гораздо проще (10) заполнять анкету в таком формате;
- от «0 — Совсем не помогло» до «10 — Очень помогло».

В Таблице № 6 приведены средние результаты ответов на эти вопросы в разных группах.

Таблица № 6: Впечатления респондентов от участия в опросе для разных условий его проведения (средние значения)

	Контрольная группа		Группа «Ступенчатые вопросы»		Группа «Капча»		Всего
	с тестом	без теста	с тестом	без теста	с тестом	без теста	
Легкость вопросов	7,5	7,5	7,6	7,5	7,8	7,8*	7,6
Наилучшие возможные ответы	4,9	4,8	4,9	4,9	4,9	4,9	4,9
Слишком длинный опрос	0,4	0,4	0,7***	0,7***	0,7***	0,7***	0,6
Понравилась интерактивность	NA		6,0	5,6 ^w	5,6	^{kk} 4,9 ^{ww}	5,5
Адекватная скорость вывода текста			,6	,3 ^{www}	^{kk} ,7	^{kkk} ,5 ^{www}	,5
Легче отвечать			6,0	5,5 ^{ww}	^k 5,6	^k 5,0 ^{ww}	5,5
Помогает сосредоточиться			5,7	5,1 ^{ww}	5,5	^k 4,6 ^{www}	5,2

Примечание: знак*** указывает на $p < ,01$; ** указывает на $p < ,05$; * указывает на $p < ,10$ для тестов на выявление значимых различий между результатами контрольных групп и групп, в которых использовались разработанные нами методы.

Для последних четырех вопросов: знак ^{www} указывает на $p < ,01$; ^w указывает на $p < ,05$; ^w указывает на $p < ,10$ для тестов на выявление значимых различий между результатами групп с тестом и без теста; ^{kkk} указывает на $p < ,01$; ^{kk} указывает на $p < ,05$; ^k указывает на $p < ,10$ для тестов на выявление значимых различий между результатами групп «Ступенчатые вопросы» и «Капча».

Что касается общего впечатления от участия в опросе, никаких различий в оценке вопросов как простых или сложных, а также в уровне качества заполнения анкет выявлено не было. Напротив, в группах, где были использованы разработанные нами функции, большее число респондентов сочло, что опрос был слишком длинным (что вполне логично, он действительно был длиннее), однако никаких различий в результатах между группами с тестом и без теста отмечено не было.

Что касается интерактивных характеристик опроса, средние результаты оказались в целом чуть более позитивными. Это подразумевает, что респонденты хорошо восприняли новые динамические опросные анкеты. Можно также отметить, что респонденты продемонстрировали более позитивную реакцию в тех случаях, когда

проводился тест на подстройку скорости вывода текста (все средние значения значимо выше). Если сравнить функции «Ступенчатые вопросы» и «Капча», последняя без теста оценивается несколько ниже, за исключением вопроса о скорости. В группах с тестом не удалось выявить четкой тенденции к более высокой оценке какой-либо из функций.

5. Выводы

Мы поставили задачу выяснить, могут ли интерактивные анкеты для проведения онлайн-опросов повысить качество получаемых данных и улучшить впечатления респондентов от участия в них без существенного увеличения числа покинувших опрос участников. Сначала мы провели предварительное исследование для выявления проблем и определения дальнейших направлений работы. Затем — лабораторный эксперимент для тестирования наиболее проблемных аспектов. Исходя из полученной информации, мы провели полномасштабное основное исследование, случайным образом распределив респондентов по группам с разными условиями прохождения опроса. Были сформированы контрольные группы и группы, в которых были использованы разработанные нами функции — вывод текста вопросов по блокам («Ступенчатые вопросы») или по отдельным символам («Капча») — при этом скорость появления текста была фиксированной или же подстраивалась по результатам теста на скорость чтения в самом начале анкеты.

Проанализировав результаты основного эксперимента, мы пришли к следующим выводам.

- **H1:** гипотеза о том, что «качество ответов повысится в группах, где используются разработанные нами функции», получила некоторое подтверждение, однако окончательные выводы зависят от учитываемых показателей. Что касается проверки на внимание и эффекта слепого заполнения, улучшения налицо. Однако в случае открытых вопросов и вопросов по типу совместного анализа никаких существенных различий выявлено не было. Это позволяет предположить, что динамичное представление вопросов может побудить респондентов

внимательнее читать инструкции, однако не может мотивировать их на написания более точных и значимых ответов на открытые вопросы.

- **Н2:** гипотеза о том, что «благодаря динамичному представлению вопросов анкеты впечатления от участия в опросе будут оценены более позитивно в группах, где используются разработанные нами функции», не получила достаточного подтверждения. Существенных различий в оценке впечатлений от участия между контрольными группами и группами, в которых были использованы разработанные нами функции, выявлено не было. Исключение составила оценка продолжительности опроса: в группах, где были использованы разработанные нами функции, опрос чаще оценивался респондентами как слишком длинный. При проведении подобных интерактивных опросов в будущем следует создавать более короткие анкеты.
- **Н3:** гипотеза о том, что «доля покинувших опрос респондентов в тех группах, где были использованы разработанные нами функции, окажется выше», получила подтверждение только в группах с тестом. Мы считаем, что необходимо провести повторный эксперимент, чтобы определить, является ли этот неожиданный результат устойчивым или случайным.
- **Н4:** гипотеза о том, что «доля покинувших опрос респондентов в тех группах, где проводился тест на скорость чтения, окажется ниже», не была подтверждена. Однако впечатления от участия в опросе в группах с тестом были оценены выше. Следовательно, подстройку скорости вывода текста с помощью теста можно назвать желательной.

Таким образом, даже учитывая, что, вопреки нашим ожиданиям, полученные результаты не оказались настолько убедительными, можно сделать вывод о том, что практику использования интерактивных опросных анкет следует развивать в будущем. По меньшей мере, это касается использования динамичного вывода текста вопросов в случаях, когда это дает описанные выше преимущества. Динамичный вывод текста вопросов (по символам или по строкам) не только заставляет респондентов не торопиться, но и побуждает

поэтапно обрабатывать поступающую информацию. Это позволяет разделить различные когнитивные этапы чтения, понимания прочитанного и ответа на вопрос, а также заставить респондентов принимать во внимание более значительные объемы представленной информации, не ограничиваясь ее частью. Тестирование на скорость чтения в целях подстройки скорости вывода текста под респондента представляется многообещающей идеей. Несмотря на то, что в целом результаты групп с тестом и без теста оказались схожими, впечатления от участия в опросе были оценены выше в случаях его применения. Разные варианты вывода текста — отдельными блоками («Ступенчатые вопросы») и посимвольно («Капча») — дали одинаковые результаты. В группах с тестом не было выявлено никаких признаков преимущества одной функции перед другой. Следовательно, для сбора данных в подобных ситуациях можно использовать любую из них.

Однако требуется продолжить исследование в целях совершенствования предложенных функций (например, эффективнее подстраивать под респондента скорость вывода текста) и проверки полученных результатов для разных типов опросов, вопросов и тем исследования. Мы уверены, что эти инновационные способы проведения онлайн-опросов необходимо развивать дальше. Мы доказали, что они вполне осуществимы на практике. Респонденты завершили опрос с невысокой долей отказов от участия при том, что показатели качества полученных данных оказались схожими или даже более высокими, чем в обычной ситуации. Онлайн-опросы предлагают невероятные возможности по разработке динамичных и интерактивных анкет, а также по адаптации условий проведения опроса под каждого респондента с учетом его предпочтений или навыков. Эксперименты, связанные с тестированием возможностей инновационных методов проведения онлайн-опросов, должны быть продолжены.

ЛИТЕРАТУРА

1. Bethlehem, J., and S. Biffignandi (2011). *Handbook of Web Surveys* (Vol. 567): John Wiley & Sons.
2. Conrad, F.G. and M.F. Schober (2008). 'New Frontiers in Standardized Survey Interviewing.' Pp. 173–188 in *Handbook of Emergent Methods*, edited by S.N. Hesse-Biber and P. Leavey. New York: Guilford Press.

3. Couper, M.P. (2000). "Web surveys: A review of issues and approaches". *Public Opinion Quarterly*. 64. P. 464–494.
4. Couper, M. P., Tourangeau, R., Conrad, F. G., and C. Zhang (2013). "The Design of Grids in Web Surveys". *Social Science Computer Review*. 31 (3). P. 322–345.
5. Green, P.E., Krieger, A.M., and Y. Wind (2001). "Thirty years of conjoint analysis: Reflections and prospects". *Interfaces* 31 (3): 556–573. doi:10.1287/inte.31.3s.56.9676
6. Green, P.E. and V.R. Rao (1971). "Conjoint measurement for quantifying judgmental data". *Journal of Marketing Research*, 8 (3): 355–363
7. Green, P.E. and V. Srinivasan (1990). "Conjoint analysis in marketing: New developments with implications for research and practice". *Journal of Marketing*. 54 (4): 3–19.
8. Greszki, R., Meyer, M., and H. Schoen (2014). "The Impact of Speeding on Data Quality in Nonprobability and Freshly Recruited Probability-Based Online Panels". In M. Callegaro, R. Baker, J. Bethlehem, A. S. Göritz, J.A. Krosnick & P.J. Lavrakas (Eds.), *Online Panel Research: A Data Quality Perspective* (pp. 238–262). Chichester. UK: Wiley.
9. Kapelner, A. and D. Chandler (2010). "Preventing Satisficing in Online Surveys". In *Proceedings of 2010 Crowd Conf*. 2010.
10. Kieruj, N.D., Mulder, J., Wijnant, A., Douhou, S., and F. Conrad (2013). "Hearing voices: supporting online questionnaires with Text-to-Speech technology". Presented at the European Survey Research Association (ESRA) conference 2013. ESRA. Ljubljana.
11. Krosnick, J.A. (1991). "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys". *Applied Cognitive Psychology* 5:213–36. doi: 10.1002/acp.2350050305
12. Krosnick, J. A., and D.F. Alwin (1987). "An Evaluation of a Cognitive Theory of Response Order Effects in Survey Measurement". *Public Opinion Quarterly*, 51 (2):201–219. Stable URL: <http://www.jstor.org/stable/2748993>
13. Malhotra, N. (2008). "Completion time and response order effects in web surveys". *Public Opinion Quarterly*. 72 (5). P. 914–934.
14. Menold, N., Winker, P., Storfinger, N., and S. Bredl (2014). "Interviewers' Falsifications in Face-to-Face Surveys". In U. Engel, B. Jann, P. Lynn, A. Scherpenzeel & P. Sturgis (Eds.), *Improving Survey Methods: Lessons from Recent Research* (pp. 86–97) .
15. Oppenheimer, D.M., Meyvis, T. and N. Davidenko (2009). "Instructional manipulation checks: Detecting satisficing to increase statistical power." *Journal of Experimental Social Psychology* (2009) 45: 867–872. doi:10.1016/j.jesp.2009.03.009
16. Revilla, M., and C. Ochoa (2015). "What are the links in a web survey among response time, quality, and auto-evaluation of the efforts done?" *Social Science Computer Review*, 2015, Vol. 33 (1) 97–114. First published online on May 14. 2014. doi:10.1177/0894439314531214
17. Schuman, H., and S. Presser (1981). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. New York: Academic Press.
18. Simon, H.A. (1957). *Models of Man*. New York: Wiley
19. Wikipedia. "Speed reading". Section "Claims of speed readers", available at: http://en.wikipedia.org/wiki/Speed_reading#Claims_of_speed_readers
20. Zhang, C. (2013). "Satisficing in web surveys: implications for data quality and strategies for reduction". Doctoral dissertation University of Michigan 2013. Available at: <http://deepblue.lib.umich.edu/handle/2027.42/97990>
21. Zhang, C., and F. Conrad (2014). "Speeding in Web Surveys: The Tendency to Answer Very Fast and Its Association with Straightlining". *Survey Research Methods*, 8 (2). 127–135.

Как влияет приглашение на отклик в онлайн-опросах?

*Мавлетова Айгуль Маратовна,
Малошенок Наталья Геннадьевна,
Терентьев Евгений Андреевич
(Национальный исследовательский университет
«Высшая школа экономики»)*

Введение

Значительная доля онлайн-опросов проводится с помощью реализации рассылки приглашений на электронные адреса респондентов. Обычно электронное приглашение содержит краткое описание исследования и ссылку на онлайн-анкету. Исследователь должен с особой тщательностью подойти к составлению приглашения, поскольку именно на его основании большинство респондентов принимают решение о заполнении анкеты. Как правильно составить приглашение и что в него включить? В данной работе мы постараемся дать практические советы, опираясь на результаты экспериментальных исследований, проведенных российскими и зарубежными исследователями. Можно выделить следующие базовые инструменты привлечения респондентов к опросу:

1. предварительное уведомление о проведении опроса;
2. приглашение к опросу;
3. напоминание об опросе.

Наиболее детально мы рассмотрим, какие факторы необходимо учесть при составлении непосредственно приглашения к опросу. Мы будем опираться на результаты методических экспериментов, измеряющих эффективность разных факторов. Учитывая результаты проведенных экспериментов, исследователи могут подобрать оптимальную комбинацию элементов приглашения для получения более высокого отклика.

Перейдем к последовательному рассмотрению результатов экспериментов. Начнём с предварительного уведомления об исследовании.

1. Предварительное уведомление о проведении опроса

Как правило, большее количество контактов приводит к увеличению отклика в опросах, в том числе и онлайн-опросах. Значимость данного фактора отмечается большинством исследователей. В рамках теории социального обмена влияние этого фактора объясняется созданием дополнительного вознаграждения, повышающего вероятность реципрокного обмена и участия в опросе [19]. Использование предварительных уведомлений повышает чувство морального обязательства со стороны респондентов [5]. Кроме того, предварительное уведомление способствует установлению более устойчивого и менее формального контакта между исследователем и респондентами [55].

Результаты мета-анализа и ряда эмпирических исследований показывают высокую эффективность использования предварительных уведомлений в почтовых и телефонных опросах [14; 22; 33; 45, но есть исключения 24; 30].

Что касается онлайн-опросов, то проведенный мета-анализ также показал эффективность использования предварительных уведомлений о проведении исследования [7]. При этом значимым фактором является способ рассылки предварительного уведомления: электронное уведомление (на электронную почту), почтовое уведомление или SMS-уведомление. Электронное уведомление является менее эффективным в онлайн-опросах. Ряд исследований показали, что предварительное уведомление по электронной почте может повысить отклик на 3–5 процентных пункта, однако влияние оказывалось не всегда статистически значимым [4; 23; 38]. Наиболее эффективными стратегиями является рассылка почтовых или SMS-уведомлений. Данные уведомления в некоторых случаях могут повысить отклик на 10–15 процентных пунктов [1; 4]. Эффекты в других исследованиях были менее существенными — до 3–5 процентных пункта, однако, как правило, являлись статистически значимыми [13; 50; 59].

Таким образом, результаты экспериментов показали эффективность использования предварительного уведомления об исследова-

нии. В зависимости от специфики исследования, а также имеющихся данных о выборочной совокупности исследователь может выслать предварительное уведомление о проведении опроса по почте, электронной почте или SMS-уведомление на мобильный телефон, что позволит привлечь к опросу большее количество респондентов.

Далее мы рассмотрим, на что исследователь должен обратить внимание непосредственно при составлении приглашения к опросу.

2. Приглашение к опросу

В приглашении к опросу мы рассмотрим влияние следующих факторов на отклик респондентов:

- отправитель;
- тема;
- обращение к респондентам;
- ожидаемая длительность заполнения анкеты;
- вознаграждение;
- ссылка на опрос;
- длина приглашения;
- другие факторы: мобильная оптимизация приглашения, обоснование отбора респондентов для участия в опросе, дефайн заполнения анкеты, упоминание о возможности получить результаты исследования.

Отправитель

Важным элементом приглашения к онлайн-опросу является отправитель приглашения. Использование в качестве отправителя человека с более высоким статусом и со статусом, вызывающим доверие, приводит к увеличению отклика. Так, использование в качестве отправителя профессора университета или декана факультета приводит к значительному увеличению участников опроса. В исследованиях Н. Гежен и С. Джэкоб было показано, что использование в качестве отправителя человека со статусом профессора приводит к увеличению отклика по сравнению с использованием отправителя со статусом студента [20]. В общем опросе интернет-пользователей

количество принявших участие в случае получения письма от профессора почти в 8 раз, а в опросе студентов в полтора раза превысило количество участвующих в случае получения письма от студента.

В исследовании А. Мавлетовой, Н. Малошонок и Е. Терентьева [62], проведенном среди студентов, была также обнаружена эффективность приглашения к опросу от имени декана. Исследователи показали, что отклик увеличился вдвое в случае отправления приглашения от имени декана по сравнению с приглашением от имени исследовательского центра.

Однако результаты исследования А. Джойнсона и У. Райпсана не выявили значимого различия в доле откликов между приглашением от имени проректора и приглашением от имени рядового сотрудника университета [31]. При этом статистически значимые различия были обнаружены при взаимодействии фактора отправителя с фактором персонального обращения. Наибольший отклик был зафиксирован среди тех, кто получил персонализированное приглашение от проректора. Наименьший отклик — среди тех, кто получил неперсонализированное приглашение от рядового сотрудника университета.

Другие эксперименты показали, что значимыми могут оказаться как имя отправителя, так и его пол. Так, Н. Гежен, С. Джэкоб и Т. Морино показали, что интернет-пользователи более склонны переходить по ссылке в том случае, если отправителем являлся их тезка. Отметим, однако, что ссылка была не на опрос, а на сайт благотворительной организации [21]. Ф. Койш показал, что среди IT-специалистов (в основном, мужской аудитории) пол отправителя оказался значимым: в случае получения приглашения от лица женского пола отклик был выше на 6 процентных пунктов по сравнению с приглашением от лица мужского пола [38].

Итак, на основе проведенного анализа для повышения отклика в онлайн-опросах мы рекомендуем высылать приглашение на опрос от человека с более высоким статусом и авторитетом. Однако при этом не должны нарушаться этические границы, и человек, от имени которого отправляется приглашение, должен дать согласие на использование своего имени.

Тема

Какую тему следует указать в приглашении к опросу? Одна из возможных стратегий — сформулировать тему в формате просьбы. Например: «Пожалуйста, помогите компании и поделитесь Вашим мнением» вместо «Поделитесь Вашим мнением с компанией». Эффективность данной стратегии была показана в ряде исследований, где увеличение отклика за счет данной стратегии достигало 5 процентных пунктов [25; 56]. М. Капловитц и его коллеги показали, что формулирование темы в формате просьбы от авторитетного лица («Вице-президент просит принять участие в опросе») повышает отклик по сравнению с более общей темой («Прими участие в опросе об охране окружающей среды в кампусе») до 10 процентных пунктов [35]. Однако в некоторых экспериментах данный эффект либо не был обнаружен [49], либо был отрицательным [40].

Р. Кент и Х. Брендел показали неэффективность использования тем, которые респонденты могут спутать с маркетингом или спамом [36]. Так, указание в теме о возможности выигрыша туристической поездки снизило долю откликов на 14 процентных пунктов по сравнению с более стандартной просьбой принять участие в опросе.

А. Мавлетова, Н. Малошенок и Е. Терентьев [62] не обнаружили различия в доле отклика в зависимости от использования одной из двух тем в опросах студентов: «Опрос студентов ведущих вузов» и «Выскажите свое мнение, помогите сделать Ваш ВУЗ лучше». Однако тема «Опрос студентов ведущих вузов» во взаимодействии с персональным обращением к респонденту по имени и отчеству увеличила вероятность участия в опросе в 7 раз. Скорее всего, оба фактора способствовали повышению воспринимаемой ценности участия в опросе.

Сложно сделать однозначный вывод о том, как правильно формулировать тему электронного письма с приглашением к опросу. В том случае, если отправитель письма вызывает у участников доверие, тема письма может не иметь большого значения. Однако в случае неизвестного отправителя тема письма не должна походить на спам или письмо, рекламирующее товары или услуги. В теме не стоит указывать возможные вознаграждение и выигрыш, которые респондент может получить за участие в исследовании. Лучше напрямую

указать, что респондента приглашают принять участие в опросе [8]. Например, тема «Опрос ведущих IT-экспертов» или «Опрос ключевых сотрудников» может подчеркнуть как цель опроса, так и повысить его ценность в глазах участников.

Обращение к респондентам

Одна из эффективных стратегий повышений отклика в онлайн-опросах — обращение к респонденту по имени или имени и отчеству. По мнению Д. Диллмана и его коллег, обращение к респонденту по имени демонстрирует ценность его участия в исследовании [12]. Для реализации данной возможности необходимо иметь базу с актуальными персональными данными респондентов. Результаты ряда методических экспериментов свидетельствуют о том, что персональное обращение к респонденту способно увеличить отклик на онлайн-опрос на 8–10 процентных пунктов [7; 27; 26; 32; 43].

Х. Заурмен и М. Роуч обнаружили большую эффективность неформального обращения по имени по сравнению с более формальным обращением по имени и фамилии среди студентов и аспирантов. Неформальное обращение увеличило отклики в два раза [52]. Подобный результат был также получен в исследовании А. Джойнсона и У. Райпса, однако величина эффекта была менее значительной [31]. Результаты других экспериментов показывают, что иногда формальное обращение может быть более эффективным. Так, в эксперименте Дж. Пирсона и Р. Левина среди выпускников университета уровень отклика была выше в случае формального обращения по фамилии («Уважаемый Мистер Джонсон») по сравнению с обращением по имени («Уважаемый Джон») среди выпускников до 30 лет и среди выпускников старше 50 лет [47]. Формальное обращение повысило долю откликов на 5 процентных пунктов. А. Мавлетова, Н. Малошонов и Е. Терентьев [62] также обнаружили большую эффективность формального обращения по имени и отчеству среди студентов. Формальное обращение почти в два раза увеличило отклик по сравнению с неперсонифицированным обращением по имени. Однако исследование проводилось в двух вузах

и подобный эффект был обнаружен только в одном университете.

В целом использование персонализации является эффективным. Использование формального или неформального обращения должно зависеть от статуса респондентов и знания респондентами организации или исследователей, отвечающих за проведение опроса.

Ожидаемая длительность заполнения анкеты

Респонденту, как правило, необходимо знать, сколько времени займёт заполнение анкеты. В этой связи значимым фактором является информация о длительности анкеты. Если следовать здравому смыслу, можно предположить, что указание сравнительно меньшего времени на заполнение анкеты приведет к увеличению отклика. Эта гипотеза проверялась в ряде эмпирических работ.

Исследование С. Кроуфорда, М. Купера и М. Ламиаса показало, что указание в тексте электронного приглашения сравнительно меньшего времени (8–10 минут), требуемого для заполнения анкеты, увеличило отклики на 4 процентных пункта по сравнению с указанием большей длительности (20 минут) [6]. В среднем заполнение анкеты занимало у респондентов 20 минут. В этой связи указание сравнительно меньшего времени на заполнение анкеты привело к увеличению количества тех, кто прерывает опрос.

Результаты исследования М. Капловица и его коллег показывают, что эффект увеличения отклика при указании сравнительно меньшего времени на заполнение анкеты имеет место не для всех исследуемых аудиторий [35]. Так, указание разной длительности анкеты не привело к значимым различиям в уровне отклика среди преподавателей и административных сотрудников университета и в то же время привело к увеличению отклика среди студентов. Возможно, при принятии решения об участии в исследовании преподаватели и административный персонал имеют другую мотивацию, нежели студенты, и фактор затраченного времени для них является не самым важным.

Некоторые исследователи не обнаружили значимых различий в доле откликов между условиями, в которых указывается разная длительность анкеты [27; 40; 56; 60]. В случае указания меньшей дли-

тельности опроса доля частично заполнивших анкету увеличивается на 5 процентных пунктов. Снижение воспринимаемых затрат на заполнение анкеты приводит к увеличению количества начинающих заполнять анкету, однако осознание того, что заполнение анкеты занимает больше времени, приводит к потере доверия организаторам опроса и дальнейшему отказу от заполнения опроса. Некоторые исследователи обнаружили взаимодействие между указанием большей длительности анкеты и темой опроса: в случае просьбы (помочь университету или компании своими ответами) и указанием большей длительности анкеты количество участников увеличивалась [40; 56].

Можно сделать вывод, что при составлении приглашения не стоит значительно уменьшать ожидаемую длительность анкеты. Хотя в некоторых группах указание сравнительно меньшего времени на заполнение анкеты способствует увеличению количества тех, кто соглашается принять участие в исследовании, оно также приводит к увеличению количества респондентов, не заполнивших анкету до конца. Кроме того, большая длительность анкеты при взаимодействии с другими факторами (например, темой) может, наоборот, повысить долю откликов.

Вознаграждение

Для увеличения отклика на опрос исследователи часто предоставляют вознаграждение или устраивают лотереи. Считается, что лотерея или вознаграждение могут играть позитивную роль при принятии решения об участии в исследовании, выступая в качестве дополнительно подкрепления [16]. Однако результаты исследований, изучавших влияние размера вознаграждения на уровень отклика, достаточно противоречивы. Вознаграждение может быть как гарантированным (для всех участников опроса, заполнивших анкету до конца), так и разыгрываться в формате лотереи (например, 10 призов разыгрываются среди всех участников опроса). Вознаграждение может предоставляться респонденту до или после участия в исследовании. Может быть денежным или предполагать получение сертификата, продукта или услуги.

Результаты проведенных методических экспериментов позволяют говорить о том, что в срезовых исследованиях более эффективным может являться проведение лотереи среди участников опроса [5], в то время как в лонгитюдных исследованиях или онлайн-панелях больший эффект имеет гарантированное вознаграждение [17].

При проведении лотереи значение для респондентов имеет не столько вероятность возможного выигрыша, сколько его размер. В исследовании Х. Заурмана и М. Роача было обнаружено, что небольшой размер вознаграждения (\$5х100) по своему влиянию на отклик не отличается от условий, когда респонденту не предлагается вознаграждение [52]. Однако варианты, в которых 5 респондентов могли выиграть \$100, или 25 респондентов — \$25, позволяли повысить отклик на 6 процентных пунктов. Т. Тутен и её коллеги показали, что безотлагательное уведомление о выигрыше в лотерею за заполнение опроса приводит к увеличению доли откликов на 6 процентных пунктов по сравнению с уведомлением о выигрыше в лотерею через месяц после заполнения анкеты [57]. Таким образом, участники хотят сразу после заполнения анкеты уже знать, выиграли они в лотерею или нет.

Дж. Бирнхольц и его коллеги показали, что вознаграждение наличными является более эффективным вознаграждением по сравнению с вознаграждением в формате подарочных сертификатов [3]. Исследователи сравнили три вида вознаграждения, размер которого был одинаков (\$5): наличные, отправка сертификата на Amazon.com с помощью почтового сервиса, отправка сертификата на Amazon.com на электронную почту. Оплата участия в исследовании наличными приводила к увеличению откликов на 21 процентный пункт по сравнению с использованием сертификатов [3].

М. Педерсен и К. Нильсен обнаружили, что вознаграждение за онлайн-опрос, которое будет переведено на благотворительные цели, менее эффективно чем получение вознаграждения самим респондентом [48].

В срезовых почтовых опросах была показана эффективность денежного вознаграждения, выплачиваемого до участия респондентов в опросе (т.е. вознаграждение заранее получают все, кто попал в выборочную совокупность), по сравнению с вознаграждением, опла-

чиваемым по факту участия респондентов в опросе. Но может ли работать вознаграждение, выплачиваемое до заполнения анкеты, в онлайн-опросах?

Ранние исследования говорили о неэффективности такой стратегии в онлайн-опросах [5]. Однако более поздние эксперименты показали её действенность [41; 42; 45; 59]. Исследования показали, что размер вознаграждения, выплачиваемого до заполнения анкеты, не обязательно должен быть большим. Так, вознаграждение в размере 20 и 50 евро незначительно превысили долю откликов по сравнению с вознаграждением в размере 10 евро. В случае вознаграждения, выплачиваемого участникам до заполнения анкеты, уровень отклика может увеличиться на 10–15 процентных пунктов, увеличив долю участвующих вдвое по сравнению с вознаграждением, выплачиваемым по факту участия в опросе [41; 42; 45].

На основе этого мы можем сделать следующие выводы. Вознаграждение, которое будет потрачено на благотворительность, не является эффективной стратегией в онлайн-опросах. В случае срезовых опросов лотерея может быть эффективной, однако респонденты больше обращают внимание на размер возможного вознаграждения, чем на вероятность его получения. Кроме того, безотлагательное уведомление о возможном выигрыше срабатывает лучше, чем отложенное уведомление. Однако более эффективной стратегией является денежное вознаграждение, оплачиваемое до участия респондентов в опросе. В случае лонгитюдных опросов или онлайн-панелей наиболее эффективной стратегией представляется выплата гарантированного денежного вознаграждения.

Ссылка на опрос

Респондент получает приглашение к участию в онлайн-опросе, в котором ему предлагается перейти по ссылке на анкету, либо дополнительно ввести идентификационный код для доступа к опросу. Расположение ссылки в приглашении, а также ее вид и способ доступа к опросу может повлиять на решение получателя письма об участии в исследовании. С. Кроуфорд, М. Купер и М. Ламиас обнаружили, что отклик на онлайн-опрос увеличивается на 4 процентных

пункта в случае получения индивидуальной ссылки по сравнению со ссылкой на опрос, которая требует дополнительного введения индивидуального кода [6]. Некоторые другие эксперименты, однако, показали противоположные результаты [28; 26].

Расположение ссылки на онлайн-анкету также может влиять на желание респондентов принять участие в исследовании. В методическом эксперименте М. Капловица и его коллег половине респондентов предъявлялись приглашения, в которых ссылка находилась в верхней части приглашения и следовала сразу после обращения к получателю письма, а другой половине рассылались приглашения со ссылкой, размещенной в конце письма после текста с описанием исследования [35]. Вопреки ожиданиям исследователей, которые предполагали, что получателям письма будет легче пропустить текст с описанием исследования и сразу перейти по ссылке, которая находится в начале, больший отклик получило приглашение, в котором ссылка размещалась в конце письма. Доля участников была выше на 6 процентных пунктов. Исследование включало студентов, преподавателей и административных сотрудников вуза. Для первых двух групп различия в отклике оказались статистически значимыми. Данный эксперимент позже был повторен в работе Ф. Барласа и его коллег на меньшей выборке [2]. Они получили другие результаты, а именно, что ссылка, расположенная в верхней части письма, привела к увеличению откликов на 6 процентных пунктов. Э. Хапп и М. Мюррей-Клозе также повторили эксперимент, однако не обнаружили различия между двумя экспериментальными условиями [29].

Итак, в случае электронного приглашения на онлайн-опрос рекомендуется высылать индивидуальные ссылки, которые не требуют дополнительного введения идентификационного кода. Насколько важно найти «правильное» место для размещения ссылки? Эксперименты не дают однозначного ответа. Можно предположить, что место для ссылки может зависеть от ряда факторов. В случае онлайн-панелей или лонгитюдных исследований ссылка может быть расположена в начале приглашения, однако главное размещать ссылку в одном и том же месте в большинстве приглашений. В случае срезовых опросов ссылка должна идти после введения и обоснования,

почему респондент получил ссылку и почему от него ожидают заполнения анкеты.

Длина приглашения

Длина текста приглашения может оказывать влияние на отклик в онлайн-опросах. Логичным выглядит предположение о том, что респонденты не захотят читать большой объем текста. В соответствии с этим многие исследователи рекомендуют делать приглашения короткими, сфокусированными и доступными для быстрого прочтения [12; 44]. Тем не менее результаты эмпирических исследований показывают, что влияние этого фактора является неоднозначным.

Исследование К. Клофстада, Ш. Булиан и Д. Бассон, проведенное среди студентов, преподавателей и научных сотрудников университета, показало отсутствие статистически значимых различий в уровне отклика в зависимости от длины текста приглашения [37]. М. Капловиц и коллеги предприняли аналогичную попытку и сравнили уровень отклика на два типа приглашений: длинное (182 слова) и короткое (80 слов) [35]. Результаты показали увеличение доли откликов на 5 процентных пунктов в случае длинного приглашения среди преподавателей и административных сотрудников. Значимых различий в доле откликов среди студентов обнаружено не было. Исследователи предположили, что в случае длинного приглашения респонденты восприняли исследование как более серьезное и важное.

А. Мавлетова, И. Девятко и Н. Малошонок исследовали влияние длины приглашения среди студентов, преподавателей и административных сотрудников российского университета. Исследователи обнаружили значимое влияние длины приглашения среди студентов и отсутствие значимого влияния среди преподавателей и административных сотрудников [40]. Длина приглашения варьировалась от 95 до 116 слов в экспериментальных группах, получивших короткое приглашение, и от 155 до 205 слов в группах, получивших длинное приглашение (длина варьировалась для студентов, преподавателей и административных сотрудников). Среди студентов, получивших длинное приглашение, отмечался более высокий отклик, хотя увеличение откликов составило всего полтора процентных пун-

кта. Ф. Барлас и его коллеги сравнили приглашение объёмом в 174 слова с объёмом в 386 слов и обнаружили, что короткое приглашение увеличило отклики на 5 процентных пунктов, хотя различие оказалось статистически незначимым [2].

Различие в экспериментах с длиной приглашения зависит от того, какое приглашение считают «коротким», а какое — «длинным». Скорее всего, эффект длины приглашения является параболическим и приглашение до 100 слов можно считать коротким, в то время как объём в 350—400 слов и больше можно считать длинным. Таким образом, оптимальное приглашение может быть объёмом в 150—300 слов, хотя это предположение требует более подробной экспериментальной проверки.

Другие факторы

Мобильная оптимизация приглашения

В связи с ростом мобильного трафика и увеличением мобильных веб-респондентов оптимизируются не только сами анкеты, но и приглашение к опросу. Так, Т. Сондерс и А. Кесслер показали, что отклик увеличивается на четверть среди мобильных пользователей в случае оптимизации приглашений [53].

Хотя мобильная оптимизация приглашений требует некоторых дополнительных затрат со стороны исследователя, скорее всего, мобильная оптимизация приглашений может в скором времени стать стандартом, как уже стала стандартом мобильная оптимизация самих опросов.

Обоснование отбора респондента для участия в опросе

Указание на то, что респондент был специально отобран для участия в исследовании, может оказать положительное влияние на уровень отклика. В рамках теории влияния Р. Чалдини селективность подчеркивает редкость данной возможности, что повышает ценность объекта. В данном случае увеличивается ценность участия в опросе для респондента [19]. Однако в исследовании С. Портера

и М. Виткома было показано, что данный фактор оказывает незначительное влияние на уровень отклика [51]. При этом увеличение отклика на 8 процентных пунктов было обнаружено во взаимодействии этого фактора с указанием даты, до которой можно было заполнить анкету (дедлайна). А. Мавлетова, Н. Малошенок и Е. Терентьев [62] обнаружили повышение доли откликов на 11 процентных пунктов среди абитуриентов-мужчин и отсутствие значимого эффекта среди абитуриентов-женщин в случае обоснования отбора респондентов для участия в опросе.

Таким образом, указание на исключительность отбора респондента для участия в опросе может повысить отклики. Однако это может произойти не во всех группах респондентов.

Указание дедлайна

Указание «дедлайна» (или финальной даты заполнения анкеты) может увеличить отклик, поскольку указывает на ограниченность возможности принять участие в исследовании [19]. Но результаты экспериментов в почтовых [61] и онлайн-опросах [18] не подтверждают наличие значимого эффекта. Можно сказать, что указание дедлайна не повышает отклики в онлайн-опросах.

Упоминание о возможности получить результаты исследования

В некоторых исследованиях респонденты имеют возможность получить краткий отчет о результатах опроса. Считается, что респондентам интересно узнать мнение других людей и сравнить себя с остальными участниками [58]. Однако экспериментальные исследования не обнаружили значимого увеличения откликов [39; 54; 62]. При этом некоторые исследования обнаружили взаимосвязь данного фактора с другими. Так, Б. Маркус и М. Босняк показали, что в случае менее интересной темы возможность сравнить свои ответы с ответами других увеличила отклики [39]. А. Мавлетова и её коллеги обнаружили, что указание дедлайна в сочетании с указанием на важность получения ответов от определенной группы студентов

(студентов 3–4 курсов) привело к статистически значимому увеличению отклика данной группы [62].

Таким образом, для эффективного использования данного фактора необходимо более детально изучить его влияние в сочетании с другими элементами приглашения.

3. Напоминания об опросе

Одним из наиболее эффективных способов повышения отклика в онлайн-опросах является использование напоминаний о возможности участия в исследовании, которые рассылаются неответившим через определенное время после рассылки первого приглашения [9; 11; 34]. На эффективность использования напоминаний могут оказывать влияние как количество напоминаний, так и временной интервал между контактами. Если по поводу первого у исследователей сложился определенный консенсус, основанный на том, что в качестве оптимального количества рассматриваются два или три напоминания [9; 10; 52], то в отношении второго высказываются различные мнения. Рассмотрим дальше, как влияет временной интервал между напоминаниями на отклики в онлайн-опросах.

Временной интервал между напоминаниями

В почтовых опросах Д. Диллман советует рассылать первое напоминание через неделю после рассылки приглашения [9]. Однако Э. Детскенс и его коллеги не нашли значимых различий в уровне отклика между напоминанием, высланным через неделю после приглашения, и напоминанием, высланным через две недели [10]. Аналогичные результаты были получены в исследовании Х. Заурмана и М. Роача [52]. С. Кроуфорд и его коллеги показали, что напоминание, высланное через два дня, более эффективно по сравнению с напоминанием, высланным через пять дней [6]. Отклики были выше на 5 процентных пунктов. В то же время Э. Хапп и М. Мюррей-Клозе значимых различий не обнаружили [29].

Какую стратегию следует использовать и когда высылать напоминания? Выбор исследователя может зависеть от специфики объекта

и временного периода, запланированного на проведение полевых работ. В случае небольшого временного периода или если объектом исследования являются интернет-пользователи, часто выходящие в сеть, есть смысл высылать напоминание в течение 2–5 дней после рассылки приглашения. В случае большого временного периода или если объектом исследования являются интернет-пользователи, редко выходящие в сеть, есть смысл высылать напоминание через неделю или две после рассылки приглашения. Вследствие того, что после второго электронного напоминания уровень отклика почти не увеличивается, дополнительные напоминания можно высылать по почте или SMS.

Изменение текста напоминаний

Для увеличения отклика на опрос Д. Диллман и его коллеги советуют изменять тексты напоминаний по сравнению с текстами приглашений [12]. Изменение текста напоминаний может положительно повлиять на отклики, когда респонденты видят, что исследователи вкладывают время и усилия в выстраивание отношений с участниками опроса. Х. Заурман и М. Роач нашли экспериментальное подтверждение данной гипотезе: тексты напоминаний с изменённым текстом, но содержащим одну и ту же информацию об опросе, что и приглашение, привели к увеличению вероятности участия в опросе на 36% [52]. В то же время в исследовании А. Мавлевой и её коллег значимых различий не было обнаружено [62].

Таким образом, имеющиеся результаты исследований не позволяют сделать однозначных выводов о влиянии изменения текста напоминаний на отклик респондентов. Возможно, апеллируя к разным мотивам участия в опросе в текстах напоминаний, можно повысить отклики, охватывая тех, кто ранее не принял участие в исследовании.

В Таблице 1 отражены результаты нашего обзора.

Таблица 1 Влияние элементов приглашения на долю откликов в онлайн-опросах

1. Предварительное уведомление об исследовании	- Эффективность зависит от <i>способа рассылки</i>. - Уведомление по <i>почте или SMS</i>-уведомление более эффективно, чем уведомление по <i>электронной почте</i>. - По электронной почте: увеличение отклика на 3–5 процентных пункта [4; 23; 38]. - Уведомление по почте или SMS: увеличение отклика от 5 до 10–15 процентных пунктов [1; 4; 13; 50; 59] .
2. Приглашение к опросу	
Отправитель	- Использование в качестве отправителя человека с <i>более высоким статусом и со статусом, вызывающим доверие</i>, приводит к увеличению отклика. - Приглашение от имени декана или профессора более эффективно, чем от студента или рядового сотрудника: отклик может увеличиться от 1.5 до 8 раз [20; 62] . - Может влиять имя и пол отправителя [21; 38] .
Тема	- Формулирование темы в <i>формате просьбы</i>: отклик может увеличиться на 5 процентных пунктов [25; 56] . - Формулирование темы в <i>формате просьбы от авторитетного лица</i> повышает отклик до 10 процентных пунктов [35]. - В некоторых экспериментах эффект темы в формате просьбы не был обнаружен [49], либо наблюдался отрицательный эффект [40] . - Неэффективность использования тем, которые респонденты могут спутать с <i>маркетингом или спамом</i>: указание в теме возможности выигрыша уикенда снизил отклик на 14 процентных пунктов [36] .
Обращение к респондентам	<i>Персональное обращение по имени/имени и отчеству/имени и фамилии</i> к респонденту способно увеличить отклик на онлайн-опрос на 8–10 процентных пунктов [7; 25; 26] .
Ожидаемая длительность анкеты	- Лучше указывать <i>время, которое респондент в действительности может затратить на опрос</i>. - Хотя указание сравнительно меньшего времени может увеличить отклики [6; 35], многие эксперименты не обнаружили значимых различий и показали увеличение доли частично заполнивших анкету [27; 40; 60] .

Вознаграждение	• В <i>срезовых опросах</i> более эффективным может являться проведение <i>лотереи</i> [5], в то время как в <i>лонгитюдных исследованиях</i> или <i>онлайн-панелях</i> больший эффект имеет <i>гарантированное вознаграждение</i> [17]. • При проведении <i>лотереи</i> большее значение для респондентов имеет <i>размер</i> возможного выигрыша, а не его вероятность. Выигрыш \$100 пятью респондентами более эффективен в увеличении отклика, чем выигрыш \$5 ста респондентами [52]. • <i>Безотлагательное уведомление</i> о выигрыше в лотерею более эффективно, чем отложенное уведомление [57]. • Пожертвование от имени респондента на <i>благотворительность</i> не увеличивает отклик [48]. • <i>Оплата вознаграждения наличными vs сертификаты</i>: оплата участия в исследовании наличными увеличила отклик на 21 процентный пункт [3]. • Наиболее эффективным является <i>денежное вознаграждение</i>, выплачиваемое до участия респондентов в опросе. Увеличивает отклик на 10–15 процентных пунктов по сравнению с вознаграждением по факту участия в опросе [1; 4]. • В случае лонгитюдных опросов или онлайн-панелей наиболее эффективная стратегия — <i>выплата денежного вознаграждения по факту участия в опросе</i>.
Ссылка на опрос	• В случае электронного приглашения на онлайн-опрос рекомендуется высылать <i>индивидуальные ссылки</i>, которые не требуют дополнительного введения идентификационного кода.
Длина приглашения	• Результаты исследований не дают однозначного ответа об оптимальной длине приглашения. • Различие в экспериментах по длине приглашения зависит от того, какое приглашение считают «коротким», а какое — «длинным». • Скорее всего, <i>оптимальное приглашение</i> может быть объемом в <i>150–300 слов</i>.
Мобильная оптимизация приглашения	• Отклики увеличиваются на четверть среди мобильных пользователей в случае <i>оптимизации приглашений для мобильных пользователей</i> [53].
Обоснование отбора респондента для участия в опросе	• Указание на исключительность отбора респондента для участия в опросе может повысить отклики. Однако это может произойти не во всех группах респондентов [53; 62].
Указание дедлайна	• Указание дедлайна не повышает отклики в онлайн-опросах [18].

Упоминание о возможности получить результаты исследования	Исследования не обнаружили значимого увеличения откликов, однако возможна взаимосвязь данного фактора с другими [39; 54; 62] .
3. Напоминания об опросе	
Временной интервал между напоминаниями	- В почтовых опросах рассылка первого напоминания через неделю после рассылки приглашения [11]. Однако подтверждения не было найдено в онлайн-опросах [10; 52]. - Напоминание, высланное через два дня, более эффективно по сравнению с напоминанием, высланным через пять дней [6] . - Стратегия может зависеть от специфики объекта и временного периода, запланированного на проведение полевых работ.
Изменение текста напоминаний	Имеющиеся результаты не позволяют сделать однозначных выводов о влиянии данного фактора.

Заключение

В работе мы рассмотрели ряд факторов, оказывающих влияние на отклик в онлайн-опросах. Мы опирались на результаты экспериментальных исследований и постарались дать практические советы исследователям, проводящим или собирающимся проводить онлайн-опросы.

Мы показали, что предварительное уведомление о проведении исследования может увеличить отклик, однако эффективность зависит от способа рассылки. Уведомление по почте или SM-уведомление более эффективно, чем уведомление по электронной почте. При отправлении непосредственно электронного приглашения исследователю необходимо обратить внимание на отправителя — он должен вызывать доверие у респондентов. При этом на отклик может влиять как имя, так и пол отправителя. Тема письма также оказывает значимое влияние на отклик, в особенности, если респонденты не знают отправителя письма. В этом случае тема не должна походить на спам или письмо, рекламирующее товары или услуги. Персональное обращение значительно увеличивает отклик в онлайн-опросе. Однако в случае, если приглашение получено от незнакомой организации или незнакомого лица, респонденту должно быть понятно, откуда у исследователей есть личные данные участников. Вознаграждение

также увеличивает отклик в онлайн-опросах. В срезовых опросах более эффективным может являться проведение лотереи, в то время как в лонгитюдных исследованиях или онлайн-панелях больший эффект имеет гарантированное вознаграждение. При проведении лотереи значение для респондентов имеет не столько вероятность возможного выигрыша, сколько его размер. Денежное вознаграждение более эффективно, чем предоставление сертификатов. Наиболее эффективным является, однако, денежное вознаграждение, выплачиваемое до участия респондентов в опросе.

Длина текста приглашения также оказывает влияние на отклик. На основе экспериментальных данных мы можем предположить, что длина оптимального приглашения варьируется в пределах 150–300 слов, однако такое предположение требует более детальной экспериментальной проверки. В ближайшем будущем исследователям нужно будет обратить особое внимание на мобильную оптимизацию приглашений. Эксперименты показали увеличение количества участников среди мобильных респондентов в случае мобильной оптимизации приглашения.

По другим факторам мы не нашли увеличения отклика, либо не нашли однозначного влияния данных факторов. В некоторых случаях лишь взаимодействие факторов приводило к увеличению отклика. Таким образом, исследователям следует обратить внимание на то, какие совместные эффекты применения двух и более факторов могут привести к увеличению количества респондентов. Несмотря на то, что данный обзор может помочь исследователям в составлении приглашения к онлайн-опросу, для более детального измерения величины эффекта каждого фактора требуется проведение мета-анализа.

Хотя мы постарались придать обзору практическую ценность, данные рекомендации не могут носить универсальный характер. Возможно, наши рекомендации будут работать не во всех совокупностях или будут работать по-разному в зависимости от специфики объекта и проблематики исследования. Несмотря на обилие экспериментов в области дизайна приглашений для онлайн-опросов, следует отметить, что эта область только в начале своего развития: предстоит изобрести еще многие стратегии и проверить их эффективность.

Кроме того, практики использования интернета меняются. Прежде всего, увеличивается количество мобильных пользователей интернета, а значит, количество мобильных респондентов. Возможно, стратегии, которые работают сейчас, перестанут работать, и нужно будет искать новые способы сделать приглашение эффективным для мобильных пользователей. Оказывают ли факторы такое же влияние на мобильных веб-респондентов? Таким образом, хотя описанные рекомендации могут применяться на практике, каждый исследователь должен применять их, учитывая специфику объекта и самого опроса.

ЛИТЕРАТУРА

1. Bandilla W., Couper M.P., Kaczmirek L. The Mode of Invitation for Web Surveys // *Survey Practice*. 2012. 5 (3).
2. Barlas F.M., Falcone A.E., Bellamy N.D., Mack A.R. The Impact of Survey Communications on Response Rates. *Proceedings of the 2013 Federal Committee on Statistical Methodology (FCSM) Research Conference*. 2013.
3. Birnholtz J., Horn D., Finholt T., Bae S. The Effects of Cash, Electronic, and Paper Gift Certificates as Respondent Incentives for a Web-based Survey of Technologically Sophisticated Respondents // *Social Science Computer Review*. 2004. 22 (3). P. 355-362.
4. Bosnjak M., Neubarth W., Couper M. P., Bandilla W., Kaczmirek L. Prenotification in Web-Based Access Panel Surveys The Influence of Mobile Text Messaging Versus E-Mail on Response Rates and Sample Composition // *Social Science Computer Review*. 2008. 26(2). P. 213-223.
5. Bosnjak M., Tuten T.L. Prepaid and Promised Incentives in Web Surveys – An Experiment // *Social Science Computer Review*. 2003. 21(2). P. 208-217.
6. Crawford S. D., Couper M. P., Lamias M. Web Surveys: Perceptions of Burden // *Social Science Computer Review*. 2001. 19. P. 146-162.
7. Cook C., Heath F., Thompson R. L. A Meta-Analysis of Response Rates in Web-or Internet-Based Surveys // *Educational and Psychological Measurement*. 2000. 60 (6). P. 821–836.
8. Couper M. *Designing Effective Web Surveys*. Cambridge University Press. New York. 2008.
9. De Rada V. D. The Effect of Follow-up Mailings on the Response Rate and Response Quality in Mail Surveys // *Quality and Quantity*. 2005. 39 (1). P. 1-18.
10. Deutskens E., De Ruyter K., Wetzels M., Oosterveld P. Response Rate and Response Quality of Internet-based Surveys: An Experimental Study // *Marketing Letters*. 2004. 15(1). P. 21-36.
11. Dillman D. A. *Mail and Internet Surveys: The Total Design Method*. Wiley. New York. 2000.
12. Dillman D.A., Smyth J.D., Christian L.M. *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. New York: John Wiley and Sons. 2009.
13. Dykema J., Stevenson J., Day B., Sellers S.L., Bonham V. L. Effects of Incentives and Prenotification on Response Rates and Costs in a National Web Survey of Physicians // *Evaluation and the Health Professions*. 2011 34 (4). P. 434-447.
14. Fox R. J., Crask M. R., Kim J. Mail Survey Response Rate a Meta-analysis of Selected Techniques for Inducing Response // *Public Opinion Quarterly*. 1988. 52 (4). P. 467-491.

15. Goldstein K. M., Jennings M. K. The Effect of Advance Letters on Cooperation in a List Sample Telephone Survey // *Public Opinion Quarterly*. 2002. 66 (4). P. 608-617.
16. Göritz A. S. Incentives in Web Studies: Methodological Issues and a Review // *International Journal of Internet Science*. 2006. 1 (1). P. 58-70.
17. Göritz A. S. The Long-Term Effect of Material Incentives on Participation in Online Panels // *Field Methods*. 2008. 20(3). P. 211-225
18. Göritz A. S., Stieger S. The Impact of the Field Time on Response, Retention, And Response Completeness in List-based Web Surveys // *International Journal of Human Computer Studies*. 2009. 67. P. 342-348.
19. Groves R. M., Cialdini R. B., Couper M. P. Understanding the Decision to Participate in a Survey // *Public Opinion Quarterly*. 1992. 56. P. 475-495.
20. Guéguen N., Jacob C. Solicitation by E-mail and Solicitor's Status: A Field Study of Social Influence on the Web // *CyberPsychology & Behavior*. 2002. 5 (4). P. 377-383.
21. Guéguen N., Jacob C., Morineau T. What is in a Name? An effect of similarity in computer-mediated communication // *Sensoria: A Journal of Mind, Brain & Culture*, 2010, 6 (2). P. 1-4.
22. Hagget S., Mitchell V. W. Effect of Industrial Prenotification on Response Rate, Speed, Quality, Bias, and Cost // *Industrial Marketing Management*, 1994, 23 (2). P. 101-110.
23. Hart A. M., Brennan C. W., Sym D., Larson E. The Impact of Personalized Prenotification on Response Rates to an Electronic Survey // *Western Journal of Nursing Research*. 2009. 31(1). P. 17-23.
24. Heberlein T. A., Baumgartner R. Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature // *American Sociological Review*. 1978. 43 (4). P. 447-462.
25. Henderson V. Increasing (or Decreasing) Response Rate by Changing the Subject of Email Invitations. AAPOR Annual Conference. 2011. May 12–15. Phoenix. USA
26. Heerwegh D., Loosveldt G. An Evaluation of the Semiautomatic Login Procedure to Control Web Survey Access // *Social Science Computer Review*. 2003. 21. P. 223-234.
27. Heerwegh D., Loosveldt G. An Experimental Study on the Effects of Personalization, Survey Length Statements, Progress Indicators, and Survey Sponsor Logos in Web Surveys // *Journal of Official Statistics*. 2006. 22 (2). P. 191–210.
28. Heerwegh D., Loosveldt G. Web Surveys: The Effect of Controlling Survey Access using PIN Numbers // *Social Science Computer Review*. 20. 2002. P. 10-21.
29. Hupp A., Murray-Close M. E-mail Experiments in Web Surveys. International Field Director's and Technologies Conference (IFD&TC). 2010.
30. Jobber D., Sanderson S. The Effects of a Prior Letter and Coloured Questionnaire Paper on Mail Survey Response Rates // *Journal of the Market Research Society*. 1983. 25. P. 339-349.
31. Joinson A. N., Reips U. D. Personalized Salutation, Power of Sender and Response Rates to Web-based Surveys // *Computers in Human Behavior*. 2007. 23(3). P. 1372-1383.
32. Joinson A.N., Woodly A., Reips U-D. Personalization, Authentication and Self-Disclosure in Self-Administered Internet Surveys // *Computers in Human Behavior*. 2007. 23 (1). P. 275–285.
33. Kanuk L., Berenson C. Mail Surveys and Response Rates: A Literature Review // *Journal of Marketing Research*. 1975. 12(4). P. 440-453.
34. Kaplowitz, M. D., Hadlock, T. D., Levine, R. A Comparison of Web and Mail Survey Response Rates // *Public Opinion Quarterly*. 2004. 68 (1). P. 94-101.
35. Kaplowitz, M.D., Lupi F., Couper M.P., Thorp L. The Effect of Invitation Design on Web Survey Response Rate // *Social Science Computer Review*. 2012. 30 (3). P. 339-349.

36. Kent R., Brandal H. Improving Email Response in a Permission Marketing Context // *International Journal of Market Research*. 2003. 45 (1). P. 489–506.
37. Klostad, C. A., Boulianne, S., Basson, D. Matching the Message to the Medium Results from an Experiment on Internet Survey Email Contacts. // *Social Science Computer Review*, 2008, 26(4). P. 498–509.
38. Keusch, F. How to Increase Response Rates in List-Based Web Survey Samples // *Social-Science Computer Review*, 2012, 30 (3). P. 380–388.
39. Marcus B., Bosnjak M., Lindner S., Pilischenko S., Schütz A. Compensating for Low Topic Interest and Long Surveys. A Field Experiment on Nonresponse in Web Surveys // *Social Science Computer Review*, 2007, 25 (3). P. 372–383.
40. Mavletova, A., Deviatko, I., Maloshonok, N. Invitation Design Elements in Web Surveys—Can One Ignore Interactions? // *Bulletin of Sociological Methodology / Bulletin de Méthodologie Sociologique*. 2014. 123(1). P. 68–79.
41. Messer B., Dillman D. Surveying the General Public over the Internet using Address-based Sampling and Mail Contact Procedures // *Public Opinion Quarterly*. 2011. 75 (3). P. 429–457.
42. Millar M., Dillman D. Improving Response to Web and Mixed Mode Surveys // *Public Opinion Quarterly*. 2011. 75 (2). P. 249–269.
43. Muñoz-Leiva F., Sánchez-Fernández J., Montoro-Ríos F., Ibáñez-Zapata J. Á. Improving the Response Rate and Quality in Web-Based Surveys through the Personalization and Frequency of Reminder Mailings // *Quality & Quantity*. 2010. 44 (5). P. 1037–1052.
44. Nielsen, J. *Designing Web Usability: The Practice of Simplicity*. New Riders Publishing. 1999.
45. Parsons N., Manierre M. Investigating the Relationship among Prepaid Token Incentives, Response Rates, and Nonresponse Bias in a Web Survey // *Field Methods*. 2014. 26 (2). P. 191–204.
46. Parsons R., Medford T. The Effect of Advance Notice in Mail Surveys of Homogeneous Groups // *Public Opinion Quarterly*. 1972. 36 (2). P. 258–259.
47. Pearson J., Levine R. A. Salutations and Response Rates to Online Surveys. The Impact of Technology on the Survey Process. *Proceedings of the Fourth International Conference on Survey and Statistical Computing*. 2003. P. 351–362.
48. Pedersen M.J., Nielsen C.V. Improving Survey Response Rates in Online Panels: Effects of Low-Cost Incentives and Cost-Free Text Appeal Interventions // *Social Science Computer Review*. 2014 DOI: 10.1177/0894439314563916 [online first].
49. Porter S. R., Whitcomb M. E. E-mail Subject Lines and their Effect on Web Survey Viewing and Response // *Social Science Computer Review*. 2005. 23 (3). P. 380–387.
50. Porter S.R., Whitcomb M.E. Mixed-Mode Contacts in Web Surveys: Paper is Not Necessarily Better // *Public Opinion Quarterly*. 2007. 71 (4). P. 635–648.
51. Porter S., Whitcomb M. The Impact of Contact Type on Web Survey Response Rates // *Public Opinion Quarterly*. 2003. 67(4). P. 579–588.
52. Sauermaun H., Roach M. Increasing Web Survey Response Rates in Innovation Research: An Experimental Study of Static and Dynamic Contact Design Features // *Research Policy*. 2013. 42 (1). P. 273–286.
53. Saunders T., Kessler A. Read Me! Click Me! Innovations in Email Invitation Design for Today's Digital World. *Digital Research Conference* 2015.
54. Scherpenzeel A., Toepoel, V. Informing Panel Members about Study Results / *Online Panel Research: A Data Quality Perspective*. Ed. By M. Callegaro, R. Baker, J. Bethlehem, A.S. Göritz, J. A. Krosnick, P. J. Lavrakas. 2014. John Wiley & Sons. Ltd, Chichester. UK.
55. Stafford, J. E. (1966). Influence of Preliminary Contact on Mail Returns // *Journal of Marketing Research*. 1966. 3. P. 410–411.

56. Trouteaud A. R. How You Ask Counts a Test of Internet-Related Components of Response Rates to a Web-Based Survey // *Social Science Computer Review*. 2004. 22 (3). P. 385–392.
57. Tuten T. L., Galesic M., Bosnjak M. Effects of Immediate versus Delayed Notification of Prize Draw Results and Announced Survey Duration on Response Behavior in Web Surveys - An Experiment // *Social Science Computer Review*. 2004. 22. P. 377–384.
58. Tyagi P.K. The Effects of Appeals, Anonymity, and Feedback on Mail Survey Response Patterns from Salespeople. *Journal of the Academy of Marketing Science*. 1989. 17 (3). P. 235–241.
59. van Veen F., Göritz A. S., Sattler S. Response Effects of Prenotification, Prepaid Cash, Prepaid Vouchers, and Postpaid Vouchers: An Experimental Comparison // *Social Science Computer Review*. 2015. DOI: 10.1177/0894439315585074 [online first].
60. Yan T, Conrad F, Tourangeau R., Couper M.P. Should I Stay or Should I Go: The Effects of Progress Feedback, Promised Task Duration, and Length of Questionnaire on Completing Web Surveys // *International Journal of Public Opinion Research*. 2010. 23(2). P. 131–147.
61. Yu J., Cooper H. A Quantitative Review of Research Design Effects on Response Rates to Questionnaires // *Journal of Marketing Research*. 1983. 20. P. 36–44.
62. Мавлетова А.М., Малошонок Н.Г., Терентьев Е.А. Влияние элементов приглашения на увеличение доли откликов в онлайн-опросах // *Социология: 4М*. 2014. № 38. С. 72–95.

Влияние знакомства на результаты рекрутинга респондентов для участия в онлайн-опросах: опыт эмпирического изучения¹

Гаврилов Кирилл Андреевич
(Национальный исследовательский университет
«Высшая школа экономики»)

Постановка исследовательской задачи

Студентов традиционно привлекают в качестве участников опросов и экспериментов, проводимых в социальных науках². Онлайн-опросы едва ли являются исключением: студенты зачастую оказываются «удобными» для решения задач исследователя респондентами³. «Удобство» заключается главным образом в доступности: опросы, особенно экспериментального характера, проводятся среди студентов, у которых тот или иной исследователь ведет курс, что минимизирует усилия по поиску респондентов и снижает затраты, связанные с вознаграждением (распространенная практика — вознаграждение в виде кредита или бонуса к итоговой оценке за соответствующий курс)⁴. В случае, если в качестве респондентов выступает

¹ В данной научной работе использованы результаты, полученные в ходе выполнения проекта No 13–05–0035, осуществляемого при поддержке Программы «Научный фонд НИУ ВШЭ» в 2013–2014 гг.

² Так, контент-анализ ведущих журналов по социальной психологии, публикующих результаты экспериментов, показал, что в 1990–2005 гг. в 91,6% статей с результатами экспериментов использовались выборки студентов [1]; см. также: [2].

³ Впрочем, у нас отсутствуют данные о частоте использования студентов в качестве участников в онлайн-экспериментах в социальных науках. Имеются основания полагать, что для онлайн-исследований доля студентов ниже, чем для экспериментальных исследований офлайн в силу более высокой доступности других категорий респондентов в онлайн-исследованиях и наличия удобных форм рекрутирования (например, посредством онлайн-панелей).

⁴ В данной работе не обсуждаются вопросы валидности выводов, сделанных на основе опросов студентов. См. в этом отношении в частности: [3].

более широкий круг студентов (например, если речь идет о мониторинговых исследованиях), возникают общеизвестные вопросы, связанные с повышением мотивации участников и обеспечением качества полученных данных. В частности, какие имеются способы повышения доли ответивших на приглашение поучаствовать в опросе? В статье предпринята попытка проанализировать влияние фактора знакомства автора приглашения и студента-участника опроса — тема, которая не была затронута в предшествующих исследованиях.

Результаты предыдущих работ и гипотезы исследования

В литературе можно найти обсуждение ряда способов увеличения доли откликов в онлайн-опросах, связанных с особенностями направленного потенциальным участникам приглашения (впрочем, далеко не все способы оказываются достаточно эффективными)⁵. Например:

- персонализация получателя приглашений (обращение по имени) [4; 5; 6; 7; 8];
- персонализация отправителя приглашений (подпись или адрес электронной почты к конкретному лицу, а не учреждения) [4];
- подчеркивание уникальной роли респондента [4; 9];
- подчеркивание некоторых деталей темы приглашения (в частности, ссылка на инициатора исследования) [10; 11; 12];
- уменьшение текста приглашения [10];
- расположение ссылки на анкету [10];
- использование графических элементов в приглашении [13; 14];
- отправление напоминаний [6]⁶.

Нашему исследованию релевантно прежде всего обсуждение характеристик *отправителя* приглашения, в частности, его статуса. Авторы ряда работ исходят из предположения, что если отправителем выступает статусная персона (например, директор подразделе-

⁵ Представляется, что обсуждение величины и способа передачи вознаграждения выходит за пределы нашей статьи, поэтому соответствующие идеи не перечислены в списке.

⁶ Обзор способов увеличения доли ответивших в офлайн- и онлайн-исследованиях см. в [15], только в онлайн-исследованиях — [16].

ния), то такое приглашение будет более эффективно по сравнению со случаем, когда приглашает менее авторитетное лицо (например, административный сотрудник) [4]. Эта идея озвучивается и в работах, посвященных более широкой теме готовности индивидов отвечать на вопросы в принципе: как отмечает Грувс и соавторы, «люди более склонны выполнять просьбу, если она исходит от субъекта, наделенного властью, кого-то, кто был санкционирован обществом на обращение с такой просьбой и на ожидание ее исполнения» [17, р. 482–483]. Результаты эмпирической проверки этого предположения противоречивы. В ряде веб-исследований влияние авторитета приглашающей стороны оказалось положительным, но не существенным [4, р. 586; 7]. В других исследованиях, напротив, выяснилось, что и студенты, и более широкая публика существенно активнее реагируют на приглашение, если оно исходит от статусной персоны [18]. А. Мавлетова и коллеги также обнаружили, что приглашение от имени декана увеличивает вдвое число откликов по сравнению с получением приглашения от исследовательского центра как среди студентов педагогического вуза, так и среди выпускников НИУ ВШЭ [9, с. 87–88].

Следует отметить, что в традиционном почтовом опросе высокий статус отправителя, как правило, положительно влияет на долю откликов, но не так сильно, как другие факторы (например, интересная анкета или наличие вознаграждения). В частности, метаанализ Эдвардса и коллег показал, что если анкеты высылают университет, то это увеличивает долю откликнувшихся в 1,3 раза, а если статусное и известное лицо — в 1,13 раз [19].

Однако едва ли только статус и авторитет — характеристики отправителя, которые могут повлиять на долю откликов. Мы предлагаем сделать акцент на личном знакомстве респондента и автора приглашения. Подчеркнем, что мы говорим именно о личном знакомстве, которое предполагает наличие истории взаимодействия сторон. В этом смысле, используя пример из предыдущих исследований, если приглашение отправляет декан факультета, то по отношению к студентам он будет, несомненно, авторитетной персоной, но далеко не всегда можно говорить о личном знакомстве.

Хотя влияние фактора знакомства кажется очевидным, для понимания возможного механизма воздействия, целесообразно рас-

смотреть некоторые теоретические положения. Для формулировки гипотез, связанных с долей ответивших, чаще всего обращаются к теории социального обмена. Как отмечает Д. Диллман, согласно этой теории действие мотивируется, прежде всего, ожиданием реакции, которая последует от других. Три элемента важны для предсказания того, будет ли совершено действие: вознаграждение, затраты и доверие. Вознаграждение — это некоторое приобретение от активности, затраты — это то, от чего субъект отказывается для получения результата, а доверие — это ожидание, что в конечном счете вознаграждения, связанные с совершенным действием, превысят затраты [20, р. 14]. Логично предположить, что доверие — это ключ к пониманию значимости фактора знакомства, поскольку знакомство с инициатором приглашения, несомненно, снижает риски того, что ожидания от совершения действия не оправдаются. Возможна и иная интерпретация, впрочем, не выходящая за пределы теории социального обмена. Согласно Р. Эмерсону, ключевое отличие этой теории от теории рационального выбора состоит в том, что она имеет дело с *длительными* отношениями обмена [21, р. 341]. Иными словами, если я приглашаю друга в гости, то дело не в ожидании, что он меня также пригласит в будущем; скорее, речь идет о том, что если такого приглашения не последует, то это поспособствует разрушению нашей дружбы. Допуская, что наличие длительных отношений обмена как раз и характеризует личное знакомство, именно эта связь и будет способствовать повышенному отклику на приглашение участвовать в опросе.

Несомненно, что в университетской среде формируются такие сети обмена. И если мы говорим о студентах как участника опроса, то не в последнюю очередь такие отношения устанавливаются между студентами и преподавателями. Таким образом, наша основная гипотеза — что фактор знакомства положительно влияет на готовность участвовать в исследовании — конкретизируется: если опрос инициирует преподаватель, ведущий (сейчас или в прошлом) у приглашаемых студентов семинары и лекции, то это повышает долю откликнувшихся на приглашение. Заметим также, что приглашение знакомым преподавателем способствует установлению доверия между исследователем и респондентом: у последнего не возникает

вопросов, почему к нему обращаются по имени, каким образом найден его адрес электронной почты и пр.

Поскольку часть опрошенных — это студенты-социологи, было решено также изучить влияние другого фактора, который мог оказаться смешивающим — «профессиональная солидарность», или «профессиональный интерес», респондентов. Речь идет о том, что для социологов-исследователей наиболее доступной категорией респондентов являются студенты-социологи, мотивация которых может отличаться от других категорий респондентов. Иными словами, вторая гипотеза состоит в том, что респонденты-социологи будут более активно реагировать на приглашение принять участие в вопросе, особенно в случае, если приглашающий — также социолог.

Если обе гипотезы верны, то наибольшая доля откликов должна наблюдаться в случае, когда студентов-социологов приглашает знакомый им преподаватель социологии.

Дизайн и материалы исследования

Для проверки предположения о влиянии указанных факторов на долю ответивших на приглашение был сконструирован экспериментальный план.

Фактор X («Знакомство с отправителем»). Уровни данного межсубъектного фактора определялись в зависимости от того, кто отправляет приглашение:

- знакомый студенту преподаватель;
- незнакомый студенту преподаватель.

Фактор Y («Профессиональная солидарность»). Уровни данного межсубъектного фактора определялись в зависимости от сочетания специализации отправителя (преподаватель-социолог или нет) и получателя приглашения (студенту факультета социологии или другого факультета):

- наличие солидарности (социолог приглашает социолог);
- отсутствие солидарности (все остальные случаи).

Второй фактор нуждается в пояснении. Мы исходили из предположения, что проверить гипотезу о влиянии профессиональной солидарности можно при варьировании двух переменных (специ-

ализация отправителя и получателя). Иными словами, профессиональная солидарность может возникнуть только в случае, если и получатель, и отправитель — социологи (при условии, что «другие» специализации отправителя / студентов не совпадают). Именно такое сочетание, как ожидалось, будет иметь наивысший уровень откликов согласно гипотезе профессиональной солидарности.

Использованные сочетания факторов представлены в *табл. 1*.

Таблица 1. Используемые сочетания факторов

Фактор 1 (X)	Фактор 2 (Y)	
	наличие солидарности (Y ₁)	отсутствие солидарности (Y ₂)
Знакомый отправитель (X ₁)	X ₁ Y ₁ =1	X ₁ Y ₂ =2
Незнакомый отправитель (X ₂)	X ₂ Y ₁ =3	X ₂ Y ₂ =4

Всем потенциальным участникам исследования направлялось по электронной почте приглашение, в котором варьировалась исключительно подпись (в зависимости от уровня Фактора X). Особенности приглашения:

1. Тема приглашения — «просьба принять участие в опросе» («напоминание об участии в опросе» — для письма с напоминанием).
2. Обращение к респонденту по имени.
3. Длина приглашения — около 450 знаков.
4. Адрес отправителя — личная электронная почта с указанием фамилии и имени (на английском языке).

Полный текст приглашения представлен в Приложении.

Напоминание было отправлено через неделю после первого приглашения. Общая длительность этапа сбора данных — 15 дней.

В качестве исследования, принять участие в котором предлагалось респонденту, был проект автора «Восприятие ценности жизни в ситуации риска: перспективы социологического анализа» [22], в рам-

ках которого была сконструирована анкета. Из данных пилотажа было известно, что среднее время ответа на анкету составляет около 15 минут, о чем респондент был проинформирован в приглашении.

Эмпирическая база исследования

Для проведения эксперимента была составлена база данных из 253 студентов 2-го курса двух факультетов НИУ ВШЭ:

- факультета социологии (132);
- Международного института экономики и финансов (МИЭФ), направление «Экономика» (121).

Приглашения рассылались от имени трех преподавателей:

- профессор-социолог, ведущий курс в МИЭФ;
- преподаватель-социолог, ведущий курс на факультете социологии;
- преподаватель (не социолог, сотрудник Института транспортной политики НИУ ВШЭ), не ведущий курс ни на одном из факультетов, где проводился опрос.

Соответственно, профессор-социолог выступал в качестве знакомого инициатора опроса для МИЭФ, для студентов-социологов он был незнакомым преподавателем. Роль преподавателя-социолога была обратной (знакомый для студентов-социологов и незнакомый для МИЭФ). Роль третьего преподавателя можно рассматривать в качестве контрольной, т.к. он не был знаком студентам и не был социологом.

Распределение количества разосланных приглашений представлено в табл. 2.

Всего на приглашение в течение 15 дней ответило 123 студента (49%). Среди откликнувшихся на просьбу поучаствовать в опросе 83 респондента (67%) заполнили анкету полностью, т.е. доля респондентов, которые прошли опрос до конца, составила 33%.

Медианное время заполнения анкеты — 8,5 минут; если исключить респондентов, не заполнивших анкету до конца, этот показатель составил 12 минут.

Таблица 2. Количество разосланных приглашений

Отправитель	Принадлежность получателя к факультету		
	Всего	Социология	МИЭФ
профессор-социолог, ведущий курс в МИЭФ	66	35 (незнакомый)	31 (знакомый)
преподаватель-социолог, ведущий курс на факультете социологии	93	48 (знакомый)	45 (незнакомый)
преподаватель (не социолог), не ведущий курс ни на одном из факультетов	94	49 (незнакомый)	45 (незнакомый)

Результаты

Было проанализировано влияние выделенных факторов на долю откликов, т.е. долю респондентов, которые отреагировали на просьбу и воспользовались ссылкой⁷. В целом разница в долях откликнувшихся на приглашения различных отправителей отсутствует: отреагировали на просьбу профессора-социолога 59% опрошенных, преподавателя-социолога — 47%, а преподавателя (не социолога) — 43%⁸. В то же время разница между профессором-социологом и преподавателем (не социологом) имеется: уровень откликов на просьбу последнего статистически значимо меньше⁹. Отчасти это указывает на наличие факторов, не учтенных в нашем исследовании: если наши гипотезы были бы верны, то наибольший уровень отклика должен был бы быть в ответ на просьбу преподавателя-социолога, а не преподавателя, ведущего занятия в МИЭФ. Одно из объяснений состоит в личных и профессиональных каче-

⁷ Согласно Американской ассоциации исследователей общественного мнения (AAPOR), речь идет о показателе Response Rate 2 (RR2) [23, p. 34–35].

⁸ $\chi^2(2) = 4,3; p = 0,11$

⁹ $\chi^2(1) = 4,2; p = 0,04$

ствах профессора-социолога, которые мы не контролировали.

В то же время проведение дисперсионного анализа¹⁰ показало наличие значимого эффекта для фактора «Знакомство с отправителем» ($F(1, 252) = 4,8; p = 0,03$): если приглашение отправлял знакомый преподаватель, то доля ответивших составила 58%, тогда как если просьба исходила от незнакомого лица — 44%. Эффект профессиональной солидарности отсутствовал — $F(1, 252) = 0,2; p = 0,66$: отклик студентов-социологов на приглашения представителя своей дисциплины составил 49%, тогда как всех остальных — 48%. Также отсутствовало взаимодействие между указанными факторами.

Разница между факультетами отсутствует: социологи откликнулись в 49% случаях, студенты МИЭФ — 48%¹¹. Однако рассмотрение факультетов в отдельности позволяет обнаружить разные тенденции. Хотя социологи более активно реагируют на незнакомого профессора-социолога (54%) и знакомого преподавателя социолога (54%), этот показатель статистически не отличается от доли ответивших в случае, если приглашает незнакомого несоциолог (41%)¹². В МИЭФ студенты чаще отвечали на приглашение знакомого им профессора (65%), нежели незнакомых преподавателя-социолога (40%) и несоциолога (44%), хотя это различие имеет невысокий уровень значимости¹³. Если же сравнить отклик на приглашение знакомого им профессора (65%) с общим откликом на незнакомого приглашающего (42%), то это различие будет статистически значимым¹⁴.

Итак, данные указывают на то, что знакомство с инициатором приглашения действительно имеет положительное влияние на ве-

¹⁰ Для проведения дисперсионного анализа зависимая переменная (отклик) кодировалась как дихотомическая (0 — нет ответа, 1 — ответ присутствует), что позволило рассчитывать среднее значение.

¹¹ $\chi^2(1) = 0,04; p = 0,8$

¹² $\chi^2(2) = 2,2; p = 0,3$

¹³ $\chi^2(2) = 4,8; p = 0,09$

¹⁴ $\chi^2(1) = 4,6; p = 0,03$

личину доли ответивших, однако это наблюдается только в МИЭФ. Не исключено, что спецификой МИЭФ (например, мотивацией студентов, обучающихся в этом институте) или особенностями знакомого преподавателя (например, высокими рейтингами конкретного преподавателя)¹⁵ можно объяснить тот факт, что при *отсутствии* профессиональной солидарности (т.е. если не социолога приглашает социолог) студенты активнее реагируют на просьбу знакомого им преподавателя (табл. 3).

Таблица 3. Доля ответивших (в %) и значимость различий

Фактор 1 (X)	Фактор 2 (Y)		Значимость различий по строке
	наличие солидарности (Y ₁)	отсутствие солидарности (Y ₂)	
Знакомый отправитель (X ₁)	54	65	<i>n.s.</i>
Незнакомый отправитель (X ₂)	46	42	<i>n.s.</i>
Значимость различий по столбцу	<i>n.s.</i>	0,03	

К сожалению, дизайн нашего исследования не позволил проанализировать влияние изучаемых факторов на долю респондентов, полностью заполнивших анкету, т.е. на Response Rate 1 (RR1) в терминологии Американской ассоциации исследователей общественного мнения. Причина — содержательные задачи исследования требовали использования нескольких наборов вопросов (типов анкет), их предъявление оказалось неслучайно распределено между отправителями приглашения. Иными словами, обнаруженные различия по RR1 могли быть обусловлены не заявленными факторами, а различиями в том, какие вопросы были предъявлены. По этой же самой причине мы не могли использовать абсолютный

¹⁵ В то же время данные, позволяющие напрямую сравнить мотивацию студентов разных факультетов или сопоставить рейтинги различных преподавателей, у нас отсутствуют. Эти факторы требуют дальнейшего изучения.

показатель времени, затраченного респондентом на заполнение анкеты¹⁶.

Вместо этого для дополнительной проверки полученных результатов мы использовали другой индикатор — доля респондентов, потративших на заполнение анкеты больше времени, чем медианное значение для данного набора вопросов. Всего было два типа анкет. Для первого медианное время заполнения анкеты составило 11 минут, для второго — 7 минут. Если респондент получал первый (второй) тип анкеты и тратил больше 11 (7) минут, то он считался более «вдумчивым».

Итак, мы исходили из предположения, что оба фактора — знакомства и профессиональной солидарности — также положительно влияют на продолжительность заполнения анкеты, в нашей интерпретации — на вероятность того, что респондент будет вдумчиво заполнять анкету. Результаты частично соответствуют этой гипотезе: дисперсионный анализ показал влияние фактора знакомства¹⁷: из студентов, начавших заполнять анкету, сделали это вдумчиво 61% в случае их приглашения знакомым лицом и 44% — незнакомым. Влияние фактора профессионального интереса оказалось скорее незначимым¹⁸.

Разница реакций на конкретного инициатора опроса, а также зависимости от факультета отсутствует. «Вдумчивость» социологов статистически не зависит от того, кто отправлял приглашения. Однако среди студентов МИЭФ были обнаружены значимые различия в зависимости от инициатора приглашения¹⁹: они тратили больше времени на анкету, если приглашение было получено от знакомого им профессора (75%) в отличие от незнакомого преподавателя (47%)²⁰. Впрочем, следует отметить, что разница в реакциях на про-

¹⁶ Именно этот показатель часто используют для характеристики качества полученных результатов. См., например: [12; 24].

¹⁷ $F(1, 122) = 0,96; p = 0,05$. При этом наблюдаемая мощность равна 0,5 (при $\alpha = 0,05$).

¹⁸ $F(1, 122) = 0,7; p = 0,09$.

¹⁹ $\chi^2(2) = 6,7; p = 0,03$

²⁰ $\chi^2(1) = 4,1; p = 0,04$

фессора-социолога (75%) и незнакомого преподавателя-социолога (61%) отсутствовала²¹.

Обсуждение и заключение

В целом полученные результаты свидетельствуют о том, что знакомство респондента с лицом, отправляющим приглашение принять участие в опросе, может положительно сказаться на доле откликов в онлайн-исследовании: в нашем случае эта доля увеличилась в 1,3 раза. Вероятно, произойдет увеличение и времени, затрачиваемого респондентом на заполнение анкеты, что косвенно указывает на рост качества полученных данных. Этот вывод, если будет подтвержден в ходе дальнейших исследований, имеет немаловажное значение для планирования онлайн-опросов в вузах.

В то же время следует помнить, что обнаруженный нами эффект связан прежде всего с реакцией студентов МИЭФ на приглашение, исходящее от преподавателя, который недавно вел у них занятия. Это обстоятельство накладывает серьезные ограничения на полученные выводы. К числу причин, которые могли способствовать более высокому отклику на просьбу знакомого профессор-социолога среди студентов МИЭФ, можно отнести то, что он:

- продолжал вести занятия у респондентов на момент опроса, тогда как преподаватель-социолог на факультете социологии вел занятия у студентов в предшествующем учебном году²²;
- имеет более высокий профессиональный статус (доктор наук, профессор, тогда как преподаватель-социолог был кандидатом наук, а преподаватель (несоциолог) не имел степени);

²¹ $\chi^2(1) = 0,8; p = 0,36$

²²Студенты МИЭФ были как минимум в ситуации большей зависимости от знакомого преподавателя: им предстояло сдавать ему экзамен. Такая интерпретация вполне соответствует теории социального обмена: участие в опросе — хотя бы подсознательно — могло рассматриваться как нечто обмениваемое позже на вознаграждение (например, оценку по курсу).

- предположительно более популярен среди своих студентов²³.

Наши данные также позволяют сформулировать предположение, что студентам-социологам, в отличие от МИЭФ, не важно, кто их приглашает: знакомый или незнакомый преподаватель. Возможная интерпретация этого эффекта состоит в том, что участие в опросах воспринимается студентами-социологами как рутинная практика, так что они придают меньшее значение фигуре инициатора опроса.

Необходимо также принимать во внимание следующие ограничения проведенного исследования:

- маленькая выборка;
- невозможность контролировать коммуникацию между респондентами с одного факультета;
- отсутствие различий в аффилиации инициаторов приглашений: все они — сотрудники НИУ ВШЭ, что делает невозможным изучение влияния корпоративной солидарности на долю ответивших.

По-видимому, в будущих исследованиях едва ли все эти ограничения могут быть преодолены. Тем не менее именно воспроизведение полученных результатов на другой выборке и с другими отправителями приглашений (в том числе — представителей другого вуза или другой организации), позволит говорить о надежности сделанных выводов. Так или иначе, в качестве промежуточного вывода можно утверждать, что стратегия, согласно которой исследователи-социологи приглашают своих профильных студентов к участию в опросе, — вероятно, не самая эффективная с точки зрения максимизации доли ответивших.

²³ Отсутствуют данные, позволяющие напрямую сравнить рейтинги двух преподавателей. Однако имеются косвенные свидетельства более высокого рейтинга у преподавателя МИЭФ: он последние 3 года признавался «Лучшим преподавателем» по итогам студенческого голосования, тогда как преподаватель факультета социологии не был «Лучшим преподавателем» ни разу.

Благодарности

Автор выражает признательность проф. В.Г. Ледяеву и А.Н. Воробьеву за помощь в проведении эмпирической части исследования, а также проф. И.Ф. Девятко, С.Г. Давыдову и коллегам по НУГ «Качество данных в онлайн-исследованиях» за замечания и предложения, высказанные в ходе подготовки текста.

ЛИТЕРАТУРА

1. Henry P.J. College sophomores in the laboratory redux: Influences of a narrow data base on social psychology's view of the nature of prejudice // *Psychological Inquiry*. 2008. Vol. 19 (2). PP. 49–71.
2. Sears D.O. College Sophomores in the Laboratory. Influences of a Narrow Data Base on Social Psychology's View of Human Nature // *Journal of Personality and Social Psychology*. 1986. Vol. 51 (3). PP. 515–530.
3. Druckman J., Kam C. Students as Experimental Participants: A Defense of the 'Narrow Data Base' (July 6, 2009). URL: <http://ssrn.com/abstract=1498843> (date of access: 15.11.2014).
4. Porter S., Whitcomb M. The Impact of Contact Type on Web Survey Response Rates // *The Public Opinion Quarterly*. 2003. Vol. 67, No. 4. PP. 579–588.
5. Heerwegh D. Effects of Personal Salutations in E-Mail Invitations to Participate in a Web Survey // *The Public Opinion Quarterly*. 2005. Vol. 69, No. 4. PP. 588–598.
6. Muñoz-Leiva F., Sánchez-Fernández J., Montoro-Ríos F., Ibáñez-Zapata J.A. Improving the response rate and quality in Web-based surveys through the personalization and frequency of reminder mailings // *Quality and Quantity*. 2010. Vol. 44 (5). PP. 1037–1052.
7. Joinson A.N., Reips U.D. Personalized Salutation, Power of Sender and Response Rates to Web-Based Surveys // *Computers in Human Behavior*. 2007. Vol. 23. № 3. P. 1372–1383.
8. Sauermann H., Roach M. Increasing web survey response rates in innovation research: An experimental study of static and dynamic contact design features // *Research Policy*. 2013. Vol. 42. P. 273–286.
9. Мавлетова А.М., Малошонов Н.Г., Терентьев Е.А. Влияние элементов приглашения на увеличение доли откликов в онлайн-опросах // *Социология*. 4М. 2014. № 38. С. 72–95.
10. Kaplowitz M., Lupi F., Couper M. and Thorp L. The Effect of Invitation Design on Web Survey Response Rates // *Social Science Computer Review*. 2012. Vol. 30, PP. 339–349.
11. Porter S., Whitcomb M. E-mail Subject Lines and Their Effect on Web Survey Viewing and Response // *Social Science Computer Review*. 2005. Vol. 23. PP. 380–387.
12. Callegaro M., Kruse Y., Thomas M., Nukulkij P. The Effect of Email Invitation Customization on Survey Completion Rates in an Internet Panel: A Meta-analysis of 10 Public Affairs Surveys. 2009. URL: http://www.knowledgenetworks.com/GANP/docs/jsm2009/The_Effect_of_Email_Invitation_Customization_JSM_2009_submitted.pdf (date of access: 15.07.2015).

13. Porter S., Whitcomb M. E-Mail Contacts: A Test of Complex Graphical Designs in Survey Research // *Social Science Computer Review*. 2004. Vol. 22. PP. 370–376.
14. Downes-Le Guin T., Baker R., Mechling J., Ruylea E. Myths and realities of respondent engagement in online surveys // *International Journal of Market Research*. 2012. Vol. 54. Iss. 5. P. 1–21.
15. Edwards P., Roberts I., Clarke M., DiGuseppi C., Wentz R., Kwan I., Cooper R., Felix L., Pratap S. Methods to increase response to postal and electronic questionnaires (Review) // *Cochrane Database of Systematic Reviews*. 2009. Iss. 3.
16. Fan W., Yan Z. Factors affecting response rates of the web survey: A systematic review // *Computers in Human Behavior*. 2010. Vol. 26. P. 132–139.
17. Groves R.M., Cialdini R.B., Couper M.P. Understanding The Decision to Participate in a Survey // *The Public Opinion Quarterly*. 1992. Vol. 56. No. 4. P. 475–495.
18. Guéguen N., Jacob C. Solicitation by E-mail and Solicitor's Status: A Field Study of Social Influence on the Web // *CyberPsychology & Behavior*. 2002. Vol. 5. № 4. P. 377–383.
19. Edwards P., Roberts I., Clarke M., DiGusepi C., Pratap S., Wentz R., Kwan I. Increasing Response Rates To Postal Questionnaires: Systematic Review // *British Medical Journal*. 2002. Vol. 324, No. 7347. P. 1183–1185.
20. Dillman D.A. Mail and Internet Surveys: The Tailored Design Method. Second edition. Hoboken: Wiley, 2007.
21. Emerson R.M. Social Exchange Theory // *Annual Review of Sociology*. 1976. Vol. 2. P. 335–362.
22. Гаврилов К.А. Целостность и ценность сообщества в исследованиях риска // *Социологический журнал*. 2013. № 3. С. 92–113.
23. American Association for Public Opinion Research. Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys. Revised 2008. URL: http://www.aapor.org/AAPORKentico/AAPOR_Main/media/MainSiteFiles/Standard_Definitions_07_08_Final.pdf (date of access: 01.08.2015).
24. Malhotra N. Completion time and response order effects in web surveys // *Public Opinion Quarterly*. 2008. Vol. 72. No. 5. P. 914–934.

Приложение

Текст использованного приглашения

Тема: просьба принять участие в опросе
Здравствуйте, Николай!

Обращаюсь к вам с просьбой поучаствовать в небольшом опросе, который проводится в рамках исследовательского проекта НИУ ВШЭ, посвященного принятию решений в условиях риска.

Пожалуйста, заполните онлайн-опросник, это займет у вас совсем немного времени — около 15 минут.

Анкета размещена по адресу:

<http://www.research.net/s/vld18?c=103660>

Спасибо заранее!

С уважением,

Гаврилов К. А.

доцент кафедры анализа

социальных институтов НИУ ВШЭ

Раздел 3

Методология онлайн-исследований: построение выборок

Нужно ли оставлять выборку на волю случая?

Дэбра Сантус
(*Sample Strategies TNS Research International*)

Питер Квок
(*TNS Research International*)

Франк Келли
(*Lightspeed GMI*)

Как при построении выборки проконтролировать переменные, на которые не установлены квоты

Предпосылки

Переход от вероятностной к неслучайной выборке был очень трудным для большинства исследователей. Индустрия исследований сейчас вовлечена в настолько же трудный переход: от неслучайной квотной выборки с широкими возможностями выравнивания для обеспечения репрезентативности к квотным выборкам, собираемым при помощи роутеров, с гораздо меньшим числом переменных, по которым выборку можно выравнивать. В то время как переход от набора случайных номеров (выборка RDD для телефонных опросов) к онлайн-панелям происходил медленно и с большой осторожностью, нынешний переход прогрессирует быстро, и у исследователя часто нет понимания того, какие именно изменения происходят.

В течение многих лет в крупных компаниях, занимающихся исследованиями рынка, считалось стандартом квотировать приглашения в выборку на основе таких ключевых демографических переменных, как географический регион, размер семьи, пол и возраст. Также в расчет часто принимались другие важные переменные, такие как наличие детей, расовая/этническая принадлежность, состав семьи и образование, но в самом опросе не использовались жесткие квоты. При таком подходе соблюдались определенные требования к ключевым переменным, в то же время соблюдался разумный баланс других демографических переменных.

Определенные изменения в индустрии исследований вынуждают нас адаптировать подходы к построению выборки. Появление роутеров, направляющих поток респондентов сразу на опрос, привело к тому, что выравнивание выборки производится на основе всего нескольких квотируемых переменных, а не как часть тщательно спланированного и приспособленного к ответам респондентов процесса. Быстрый рост выборок с динамическими источниками и поточных выборок (из источника сразу в опрос) в сочетании с технологией роутера создали ситуацию, когда балансирование переменных не всегда доступно или возможно только для части выборки. Использование мобильных устройств для сбора данных дополнительно усложнило процесс приглашения к опросам, т.к. многие анкеты не адаптированы для заполнения на мобильных устройствах. Устройство, используемое человеком, часто влияет на то, к каким опросам он приглашается.

На сегодняшний день в большинстве программных сред, предназначенных для сбора данных, используются алгоритмы, целью которых является привлечение максимального числа респондентов к опросу. Технология роутера редко учитывает какие-либо дополнительные переменные, кроме возраста, пола и региона, при заполнении опросных квот, так как выборки, формируемые роутерами, не предназначены для конкретных исследований (иными словами, роутер задает респонденту несколько соцдем вопросов, а потом приглашает его к любому находящемуся в системе проекту, к которому он подходит). Таким образом, у исследователя нет возможности удостовериться в том, что другие неконтролируемые или используемые

для выравнивания выборки переменные будут репрезентативными. Нужно учитывать важность этих неконтролируемых переменных, чтобы быть уверенными в том, что мы формируем корректную выборку для нашего проекта.

Выбор подходящего метода квотирования является критичным, если мы хотим быть уверенными, что наши выборки позволяют обеспечивать точные, воспроизводимые и высококачественные результаты. Принимая это во внимание, мы сфокусировали дизайн нашего исследования и анализ его результатов на важности установки квот как при приглашении к опросу, так и на завершённые интервью в отличие от стандартного подхода, когда квотируются только завершённые интервью. Создание исходящей выборки, которая сбалансирована с использованием соответствующих переменных и далее скорректирована с учетом уровня отклика, позволяет быть уверенными, что группа респондентов, заполняющих опрос, будет соответствовать целевой аудитории.

Ниже приведены определения для тех типов квот, которые мы разбираем:

Квоты для выравнивания выборки: квоты задаются на той фазе исследования, когда определяется выборка. Мы также называем такое квотирование балансировкой, дизайном приглашений в выборку или переменными, используемыми при построении выборки. Основным источником для квотирования выборки обычно являются данные переписи населения, однако для квотирования могут использоваться и иные установочные исследования (например, данные полстера Pew Research или Департамента контроля и предупреждения заболеваний (CDC) в США), а также квоты, предоставленные клиентом. Квотирование выборок было стандартным подходом с момента появления первой access-панели, созданной компанией NFO ещё в 1940-х годах.

Опросные квоты: квоты, добавленные в опрос, чтобы определить демографический профиль или ответы на определенные вопросы в анкете. Стандартные квоты — это обычно регион, возраст и пол, к которым могут добавляться дополнительные переменные согласно требованиям конкретного исследования.

Квоты для выборки являются фундаментальными для обеспечения репрезентативности нашей выборки. Есть несколько обстоятельств, которые стоит принимать во внимание при использовании квотных выборок в вашем исследовании:

- некоторые группы населения труднее найти, чем другие, поэтому на приглашения к опросу и заполненные анкеты нужно поставить квоты; это обеспечит репрезентативность исследования;
- границы и дизайн выборки могут не отражать профиль целевой аудитории из-за особенностей таргетирования или других характеристик исследования;
- процент откликов на приглашения к опросу значительно отличается в зависимости от демографического профиля респондентов, поэтому квотирование может быть единственным способом репрезентировать социальные группы с низким откликом;
- использование квот может повлиять на стоимость и длительность проекта;
- требования клиента к репрезентации определенных групп могут потребовать использование квот.

Дизайн исследования

Данное исследование было специально спланировано таким образом, чтобы позволить всем приглашенным респондентам принять участие. Это позволяет обеспечить отсутствие смещения, связанного с достижимостью или особенностями отбора. Также респондентам было позволено принимать участие в опросе на любом устройстве, включая компьютер, мобильный телефон и планшет.

Дизайн и характеристики исследования включают:

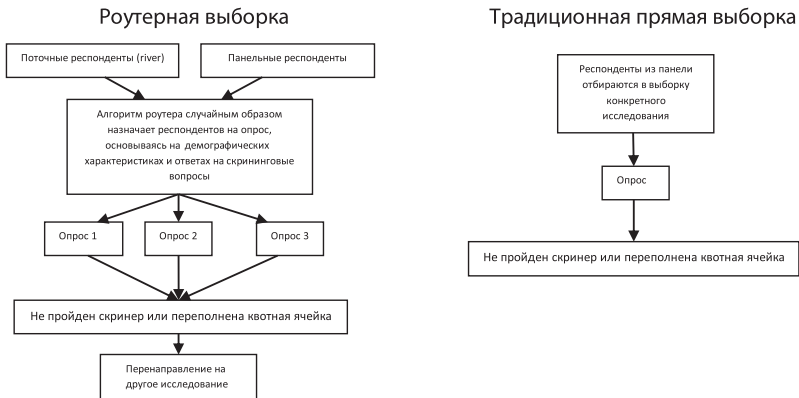
- время проведения опроса — 8 дней в начале декабря 2014 г.;
- длительность интервью составляла 10–0 мин.;
- простая анкета была посвящена покупке повседневных товаров и включала:
 - вопросы, связанные с новыми технологиями покупки продуктов;

- одну небольшую таблицу;
- открытые вопросы отсутствовали;
- 100% достижимость (то есть все респонденты подходили для участия, скринера не было);
- опрос можно было проходить на любом устройстве;
- в качестве источника респондентов использовалась онлайн-панель My Survey компании Lightspeed.

Дизайн выборки

Как было отмечено ранее, за последние несколько лет подходы к построению выборки значительно изменились. Сегодня мы можем встретить два основных типа выборки: выборка при помощи роутера (в условиях общего недостатка респондентов после прохождения одной анкеты респондент автоматически приглашается в другой опрос из числа активных, к которому он подходит по социально-демографическим характеристикам) и классическая прямая выборка (по приглашению). В случае с роутерами респонденты не отбираются специально для конкретного исследования, поэтому управление квотами происходит на уровне распределения трафика респондентов, а не при помощи плановой рассылки приглашений. В отличие от описанного, традиционный подход предполагает отбор респондентов для конкретного опроса, когда мы каждый раз ставим квоты на приглашения к опросу и на заполненные интервью. метода респонденты отбираются для конкретного исследования, квоты в выборке распланированы, а квоты в опросе устанавливаются на основе одного исследования. Для данного исследования мы использовали традиционный подход к построению выборки.

Краткое описание каждого процесса выборки представлено ниже:



Подход к дизайну данного исследования включал в себя два проекта (ячейки) с прямой выборкой. Каждая ячейка приводилась в соответствие с данными переписи населения США для лиц старше 18 лет. В качестве переменных выступали географический регион, семейный доход, размер домохозяйства и связанные пол и возраст.

Приведенная ниже таблица представляет демографический профиль населения США.

Перепись США			
Регион		Возраст опрошенных	
Северо-восток	18.1%	Мужчины 18–29	10.9%
Средний Запад	21.3%	Мужчины 30–39	8.3%
Юг	37.1%	Мужчины 40–49	8.7%
Запад	23.4%	Мужчины 50–59	8.9%
Семейный доход		Мужчины 60+	11.6%
<30 000\$	23.3%	Женщины 18–29	10.7%
30 000\$–<50 000\$	17.8%	Женщины 30–39	8.5%
50 000\$–<75 000\$	18.4%	Женщины 40–49	9.0%
75 000\$+	40.5%	Женщины 50–59	9.4%
Размер домохозяйства		Женщины 60+	14.1%
1 человек	14.2%		
2 человека	34.1%		
3 и более человек	51.8%		

Исследование было разделено на две ячейки с различными квотами для выборки. В Ячейке 1 мы контролировали все переменные на уровне выравнивания квот/приглашений к опросу; в Ячейке 2 для выравнивания выборки использовался только регион и связанные половозрастные квоты. В обеих ячейках опросные квоты (на заполненные анкеты) были установлены на одни и те же переменные.

	Ячейка 1		Ячейка 2	
Переменные	Квоты выборки	Квоты опроса	Квоты выборки	Квоты опроса
Регион	✓	✓	✓	✓
Доход	✓			
Размер домохозяйства	✓			
Связанные возраст и пол	✓	✓	✓	✓

В данном исследовании мы изучали эффект взвешивания на основе сравнения двух ячеек. Ячейка 1 после прескринирования и сбора данных показала намного более высокую эффективность при взвешивании и оказалась более устойчива к недостаточному количеству доступных данных. Ячейка 2 была сбалансирована только по некоторым первичным параметрам, что является типичным для современной исследовательской практики.

Чтобы оценить сходство, мы использовали «случайный интерактивный метод» взвешивания (англ. RIM weighting) для приведения наших данных к данным переписи США. Случайный интерактивный метод взвешивания считается наиболее эффективным для проверки разброса выборки. Данные приводились в соответствие пропорционально по ячейкам для каждой из перечисленных переменных. Для упрощения веса применялись к предельным, а не к связанным распределениям.

Приведенная ниже таблица отражает различия в подходах к взвешиванию для каждой ячейки по параметрам региона, дохода, размера домохозяйства и возраста/пола. Эффективность взвешивания для Ячейки 1 составила 95.9%, тогда как в Ячейке 2 она оказалась значительно ниже — 76.4%. Обнаружен значительный разброс между группами населения с различными доходами и базовой демографической выборкой.

Взвешивание

Данные опроса были взвешены относительно данных переписи США.

Для повышения точности приведения выборочных данных к данным переписи США был использован «случайный интерактивный метод» взвешивания (RIM weighting)

- данные пропорционально приводились к целевым;
- для упрощения веса применялись к предельным, а не к связанным распределениям;
- подход показал свою эффективность для проверки разброса выборки.

Джон Бремер (2013)

«Наиболее эффективное использование взвешивания — это взвешивание на этапе построения выборки».

Анализ определенных групп населения в ячейках позволяет глубже понять разброс выборки для контролируемой и неконтролируемой демографии.

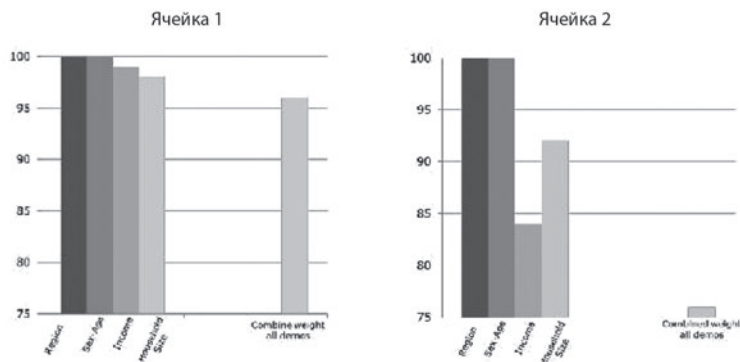
Общие результаты — Эффективность взвешивания

	Веса		
	Эффективность	Минимальный	Максимальный
	взвешивания	вес	вес
Ячейка 1	95,9%	0,66	1,68
Ячейка 2	76,4%	0,39	3,42

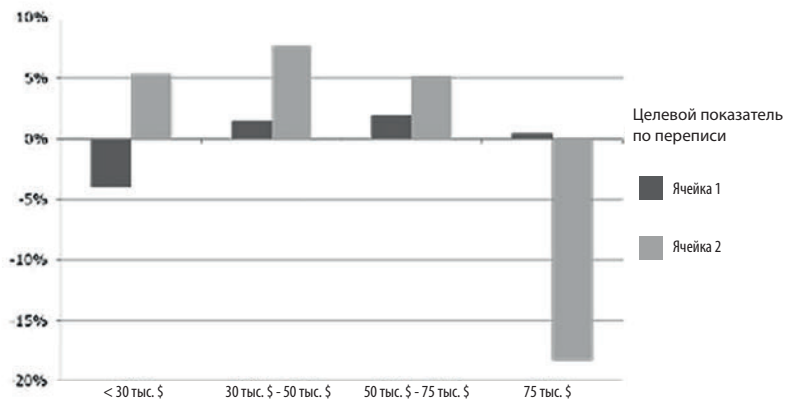
Основные нормы эффективности взвешивания:

- > 75% эффективности;
 - Значение 3 или меньше для весов.
- ✓ 100%-я эффективность взвешивания достигается, когда пропорции по всем взвешенным переменным в выборочной совокупности полностью совпадают с населением.
- ✓ Чем ниже эффективность весовой обработки, тем менее надежны данные.

Эффективность выборки по взвешенным переменным внутри ячеек

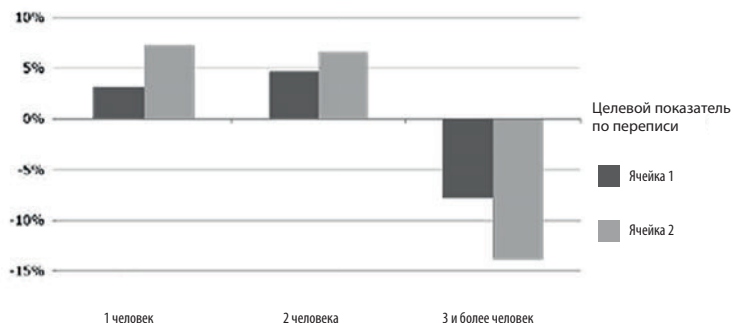


Детальный демографический анализ – Доход

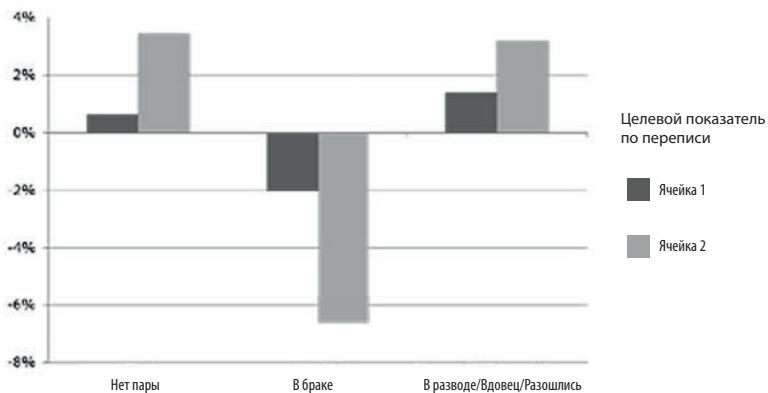


В Ячейке 2 содержались значимые различия с данными переписи до взвешивания

Детальный демографический анализ – Размер домохозяйства



Детальный демографический анализ – Семейное положение



Мы добавили в нашу анкету несколько вопросов, связанных с покупкой продуктов питания. Различия между Ячейками 1 и 2 были очевидны; однако после взвешивания некоторые из различий свелись к минимуму, в то время как другие остались без изменений.

Различия недемографических переменных

	Данные до взвешивания		Данные после взвешивания		Различие между взвешенными и невзвешенными данными	
Данные	Ячейка 1	Ячейка 2	Ячейка 1	Ячейка 2	Ячейка 1	Ячейка 2
Есть ли у Вас смартфон?	72%	68%	72%	72%	0%	+4%
Как Вы оплачиваете покупки?						
Наличные	16%	19%	16%	17%	0%	-2%
Кредитная карта	32%	27%	31%	31%	-1%	+4%
Другое	53%	55%	54%	52%	+1%	-3%
Состоите ли Вы в программах лояльности магазинов?						
Costco	27%	23%	27%	28%	0%	+5%
Sam's Club	24%	23%	24%	25%	0%	+2%
Не состою в программе лояльности	44%	49%	45%	42%	+1%	-7%
Используете ли Вы смартфон/ноутбук/ компьютер/планшет при покупке продуктов?	40%	34%	39%	37%	-1%	+3%

Хотя в опросе можно задать квоты для целого ряда различных переменных, их трудно будет набирать, и это может повлечь увеличение сроков и стоимости проекта. В наиболее типичных случаях мы наблюдаем квотирование по связанным полу и возрасту вместе с еще одной дополнительной переменной, например, регионом. При выходе за рамки такой квотной структуры получают очень мелкие ячейки, которые трудно заполнить. В результате, в случае с динамически формируемыми выборками и в частности с роутерными выборками, мы обычно наблюдаем квотную структуру, подобную Ячейке 2.

Переход от случайной вероятностной выборки к подходам с использованием неслучайной выборки происходил в течение десятилетнего периода и породил обширные исследования. Переход от того, что мы называем традиционной прямой выборкой, к роутерной выборке произошел за период всего в несколько лет, и мы

не нашли значимых исследований на эту тему. В настоящий момент роутерная выборка является *de facto* нормой в индустрии исследований, однако мы не уверены, что побочные эффекты, которые сопровождают данный подход к формированию выборки, были приняты осознанно. Необходимо найти способ контроля неквотируемых переменных, особенно если файлы данных не подвергались взвешиванию, как можно наблюдать в нашем исследовании.

Мы предлагаем участникам рынка провести необходимые исследования, чтобы определить наиболее значимые переменные для различных категорий товаров и услуг. Мы протестировали подход, при котором из копии исходной выборки удалялись респонденты с определенными демографическими параметрами, такими как высокий доход. Если мы удалим из выборки половину группы с самыми высокими доходами, как это повлияет на эффективности взвешивания для данной выборки? Чтобы определить переменные, которые сильнее всего влияют на качество выборки (измеряемое с помощью эффективности взвешивания), мы провели такие тесты для целого ряда переменных в нашем исследовании. Возможно, что таким путем мы сможем определить одну или две переменных, помимо типичных возраста/пола/региона, которые можно контролировать в роутерной выборке.

Данный подход является проблематичным, т.к. каждый источник выборки, используемый в роутере, имеет определенный набор доступных демографических переменных. Во многих ситуациях вы можете захотеть использовать такую переменную, как размер домохозяйства, но обнаружите, что эти данные доступны только для части респондентов. Будете ли вы ограничивать используемую выборку только той выборкой, в которой доступны требуемые демографические данные? Или вы зададите нужный вопрос в соцдем блоке анкеты, чтобы процесс отбора осуществлялся в самом опросе? В этом случае эффективность роутера значительно снизится.

Неслучайная выборка взвешивается по некоторому количеству переменных так, чтобы ее распределение соответствовало населению или другому эталону. Предполагается, что другие переменные, не использованные при взвешивании, более или менее не зависят от использованных переменных. Если это предположение не работает, тогда переменные, не участвующие в процессе взвешивания,

могут влиять на значительные смещения выборки. Исходная мысль заключается в том, чтобы включить все интересующие нас переменные в процесс взвешивания и контролировать все отклонения. Однако при таком подходе размер выборки должен быть достаточно большим, чтобы гарантировать отсутствие пустых ячеек для взвешивания. Окажется, что легко набираемые квотные ячейки будут переполнены, в то время как для трудных параметров выборка будет недостаточной. Это приводит к тому, что выборка не будет сбалансирована с населением, а в весовых коэффициентах будет наблюдаться существенный разброс. Процесс сбора данных становится дорогим, в то время как качество данных падает.

Как обсуждалось ранее, одно из предлагаемых решений состоит в том, чтобы предварительно отсеивать респондентов по некоторым из ключевых переменных, по которым происходит взвешивание. В этом и состоит смысла роутерной выборки, которая похожа на традиционную двухфазовую выборку, за исключением того, что финальная выборка строится в процессе, а не по завершении первой фазы. Если предположить, что ни один паттерн ответов не соотносится значимо с этими переменными для взвешивания, то финальный набор респондентов может еще до начала поля быть задан так, чтобы симитировать распределение населения. Таким образом, различия в весовых коэффициентах будут небольшими, а стоимость исследования будет ниже при сохранении качества данных. Остается большой вопрос, как это может повлиять на переменные, не используемые при взвешивании.

В дополнение к «случайному интерактивному методу» взвешивания (RIM weighting) мы провели проверку чувствительности весов с целью дальнейшего анализа различий в таблицах. В таблице ниже мы учли три сценария:

Сценарий 1: Все ожидаемые случаи были сохранены.

Сценарий 2: Только 25% ожидаемых случаев высокого дохода были сохранены. Все случаи из классов с другим уровнем дохода были сохранены.

Сценарий 3: Только 25% ожидаемых онлайн-покупателей были сохранены. Все случаи, когда респонденты не совершали покупок в интернете, были сохранены.

Мы проверили эффективность выборки, чтобы увидеть эффект распределения весов, а также процент использования купонов во взвешенных и невзвешенных данных. Результаты представлены в таблице ниже:

Сценарий	Показатель	Ячейка 1 (Сбалансирована по 4 Переменным, включая Доход)	Ячейка 2 (Сбалансирована по 2 Переменным, исключая Доход)
Все ожидаемые случаи	Эффективность выборки	95,93%	76,38%
	Взвешенное использование купонов	74,08%	70,93%
	Невзвешенное использование купонов	74,34%	69,96%
25% случаев ожидаемого высокого дохода (среднее на 1000 повторений)	Эффективность выборки	61,97%	31,46%
	Взвешенное использование купонов	74,06%	71,08%
	Невзвешенное использование купонов	72,02%	69,29%
25% случаев ожидаемых покупателей в интернете (среднее на 1000 повторений)	Эффективность выборки	94,49%	70,51%
	Взвешенное использование купонов	68,61%	64,74%
	Невзвешенное использование купонов	69,28%	65,38%

Когда отсутствующие случаи соотносятся с переменной, использованной для взвешивания (как во втором сценарии), снижение эффективности выборки можно быть серьезным (более 30 процентных пунктов, когда 75% ожидаемой подгруппы не представлено). Когда отсутствующие случаи не соотносятся или слабо соотносятся с переменными, по которым происходит взвешивание (как в случае третьего сценария), снижение качества выборки может быть не таким существенным.

Разница между ожидаемыми взвешенными и невзвешенными показателями не превышает двух процентных пунктов. Наиболее сбалансированная группа стабильно превосходила по качеству минимально сбалансированную группу, т.к. влияние ключевые метрики

было менее значительное. Когда отсутствующие случаи соотносятся с переменной (доход), которая не влияет на использование купонов, падение оценок в обеих группах практически не наблюдается. Однако, когда отсутствующие случаи соотносятся с переменной (покупки в интернете), которая влияет на использование купонов, в наиболее сбалансированной группе отклонения были меньше, чем в минимально сбалансированной группе.

Мы показали, что может не применять взвешивание, если ожидаемая погрешность не превышает двух процентных пунктов. Более того, предварительное балансирование выборки по ключевым переменным делает ее более устойчивой к отсутствующим случаям. Следует предупредить, что не все отсутствующие случаи одинаковы: некоторые из них могут не изменять значения, в то время как другие оказывают влияние, которое нельзя игнорировать.

Предложения на будущее

Киш однажды сказал: «Неравенства в рамках выборок должны быть измерены, если они не были исправлены в процессе отбора. Если они влияют на небольшую долю выборки или отличаются только по маленькому фактору, их общий эффект может быть измерен и впоследствии им можно пренебречь, но только если размер или фактор не велик» [Киш, 1990].

Это как раз те обстоятельства, когда польза от неслучайной выборки перевешивает ее стоимость. В нашей работе мы показали, что такие обстоятельства могут возникнуть и при проведении маркетинговых исследований. Мы не налагали ограничений на то, кто может участвовать в наших опросах, однако у некоторых покупателей, которые могли бы подойти для нашего исследования, не было шанса принять в нем участия. Мы не пытались исследовать эффект недостаточного покрытия методики рекрутирования участников.

Наоборот, в нашей выборке мы намеренно снизили число респондентов в крупных подгруппах, чтобы увеличить недостаток доступных случаев. Мы увидели, что, по крайней мере, в нашем примере, даже в такой катастрофической ситуации, когда всего 25% ожидаемого размера подгруппы участвует в исследовании, ключе-

вые значения не обязательно отличаются на значимую величину. Это помогает сохранить тенденции по ключевым показателям.

Мы также показали, что в некоторых обстоятельствах предпочтительно использовать неслучайный дизайн выборки. Возможное будущее направление исследований — отдельно рассматривать данные обстоятельства и исследовать их ограничения.

Заключение

Итоговый результат нашего исследования показал, что по демографическим переменным, использованным для взвешивания, Ячейка 1 больше соответствует данным переписи населения, чем Ячейка 2. Контроль исходящей выборки по ключевым социально-демографическим показателям оказался действенным способом увеличить эффективность взвешивания без использования дополнительных переменных для взвешивания, таких, как владение домом или семейное положение. В некоторых случаях дизайн исходящей выборки для Ячейки 1 минимизировал необходимость взвешивания данных после заполнения анкет.

Дальнейшие исследования должны проводиться с использованием большего числа различных источников выборки, включая динамическую выборку, и анализировать результаты и последствия управления квотами.

ЛИТЕРАТУРА

Kish L. (1990) "Weighting: Why, When, and How?" Proceedings of the Section on Survey Research Methods, Alexandria, VA: American Statistical Association. p. 121–130.

Применение выборки, направляемой респондентом, в онлайн-исследованиях: методологические вызовы и исследовательские перспективы

Моисеев Станислав Павлович
(*Национальный исследовательский университет*
«Высшая школа экономики»)
Савинкова Юлия Константиновна
(*ООО «Ипсос»*)

Введение

Благодаря развитию информационных технологий проблемы, которые осложняли проведение онлайн-исследований пятнадцать лет назад, на сегодняшний день потеряли свою актуальность: доступ в интернет стал значительно проще, существенно возросла скорость и надежность интернет-соединения, пользователи получили доступ к различным браузерам и т.д. Тем не менее возможности применения онлайн-исследований по-прежнему ограничены рядом методологических затруднений. В частности, к таким затруднениям можно отнести отсутствие общепринятого выборочного метода [1, р. 874].

Начиная с 40-х годов XX века эталонной моделью выборки в социальных исследованиях являлась вероятностная выборка. Переход исследований «в сеть» заставил ученых задуматься о способах, которые позволяют приблизиться к данному эталону или добиться получения эквивалентных результатов в онлайн-исследовании. Одним из них является применение выборки, направляемой респондентом, или Respondent-Driven Sampling (RDS). RDS представляет

собой разновидность сетевой выборки, которая позволяет оценивать параметры генеральной совокупности на основе специальной математической модели. Возможность использования RDS для проведения онлайн-опросов была представлена научному сообществу около семи лет назад. В программной статье «WebRDS» характеризовалась как высокопродуктивная и эффективная выборочная методика, результатом которой являются неточные, но обоснованные оценки параметров генеральной совокупности [2, р.105]. За прошедшее время зарубежные и отечественные исследователи провели ряд экспериментов и методологических исследований по применению RDS и ее онлайн-разновидности. Эти результаты стали основой для серьезной научной дискуссии, которая продолжается и сегодня [3, с.67]. Часть авторов занимает оптимистичную позицию и отмечает, что WebRDS является перспективным методом отбора респондентов в интернет-исследованиях [4; 5; 6]. Их оппоненты указывают на сложности в применении выборки, направляемой респондентом, а главное, подчеркивают то, что условия, необходимые для корректной оценки параметров генеральной совокупности далеко не всегда реализуются на практике [7; 8; 9]. Такие противоречия указывают на то, что методика WebRDS нуждается в доработке и требует дополнительной методологической экспертизы.

Данная статья ставит перед собой задачу зафиксировать актуальное состояние научной дискуссии вокруг возможностей применения выборки, направляемой респондентом. В силу того, что критика базовых допущений выборки, направляемой респондентом, имеет равное значение для использования RDS в традиционных и онлайн-исследованиях, мы будем опираться на корпус литературы, посвященной RDS в целом. Рассуждение о методологических вызовах и исследовательских перспективах WebRDS начинается с описания предыстории появления WebRDS и обсуждения ее методологических допущений. Далее в работе рассматриваются критические замечания о возможности соблюдения данных требований в реальной исследовательской практике. Наконец, в статье проводится сравнительных анализ попыток применения WebRDS, которые предпринимались некоторыми зарубежными и отечественными авторами.

Методика и базовые допущения WebRDS

Для того чтобы правильно понимать специфику WebRDS, необходимо учитывать, что выборка, направляемая респондентом, разрабатывалась как метод изучения труднодоступных групп. Первоначально для их исследования использовались выборка снежного кома, выборка ключевого респондента и таргетированная выборка [10, р. 174–176]. По сравнению с ними RDS обладала рядом преимуществ: она обеспечивала анонимность респондентов, позволяла снизить смещения, связанные с отбором более лояльных участников, а главное — оценить размер и характеристики социальной группы. С учетом этих свойств, а также острой необходимости в информации о труднодоступных группах RDS получила достаточно широкое распространение: к 2013 году она применялась для исследований в десятках стран по всему миру [9].

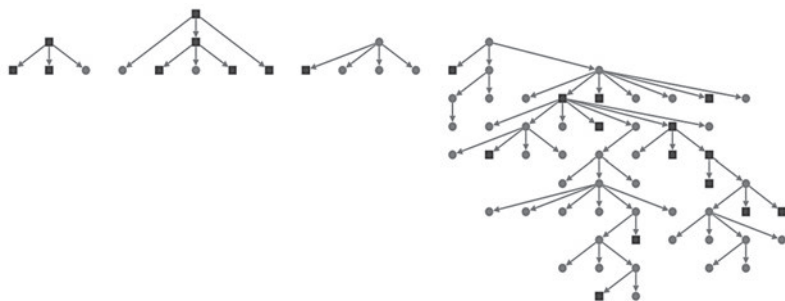
Мотивация переноса технологии RDS в область онлайн-исследований была связана с потенциальной возможностью получать достоверные оценки параметров различных совокупностей интернет-пользователей при сохранении низкой стоимости и высокой скорости опроса. При этом переход на использование средств интернет-коммуникации стал практически единственным принципиальным изменением в методике WebRDS по сравнению с RDS [см. 2].

Одно из главных отличий RDS от традиционной выборки снежного кома состоит в том, что ее реализация основана на использовании двойной системы вознаграждения: каждый участник получает средства не только за собственное участие, но и за рекрутирование своих знакомых и друзей. Согласно теоретико-методологической модели RDS такое решение позволяет задействовать вторичные мотивирующие стимулы, основанные на внешнем социальном контроле. В дополнение к этому респондент не должен раскрывать личности знакомых исследователю, что имеет важнейшее значение для закрытых социальных групп и должно стимулировать кооперацию и открытость респондентов из любого другого сообщества [9, р. 178–179].

Согласно принятой методике построение RDS начинается с целенаправленного отбора определенного числа первичных ре-

спондентов, так называемых «семян» [10, р. 179]. Каждый из них получает задание пригласить в исследование ограниченное¹ количество своих друзей и знакомых. Процедура приглашения обычно основана на использовании уникальных купонов (или других идентификационных данных), которые рекрутер передает следующему респонденту. Процесс набора респондентов продолжается до достижения целевого объема выборки или прохождения точки равновесия — момента, после которого характеристики конечных единиц наблюдения становятся независимыми от характеристик семян. Это становится возможным благодаря представлению цепочки опрашиваемых как марковской цепи [7, с. 38]. Такой принцип позволяет рассматривать процесс рекрутирования с точки зрения математической модели случайного блуждания, в которой каждый новый шаг отбора является независимым от предыдущего.

Рисунок 1. График компонента сети WebRDS, исследование студентов НИУ ВШЭ [8]. Квадраты на графе обозначают студентов мужского пола; круги — женского.



¹ Обычно речь и идет о 3-х респондентах. Такое ограничение призвано сократить смещение, связанное с наличием в выборке очень популярных респондентов, и направлено на удлинение цепочек приглашения [2, р.115].

Прямым результатом применения RDS является направленная сеть², состоящая из респондентов, связанных друг с другом отношениями рекрутирования (Рисунок 1). На основании данных о значимых критериях отбора (например, пол, возраст, национальность), собираемой информации о числе связей респондентов в изучаемой группе и информации о диадах «рекрутер-респондент» строится вероятностная модель полученной сети. Таким образом, проверка статистических гипотез о распределении изучаемых параметров в социальной группе проводится на основе сравнения реальной выборочной совокупности с заданным числом случайных.

Возможность реализации такой статистической операции требует соблюдения ряда базовых допущений относительно процесса отбора и характеристик сети изучаемой группы. Их состав и группировка могут незначительно варьироваться в зависимости от используемой модели оценки параметров. В Таблице 1 представлен перечень допущений, который принят для модели оценки, предложенной Дугласом Хекаторном и Эриком Вольцем [9].

В целом наличие базовых допущений является стандартным условием для построения определенной статистической модели. Однако представленная совокупность допущений, необходимых для корректной реализации выборки, направляемой респондентом, вызвала скептическую реакцию у многих исследователей. Следствием этой реакции стало появление ряда методологических исследований, которые ставили под сомнение и проверяли реалистичность тех или иных допущений и апробировали методику реализации RDS.

Критика базовых допущений RDS

Рефлексивный настрой исследователей, которые проводили эксперименты, связанные с реализацией RDS, коснулся практически всех базовых допущений методики. Тем не менее часть из них не получила серьезной критической оценки. Во многом это связано с тем, что они не могли пройти строгую методологическую проверку

² Под «направленной» имеется в виду сеть, в которой для каждой связи известен ее источник и реципиент.

Таблица 1. Базовые допущения модели Хекаторна-Вольца³

Цель \ Допущения	О сети	Об отборе
Обеспечение модели случайного блуждания	Достаточно большой размер сети	Возможность повторного отбора респондента
		Единая, не ветвящаяся цепочка отбора
Обеспечение независимости семян (т.е. состояния равновесия)	Достаточно слабое влияние эффекта гомофилии	Достаточное количество волн рекрутирования, необходимое для достижения равновесия
	Минимальное число «узких мест», т.е. сложных, маловероятных путей, ведущих к определенной группе респондентов	
	Принадлежность всех членов группы к одному компоненту сети*	
Обеспечение предполагаемого поведения респондентов	Реципрокность связей в сети	Точные оценки числа социальных связей всеми членами группы
		Случайность отбора в процессе рекрута респондентов

* Под «компонентом сети» подразумевается совокупность членов группы, в который каждый член группы связан хотя бы с одним другим. Т.е. такой компонент не имеет изолированных элементов.

и выйти за предел аксиоматических утверждений. В первую очередь данное замечание справедливо для допущения о «достаточно большом размере сети». Исследования, в которых приходится применять RDS, зачастую отличаются тем, что группа ученых не имеет возможности получить предварительные оценки размера генеральной совокупности и, более того, ее социально-демографической структуры. В отсутствие этих параметров утверждение о том, что размер сети целевой группы является достаточно большим для получения каче-

³ Адаптированная версия таблицы из статьи «Diagnostic for Respondent Driven Sampling» [9, p. 4].

ственных данных, может носить исключительно гипотетический характер. Единственным основанием, указывающим на справедливость такой гипотезы, служат косвенные данные или имеющиеся экспертные оценки. Похожая логика работает и для допущения о том, что представители исследуемой группы принадлежат к одному компоненту сети. В большинстве случаев исследователи могут лишь предполагать, что наличие той или иной характеристики группы объединяет всех ее членов между собой и не оставляет изолированных элементов. В частности, такой характеристикой может быть использование схожих социальных медиа (Вконтакте, Facebook, Twitter и др.), принадлежность к корпоративной сети и т.п. Тем не менее проверить справедливость такой гипотезы на практике трудно, так как для этого может потребоваться проведение сплошного исследования.

Оттакиваясь от имеющихся возможностей, исследователи сфокусировали свои усилия на тестировании ряда проверяемых допущений и смогли добиться определенных результатов в этом направлении.

Допущение о достаточном количестве волн рекрутирования, необходимом для достижения равновесия. Методика RDS предполагает, что получаемая сетевая выборка должна достигать точки равновесия, т.е. момента, после которого характеристики рекрутируемых респондентов перестают зависеть от признаков семян. Тем не менее реализация RDS не всегда позволяет набрать требуемое количество волн и получить ожидаемый результат [9, р.6]. При этом оказывается достаточно сложно установить, что стало причиной недобора респондентов и прекращения роста сети: низкая мотивация участников, недостаточный размер сети или же достигнутый предел достижимости определенной группы в популяции. Любой из этих факторов может привести к затуханию роста одной из цепочек выборки или даже всей выборки в целом.

Случайность отбора в процессе рекрутирования респондентов. Идеальная модель выборки, направляемой респондентом, требует того, чтобы рекрутеры отбирали респондентов случайным образом. Однако исследования показывают, что это представление достаточно сильно отличается от того, что происходит в реальности: выбор участников является не случайным и представляет собой слож-

ный процесс. Фактически, рекрутирование состоит не из одного, а из трех последовательных решений, которые влияют на то, присоединится ли конкретный человек к выборке [9, р. 18–19]:

1. решение рекрутера пригласить конкретного друга или знакомого;
2. решение данного друга или знакомого принять приглашение;
3. решение друга или знакомого использовать это приглашение и поучаствовать в исследовании.

Каждое из этих решений влияет на вероятность попадания в выборку конкретного респондента и значительно усложняет задачу расчета этой вероятности.

Достаточно слабое влияние эффекта гомофилии. Ситуация затрудняется и тем, что отбор новых участников происходит под влиянием ряда субъективных и ситуативных обстоятельств. Чаще всего рекрутеры обращаются к подходящим знакомым, которые находятся в близком социальном окружении. Например, в исследовании студентов, обучающихся в НИУ ВШЭ, была зафиксирована тенденция к приглашению в исследование учащихся своего пола, также проживающих в общежитии [8, с. 47]. В исследовании Хекаторна и Вайнерта было отмечено смещение распределения респондентов по расовому признаку — представители группы «белые» значительно реже приглашали азиатов, и наоборот [2, р. 128]. Подобные примеры указывают на обстоятельства, которые приводят к несоблюдению допущения о слабом влиянии эффекта гомофилии.

Точные оценки числа социальных связей всеми членами группы. Допущение о том, что все респонденты могут достаточно точно оценить размер своей эго-сети, является одним из необходимых условий для получения несмещенных оценок на основе RDS. Однако различные эмпирические исследования ставят его реалистичность под сомнение. Сложность заключается в том, что точность получаемых оценок зависит от когнитивных возможностей респондента, особенностей его памяти и восприятия понятийных конструктов, которые используются для измерения эго-сети. Например, согласно выводам Киприана Вайнерта, респонденты приписывают различные значения понятиям «друг» и «близкий друг» [11, р. 90]. Возможным выходом из подобного затруднения является постановка более

конкретного вопроса о характеристиках эго-сети⁴. В дополнение к этому в практике исследований на основе RDS принято задавать вопросы о числе контактов в телефонной книге, числе друзей в социальной сети — т.е. показателях, которые респондент может проверить и указать максимально точно.

Возможность повторного отбора респондентов. Возможность повторного отбора является формальным требованием для многих моделей оценки статистических параметров [9, р.7]. Несмотря на это, в большинстве известных исследований с применением RDS реализуется бесповторный отбор. Такое нарушение условий модели не имеет критического значения, если изучаемая популяция имеет большой размер. В обратном случае ограничение на повторный отбор может привести к появлению смещений и практических затруднений [9, р.8]. В рамках специального исследования Кристи Джиллиан и ее коллег не менее 25% респондентов в трех примерах использования RDS пытались пригласить знакомого, который уже принимал участие в опросе [9, р.6]. Помимо смещений, подобные ограничения возможностей рекрутирования могут привести к замедлению роста сети и к тому, что в исследовании примет участие недостаточное количество респондентов.

Приведенные замечания и результаты эмпирических исследований относятся к наиболее важным аспектам критики методологии RDS. В совокупности они указывают на то, что выборка, направляемая респондентом, не может быть реализована в соответствии со всеми требованиями формальной модели. Строго говоря, такое заключение не является основанием для прекращения использования этого метода: любая модель формирует определенный идеальный образ, который никогда не реализуется на практике. Тем не менее необходимо принимать во внимание то, что методология RDS нуждается в развитии и усовершенствовании как со стороны методики, так и со стороны существующих моделей получения статистических оценок. В следующем разделе данной статьи рассматривается опыт

⁴ Например, просьба дать оценку числу знакомых, с которыми респондент взаимодействовал (при личной встрече, по телефону, в переписке) в определенный период времени.

применения WebRDS, который указывает на то, что некоторые из имеющихся методических проблем уже нашли свое решение.

Анализ примеров применения WebRDS: решения и рекомендации

Несмотря на то, что выборка, направляемая респондентом, получила определенное распространение в исследованиях труднодоступных и закрытых социальных групп, опыт применения WebRDS ограничивается небольшим числом известных примеров. С целью его обобщения при подготовке данной статьи были проанализированы семь кейсов из зарубежной и российской исследовательской практики⁵ (Таблица 2). Сопоставление рассматриваемых исследований, а также анализ общих практических замечаний показывают, что реализация WebRDS зависит от многих обстоятельств, которые должен принимать во внимание исследователь.

Отбор первичных респондентов. Несмотря на то, что ожидаемым результатом применения WebRDS является независимость выборки от семян, отбор первичных респондентов предполагает соблюдение некоторых принципов. Так, с целью обеспечения роста сети исследователи стремятся найти «звезд» изучаемой группы, т.е. социально активных и наиболее общительных респондентов (Кейс 3 и Кейс 7). Кроме того, исследователи должны отобрать участников, которые отличаются друг от друга по значимым для анализа характеристикам, для того чтобы минимизировать вероятность смещения выборки в сторону определенной подгруппы. Например, в рамках кейса 4 исследователи включили в состав семян 22-х молодых взрослых, проживающих в разных частях США и имеющих различную расовую принадлежность [6, p. 835].

Свобода респондентов/рекрутеров в приглашении участников. Общая логика RDS предполагает, что после отбора семян ис-

⁵ Отбор кейсов был основан на задаче представить результаты первых экспериментов по применению WebRDS, а также опыт применения WebRDS для исследования обычных и труднодоступных групп, вариативный с точки зрения отдельных элементов методики

Таблица 2. Сводное описание методики и результатов рассматриваемых исследований на бае WebRDS

№	Авторы	Год	Длительность сбора данных	Целевая аудитория	Число семян	Размер выборки	Возна-граждение за опрос	Возна-граждение за рекрут	Макси-мальное число при-глашений
1	Киприан Вейнерт, Дуглас Хекаторн	2004	72 часа	студенты университета	9	159	10\$ (288,1 руб.)	15\$ (432,15 руб.)	3
2	Киприан Вейнерт	2008	2 месяца	студенты университета	9	378	до 25\$ (621 руб.) за опрос и рекрут		3
3	Айгуль Мавлетова	2009	4 месяца	гемблеры	4	99	нет	нет	6
4	Жозе Бауэрмейстер и др.	2011	2,5 месяца	молодые взрослые (18–24)	22	3448	20\$ (588 руб.)	10\$ (294 руб.)	5
5	Линус Бенгтссон и др.	2011	3 месяца	мужчины, которые когда-либо вступали в сексуальную связь с мужчиной, 18+, про-живающие во Вьетнаме	20	676	2,45 (72 руб.)	2,45\$ и др. (72 руб.)*	4
6	Станислав Моисеев, Юлия Савинкова	2013	2 месяца	студенты университета	6	72	150 руб. за опрос и рекрут		3
7	Станислав Моисеев, Юлия Савинкова	2013	1 месяц	студенты университета	5	79	150 руб. за опрос и рекрут		3

* В общей сложности система мотивации в данном исследовании состояла из пяти элементов: 2,45\$ в качестве аванса на баланс счета мобильного телефона и аналогичное вознаграждение за каждого из 4-х привлеченных респондентов; возможность пожертвовать средства в пользу объединения мужчин, которые вступали в сексуальную связь с мужчиной по выбору респондента; потеря с возможностью выиграть iPad; текст, подчеркивающий значимость участия для поддержки мужчин, которые вступали в сексуальную связь с мужчиной во Вьетнаме; возможность сравнить собственные ответы с анонимными результатами опроса других участников. Для того, чтобы повысить интерес к сравнению в исследовательский инструментарий были включены 8 дополнительных вопросов [5, p. 3].

следователь должен довериться социальному опыту представителей изучаемой группы — вероятнее всего, они найдут лучший способ обратиться к своему другу или знакомому. Во-первых, это касается выбора канала коммуникации. Так, решение не ограничивать способ онлайн-коммуникации в исследовании молодых взрослых позволило повысить число рассылаемых приглашений [6, р.835]. Во-вторых, рекрутеры должны иметь возможность обращаться к потенциальным респондентам в удобной для них форме. Наличие строгого шаблона приглашения несколько упрощает их задачу, однако полностью нивелирует эффект персонального приглашения. Для того чтобы облегчить работу рекрутера, его можно снабдить необходимыми информационными материалами, которые он будет рассылать вместе с предложением принять участие в опросе.

Мониторинг процесса приглашения. Одновременно с этим реализация WebRDS требует контроля за процессом рекрутирования со стороны исследователя. Благодаря такому мониторингу он получает возможность оптимизировать методику, что может оказаться очень важным на начальном этапе. Например, для того, чтобы повысить скорость роста выборки в кейсах 4 и 5 исследователи меняли дизайн WebRDS после начала сбора данных: проводили добор семян [5, р. 3] и увеличивали максимальное число приглашений для каждого участника [6, р. 835].

Мотивация респондентов/рекрутеров. Как показывает практика, методика WebRDS работает и в отсутствии денежного стимула для участников исследования (кейс 3). Тем не менее его наличие, а главное, объем позволяют повысить готовность респондентов к кооперации. Оптимальный размер вознаграждения должен определяться в зависимости от типа целевой аудитории. Опыт изучения студентов университета Корнелла показал, что уменьшение максимального вознаграждения с 55 до 25 долларов не оказало негативного влияния на размер WebRDS [11, р. 81]. Материальные поощрения могут дополняться системой символической мотивации: возможностью сравнения ответов респондента с ответами других участников, перевода средств на благотворительность и т.п. [5, р. 3]. Кроме того, эффективность WebRDS зависит от лояльности участников. Результаты исследования, проведенного авторами данной

статьи, дают основания для предположения о том, что рост размера WebRDS происходит быстрее, если первичные респонденты являются знакомыми или друзьями исследователя [8, с. 48].

Динамика роста сети. В рассматриваемых кейсах длительность исследования варьировалась от нескольких дней (кейс 1) до четырех месяцев (кейс 3). При наличии идентичных или схожих условий подобный разброс объясняется влиянием ситуативных условий прохождения опроса и приглашения новых участников. В частности, второе исследование студентов Корнельского университета совпало с периодом между весенними каникулами и экзаменационной сессией, что оказало негативное влияние на вовлеченность студентов [11, р. 81]. В общем случае динамика WebRDS зависит от мотивации участников и установления определенных временных ограничений: рекомендуемого/обязательного срока заполнения онлайн-анкеты и приглашения новых участников.

Использование данных, собранных до достижения точки равновесия. При успешном завершении процедуры сбора данных исследователь сталкивается с необходимостью решения важного вопроса: как поступить с данными, собранными до прохождения точки равновесия [11, р. 90]? Первый и наиболее радикальный путь заключается в отказе от их использования. Фактически это приводит к потере существенной части собранной информации, однако избавляет оценки параметров от потенциальных смещений. Его альтернатива — использование всех собранных данных. Такое решение кажется наиболее простым, однако, если исследователь не имеет обоснованных представлений о распределении ключевых параметров генеральной совокупности, он не может обоснованно гарантировать того, что дополнительные данные не исказят общую картину. Для того чтобы найти однозначный ответ на поставленный вопрос, необходимо проводить больше практических исследований, одной из задач которых является сравнение результатов анализа, полученных по полной выборке, с оценками параметров, полученными по данным от части респондентов, рекрутированных после достижения точки равновесия.

Нарушение анонимности. Одним из главных преимуществ RDS является возможность анонимного сбора данных, который исклю-

чает контакт респондента и исследователя. Тем не менее проведение WebRDS требует того, чтобы информанты оставляли сведения, по которым их можно идентифицировать — адрес электронной почты, телефон и т.п. (это условие описывается в кейсах 4, 5, 6 и 7). Такая особенность не является существенным препятствием для сбора данных, в том числе в труднодоступных группах. Одним из способов смягчения требования предоставления персональных сведений является соответствующая открытость со стороны исследователя — сообщение персональных данных организаторов, которые могут использоваться для обратной связи и контакта по любым возникающим вопросам [8].

Применение специальных технологических инструментов. Анализ примеров реализации WebRDS в зарубежных исследованиях показывает, что для подобных проектов нередко разрабатываются специальные технологические инструменты. Так, в кейсе 1 описывается система, которая собирала опросные данные, делала их резервные копии и блокировала возможность повторного прохождения опроса одним и тем же респондентом [2, р. 114]. При сборе данных о молодых взрослых (кейс 4) исследователи использовали специальный интернет-портал, на котором респондент создавал собственный аккаунт и затем проходил опрос [6, р.835]. Исследовательская группа кейса 5 разработала веб-сайт, который позволял собирать данные, а также делать их наглядное представление, с которым мог ознакомиться каждый респондент [5, р. 3]. Создание подобных опросных систем безусловно упрощает процедуру сбора данных и может повышать их эффективность. В то же время такие решения могут затруднить последующее воспроизводство реализуемой методики и приводят к ее удорожанию.

Приведенные решения и рекомендации показывают, что проведение исследования на базе технологии WebRDS требует принятия большого количества методических и организационных решений. При этом тщательное планирование полевого этапа не исключает того, что в разработанную схему придется внести ряд определенных изменений. Исследователю необходимо помнить о том, что важным источником обратной связи об адекватности и эффективности процедуры опроса являются первичные респонденты (семена), которые

специально отбираются исследователем и чаще всего настроены на сотрудничество.

К вопросу о перспективе

Основная задача данной статьи заключалась в том, чтобы зафиксировать актуальное состояние научной дискуссии вокруг возможностей применения выборки, направляемой респондентом. На первый взгляд, обобщенные результаты методологических и эмпирических исследований по применению RDS и WebRDS представляют достаточно пессимистичную картину. В статье показано, что с точки зрения ряда базовых допущений, формальная модель выборки, направляемой респондентом, не соответствует практике ее реализации; использование (Web) RDS требует от исследователя решения множества методических и организационных вопросов; сочетание этих условий снижает вероятность получения качественного и надежного результата при применении (Web) RDS. Очевидным следствием подобных выводов является необходимость усовершенствования методологии RDS.

Однако они не означают того, что применение RDS в традиционных и онлайн-исследованиях не имеет дальнейших перспектив. Этот тезис подтверждается несколькими эмпирическими кейсами, которые показывают, что перенос технологии RDS в область онлайн-исследований позволяет формировать выборки достаточно большого размера и получать несмещенные оценки параметров исследуемой совокупности.

Хотелось бы отметить, что наличие методических и организационных проблем требует увеличения числа методологических экспериментов по применению RDS и WebRDS. Их содержание должно быть направлено не только на решение новых проблем: можно предположить, что в ближайшее время особое значение будут иметь работы, направленные на проверку результатов, полученных в более ранних исследованиях. Еще одним важным аспектом развития методологии WebRDS является усовершенствование и разработка моделей оценки параметров, используемых при работе с данными сетевой выборки.

ЛИТЕРАТУРА

1. Brick M.J. The Future of Survey Sampling // *Public Opinion Quarterly*. Vol.75. №3. 2011. P. 872-888.
2. Heckathorn D. Web-Based Network Sampling Efficiency and Efficacy of Respondent-Driven Sampling for Online Research / D. Heckathorn, C.Wejnert // *Sociological Methods & Research*. 2008. №. 37. P. 105-134.
3. Яковлева А.А. Исследования в труднодоступных группах: опыт использования выборки, управляемой респондентом, и выборки «место-время» // *Социология* 4М. 2011. №33. С. 57-79.
4. Печенкин В.В., Зайонц В.В. Сетевые методы изучения виртуальных сообществ // *Теория и практика общественного развития*. 2011. №5. С. 71-75.
5. Bengtsson L. Implementation of Web-Based Respondent-Driven Sampling among Men Who Have Sex with Men in Vietnam / L. Bengtsson, X. Lu, Q.C. Nguyen, M. Camitz, N.L. Hoang, T.A. Nguyen, F. Liljeros, A. Thorson // *PLoS One*. 2012. Vol.7. №11. P. 1-9.
6. Bauermeister J.A. Innovative Recruitment Using Online Networks: Lessons Learned From an Online Study of Alcohol and Other Drug Use Utilizing a Web-Based, Respondent-Driven Sampling (web RDS) Strategy / J.A. Bauermeister, M.A. Zimmerman, M.M. Johns, P. Glowacki, S. Stoddard, A. Volz // *Journal Of Studies On Alcohol And Drugs*. 2012. Vol.73. P. 843-848.
7. Мавлетова А.М. Проведение онлайн RDS эксперимента: изучение мотивации гемблеров // *Вестник Российского университета дружбы народов*. Серия: Социология. 2010. № 4. С. 36-49.
8. Моисеев С.П., Савинкова Ю.К. Выборка, направляемая респондентом, в онлайн-опросе: к вопросу о динамике и качестве // *Мониторинг общественного мнения*. 2014. № 6. С. 42-50.
9. Gile K.J. Diagnostics for respondent-driven sampling / K.J. Gile, L.G. Johnston, M.J. Salganik // *Journal of the Royal Statistical Society*. 2015. Series A (Statistics in Society). P. 1-17.
10. Heckathorn D. Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations / D. Heckathorn // *Social Problems*. 1997. Vol. 44. №2. P. 174-199.
11. Wejnert C. An Empirical Test of Respondent-Driven Sampling: Point Estimates, Variance, Degree Measures, and Out-of-Equilibrium // *Sociological Methodology*. 2009. Vol. 39. P. 73-166.

Раздел 4

Методология онлайн-исследований: онлайн-панели

Эффект самоотбора при формировании онлайн-панели: опыт первого года лонгитюдного исследования студентов НИУ ВШЭ

*Малошенок Наталья Геннадьевна,
Семенова Татьяна Вадимовна,
Терентьев Евгений Андреевич
(Национальный исследовательский университет
«Высшая школа экономики»)*

С широким распространением сети интернет все больше исследователей обращаются к формату онлайн-опросов, который позволяет снизить временные и денежные затраты на проведение полевого этапа, а также сделать процесс заполнения анкеты более удобным и интересным для респондента [1; 2]. Однако такой формат сбора данных не лишен своих недостатков, которые прежде всего связаны с трудностями обеспечения репрезентативности выборочной совокупности, что накладывает ограничения на качество получаемых данных. Как правило, выделяют два феномена, которые приводят к снижению качества данных в онлайн-исследованиях: неполнота охвата (under-coverage) и эффект самоотбора (self-selection) [2]. Неполнота охвата возникает в том случае, когда генеральная совокупность, из которой формируется выборка, не вполне соответствует целевой совокупности (примером является отсутствие у исследователя полного списка всех электронных адресов потенциальных респондентов, на которые рассылается приглашение к участию в опросе) [1]. Эффект самоотбора возникает в том случае, когда исследователь не может контролировать процесс отбора респондентов в выборочную совокупность и определенные характеристики респондентов оказывают значимое влияние на принятие ими решения относительно участия в опросе [2; 3]. Особенно это характерно для онлайн-опросов, проводимых на интернет-сай-

тах, участниками которых становятся только те респонденты, которые, во-первых, имеют доступ к сети интернет, во-вторых, зашли на страницу сайта, в-третьих, увидели приглашение к участию в исследовании и, в-четвертых, решили принять участие в опросе. При наличии эффекта самоотбора характеристики респондентов, принявших участие в опросе, значительно отличаются от характеристик респондентов, отказавшихся от заполнения онлайн-анкеты [3]. Поэтому ответы респондентов, принявших участие в исследовании, могут значительно отличаться от ответов не-респондентов, что влияет на валидность данных и приводит к нарушениям принципов вероятностной выборки [2].

Среди исследователей распространение получили четыре основных метода оценки эффекта самоотбора [4; 5]: 1) сравнение ответов ранних (заполнивших анкету до рассылки напоминаний о возможности участия в опросе) и поздних (заполнивших анкету после рассылки напоминаний о возможности участия в опросе) респондентов. В основе применения такого метода лежит гипотеза о том, что поздние респонденты близки по своим характеристикам к тем, кто не ответил (модель континуума сопротивления) [6]; 2) отслеживающее исследование не-респондентов, когда для тех, кто не принял участие в исследовании, проводится дополнительный опрос (обычно с использованием стимулирующих мероприятий для обеспечения отклика) и сравниваются ответы участников основного и отслеживающего опросов; 3) панельное исследование, предполагающее проведение нескольких опросов на одной и той же когорте респондентов. В результате основой для сравнения респондентов и не-респондентов становятся данные, полученные в ходе первого опроса, охватившего всех участников панели; 4) использование дополнительных данных (обычно социально-демографических) о генеральной совокупности и сравнение на их основе характеристик респондентов и не-респондентов. Каждый из методов обладает своими преимуществами и недостатками, поэтому рекомендуется использовать сложные дизайны исследований, позволяющие сочетать различные методы оценки эффекта самоотбора [4].

Результаты значительного количества исследований, проведенных на разных аудиториях, установили наличие влияния

на готовность участвовать в опросе таких характеристик, как заинтересованность в проблематике/предмете исследования [3; 7; 8; 9; 10], а также уровень дохода [3; 7]. Так, чем выше уровень дохода у респондента, тем выше вероятность того, что он примет участие в опросе. Кроме того, чем выше уровень заинтересованности респондента в проблематике исследования, тем выше вероятность того, что он согласится участвовать в обследовании. При этом такие социально-демографические характеристики, как пол и уровень образования респондента, оказывали неоднозначное влияние на готовность участвовать в опросе. Так, в большом количестве исследований было показано, что женщины в среднем более склонны участвовать в опросах [11; 12; 13; 14]. Однако, например, в исследовании, проведенном Дж. Уайтхэд и коллегами, было показано, что более склонны принимать участие в исследовании мужчины [7]. Такие различия во многом объясняются спецификой темы исследования. Аналогичные различия были обнаружены и в отношении уровня образования. При этом большинство исследований показывают, что люди с более высоким уровнем образования более склонны участвовать в опросах [9; 12; 15; 16; 17].

Эффект самоотбора в исследованиях студенческих аудиторий находился в фокусе ряда эмпирических работ [4; 18; 19]. Так, Э. Дей на основании анализа данных национального лонгитюдного исследования студентов колледжей, реализуемого Министерством образования США и Кооперативной программой институциональных исследований (CIRP)¹, показал, что на готовность принять участие в исследовании позитивно влияют такие факторы, как высокий уровень успеваемости в школе, высокая самооценка академических способностей, евроамериканская этническая принадлежность, сравнительно большее количество лет, в течение которых студент изучает иностранный язык в высшей школе, ориентация на научную деятельность, совместное проживание родителей в браке, женский пол, наличие ожидания получить степень бакалавра, высокая

¹ Данные были получены с помощью сравнения результатов основного и отслеживающего (follow-up) исследований. Исследование было проведено в традиционном бумажном формате.

самооценка математических способностей и сравнительно большее количество лет изучения математики в высшей школе. Сильная негативная корреляция с уровнем отклика наблюдалась в отношении таких факторов, как принадлежность афроамериканскому сообществу, ориентация на предпринимательскую деятельность и экономическое благополучие, негативные установки в отношении сексуальных меньшинств и социальная активность, выраженная в активном посещении вечеринок [18].

В работе А. Сакс и коллег на основе данных двух национальных американских опросов студентов-первокурсников² (CIRPF freshmen Survey и YFCY), было обнаружено, что среди студентов женского пола и американо-азиатской или евроамериканской этнической принадлежности отклик был значимо выше [19]. При этом для группы студентов, получивших приглашение заполнить онлайн-анкету по e-mail со стимулированием³, значимыми факторами, повышающими вероятность участия в исследовании, также являлись высокие результаты стандартизованного вступительного экзамена (SAT) и принадлежность к психологическому типу, ориентированному на достижение социального статуса. В группе получивших приглашение заполнить онлайн-анкету по e-mail без стимулирования позитивное влияние на отклик оказали такие факторы, как обучение по гуманитарным специальностям или специальностям, связанным с английским языком и изобразительным искусством, а также принадлежность к психологическим типам лидера, социального активиста или человека искусства.

С. Портер и М. Уиткомб исследовали влияние социально-демографических и психологических характеристик студентов, а также показателей их социальной вовлеченности на уровень отклика на основе данных четырех исследований, проведенных в одном из американских колледжей, специализирующихся в области свободных искусств [4]. Стоит отметить, что в этой работе использовался более сложный

² Исследование проведено в бумажном и онлайн-форматах.

³ В качестве стимулирования выступала информация от исследователей о том, что после завершения исследования будет направлен всем участникам опроса отчет с результатами обследования.

дизайн исследования и более широкий круг данных, чем в двух рассмотренных выше работах. Так, оценка эффекта самоотбора производилась не на основании отслеживающего исследования, а на основании сравнения характеристик ответивших с характеристиками генеральной совокупности, полученными из административной базы данных вуза. Кроме того, в качестве независимых переменных также рассматривались некоторые показатели, полученные в ходе проведения национального американского опроса первокурсников (CIRPF freshmen Survey), в рамках которого были охвачены более 90% респондентов — участников опросов в вузе. Важным преимуществом этой работы также является то, что благодаря включению в анализ данных четырех исследований, посвященных различным темам (употреблению алкоголя и наркотиков, еде, студенческой вовлеченности и удовлетворенности обучением в вузе), исследователи могли контролировать фактор темы исследования, который сам по себе может оказывать значимое влияние на желание студентов принять участие в исследовании. Результаты анализа во многом подтвердили результаты предыдущих исследований и показали, что более склонны принимать участие в исследованиях женщины по сравнению с мужчинами, более успевающие студенты по сравнению с менее успевающими, и люди из семей с благополучным материальным положением по сравнению с теми, кто получает материальную помощь. Кроме того, на желание принять участие в исследовании оказывают влияние также некоторые психологические характеристики студентов и показатели их социальной вовлеченности. Так, сравнительно большая доля тех, кто принял участие в опросе, была зафиксирована среди студентов с исследовательским типом личности, а сравнительно меньшая — среди студентов с предпринимательским типом личности. Также было обнаружено, что большая вероятность участия в опросах отмечается среди социально активных студентов.

Считается, что сильнее эффект самоотбора влияет на качество данных, получаемых в тех исследованиях, которые предполагают долгосрочное участие в исследовании, например, в лонгитюдных обследованиях. Однако эмпирической оценки эффект самоотбора в лонгитюдных онлайн-исследованиях не получил. Поэтому в данной работе была поставлена *цель* оценить эффект самоотбора в лонги-

тюдном исследовании с помощью определения социально-демографических характеристик и установок, по которым респонденты, согласившиеся на долгосрочное участие, отличаются от отказавшихся, и выявления влияния характеристик респондентов на готовность принять участие в обследовании.

Процедура набора респондентов в панель и их удержания в течение первого года лонгитюдного исследования

Данная работа основывается на лонгитюдном исследовании «Траектории и опыт студентов университетов России»⁴, которое проводится на студентах, поступивших в вуз в 2014 году, и координируется Институтом образования НИУ ВШЭ. В этом исследовании участвуют 10 университетов из числа участников программы «5–100», а также федеральных университетов. Однако в статье мы рассмотрим опыт проведения лонгитюдного исследования в Национальном исследовательском университете «Высшая школа экономики», поскольку к данному моменту в этом университете проведены две волны исследования.

Дизайн исследования предполагает проведение шести волн онлайн-опроса: первая волна — в начале первого года обучения в университете, затем опросы будут проводиться весной первого, второго, третьего и четвертого курса, а также через год после окончания университета.

В НИУ ВШЭ набор студентов в исследовательскую панель был осуществлен среди первокурсников 6 направлений подготовки: политология, прикладная математика и информатика, программная инженерия, психология, социология и экономика (совместный бакалавриат ВШЭ и РЭШ, экономика, экономика и статистика).

Процедура набора респондентов осуществлялась в два этапа. 1 сентября 2014 г. все студенты первых курсов соответствующих специальностей получили электронное письмо с приглашением принять участие в опросе первокурсников. В тексте приглашения

⁴ http://ioe.hse.ru/collaborative_project/

было указано, что целью опроса является знакомство руководства вуза с новыми студентами университета и получение информации об их предпочтениях и опыте обучения в школе. После заполнения 5-минутной анкеты студентам показывалось сообщение, в котором респондента благодарили за данные им ответы, а также информировали, что его факультет (департамент) стал участником лонгитюдного исследования и приглашали стать участником данного исследования, предоставляя информацию о его задачах и условиях участия в нем. За участие в исследовании респондентам было обещано гарантированное ежегодное вознаграждение, размер которого варьировался от 300 до 1000 рублей в зависимости от года участия (максимальная сумма полагалась за участие в последней волне исследования). Для того чтобы стать участником исследования, студент должен был заполнить специальную форму, указав в соответствующих полях свои ФИО, адрес электронной почты, номер мобильного телефона и аккаунт в социальной сети «ВКонтакте». После этого ему предлагалось заполнить первую анкету лонгитюдного исследования, что в среднем занимало около 20 минут. Если студент отмечал, что в данный момент ему неудобно заполнять анкету, он мог выбрать возможность заполнить анкету через пару дней. В таком случае через два дня на указанный адрес электронной почты дополнительно высылалась ссылка на анкету лонгитюдного исследования. Участником панели становился тот, кто полностью заполнял эту анкету.

Помимо первого письма-приглашения студентам, не заполнившим короткую анкету первокурсника или заполнившим ее частично, высылались дополнительно 3 напоминания о возможности принять участие в опросе. Напоминания были разосланы 3, 5 и 8 сентября 2014. Текст напоминаний незначительно отличался от текста первых приглашений. Тем, кто полностью заполнил короткую анкету первокурсника и отметил, что хочет стать участником лонгитюдного исследования, но не заполнил первую анкету лонгитюдного исследования полностью, также дополнительно высылались 3 напоминания о возможности стать участником лонгитюдного исследования. В результате на приглашение к участию в опросе первокурсников откликнулись 634 человека, что составляет 69% от генеральной совокупности. Из них полностью заполнили короткую анкету перво-

курсника 590 человек (65% от генеральной совокупности). 382 человека (42% от генеральной совокупности и 65% от тех, кто полностью заполнил анкету первокурсника) согласились принять участие в лонгитюдном исследовании. Из них полностью заполнили анкету лонгитюда и стали участниками панели 283 человека, что составляет 31% от генеральной совокупности и 48% от тех, кто заполнил полностью анкету первокурсника и мог стать участником лонгитюда.

Вторая волна лонгитюдного исследования в НИУ ВШЭ была запущена 8 апреля 2015 года. К участию в ней приглашались только те первокурсники, которые полностью заполнили первую анкету лонгитюдного исследования. Опрос проходил в два этапа. На первом этапе ссылка на вторую анкету лонгитюдного исследования высылалась студентам на электронную почту, указанную во время первой волны. Затем студентам, которые не ответили на опрос, высылались напоминания (10 и 15 апреля 2015). С 21 апреля (второй этап опроса) был организован обзвон участников лонгитюдного исследования, не ответивших на вторую анкету, который позволил получить заполненные анкеты дополнительно от 89 респондентов. Обзвон производился интервьюерами, которые напоминали первокурсникам об их участии в лонгитюдном исследовании и повторно направляли ссылку на опрос. Собственно заполнение анкеты респондент производил в онлайн-формате. Общее количество опрошенных во вторую волну составило 257 студентов НИУ ВШЭ, еще 6 первокурсников перешли по ссылке на анкету, но не закончили ее заполнение. Таким образом, отклик во второй волне составил 90%, а при учете отчисленных студентов — 97%⁵. В НИУ ВШЭ, в отличие от других вузов, респондентам выплачивалось вознаграждение в размере 300 рублей за первый год участия в исследовании.

Учитывая дизайн исследования, эффект самоотбора оценивался с использованием следующих типов данных:

1. Административные данные, собираемые о студентах при подаче документов и при поступлении, включающие информа-

⁵ 19 студентов из 283 участников первой волны лонгитюдного исследования было отчислено к моменту начала второй волны. Они не заполняли вторую анкету исследования.

цию о возрасте, месте постоянного проживания, баллах ЕГЭ, основании поступления, других направлениях подготовки, на которые студент подавал документы в НИУ ВШЭ, и приоритетном для поступления направлении подготовки. Данные собираются и хранятся в статистической базе «Абитуриент-Студент-Аспирант-Выпускник» (АСАВ).

2. Данные бумажного опроса абитуриентов бакалавриата, который проводится Центром внутреннего мониторинга НИУ ВШЭ при подаче документов для поступления. Опрос проводился с 20 июня по 25 июля 2014 года. Отклик составил 83% от общего числа абитуриентов НИУ ВШЭ 2014 года. В анкету были включены вопросы о подготовке к поступлению в вуз, причинах выбора НИУ ВШЭ, вузах, в которые абитуриент подавал документы на поступление помимо НИУ ВШЭ, участии во всероссийских олимпиадах школьников, наличии спортивных и творческих достижений.
3. Данные опроса первокурсников. В анкету были включены вопросы о любимых предметах в школе, причинах выбора направления подготовки, пользовании социальными сетями, культурном капитале, материальном положении семьи, целях обучения и т.д.

Результаты анализа эффекта самоотбора

На основе имеющихся данных мы проанализировали различия между тремя группами студентов:

1. респондентами, согласившимися на участие в лонгитюдном исследовании и заполнившими первую анкету лонгитюдного исследования до конца;
2. респондентами, принявшими участие в опросе первокурсников, но отказавшимися от участия в лонгитюдном исследовании;
3. студентами, не заполнившими ни анкету первокурсника, ни анкету лонгитюдного исследования.

Сравнения трех групп первокурсников производилось по следующим параметрам:

1. Характеристики обучения:
 - форма финансирования обучения (данные АСАВ);
 - направление подготовки (данные АСАВ);
 - основание поступления (данные АСАВ).
2. Социально-демографические характеристики студента:
 - пол (данные АСАВ);
 - регион проживания до поступления в НИУ ВШЭ (данные АСАВ);
 - тип населенного пункта (данные АСАВ);
 - материальное положение семьи (данные бумажного опроса абитуриентов).
3. Уровень исходной подготовки и академическая успеваемость:
 - балл ЕГЭ по предметам (данные АСАВ);
 - средний балл в первом и втором семестрах в НИУ ВШЭ (данные АСАВ).
4. Лояльность к НИУ ВШЭ:
 - подача документов на поступление только в НИУ ВШЭ (данные бумажного опроса абитуриентов);
 - осознанность выбора направления подготовки (данные бумажного опроса абитуриентов);
 - вовлеченность в поиск информации о НИУ ВШЭ на сайте (данные бумажного опроса абитуриентов);
 - время принятия решения о поступлении в НИУ ВШЭ (данные бумажного опроса абитуриентов).
5. Цели поступления в НИУ ВШЭ (данные бумажного опроса абитуриентов).

Характеристики обучения

В данном разделе будут рассмотрены отличия участников лонгитюдного обследования от других первокурсников, входящих в генеральную совокупность нашего исследования, по формальным характеристикам их обучения в НИУ ВШЭ: форме финансирования обучения, направлению подготовки, основанию поступления.

Результаты анализа показывают, что эффект самоотбора не наблюдается в отношении такой характеристики, как форма финансирования

ния обучения: доли студентов, обучающихся на бюджетном и платном местах, не отличаются статистически значимо в зависимости от статуса участия студента в лонгитюдном исследовании, что позволяет говорить о репрезентативности полученной нами выборки генеральной совокупности, из которой производился набор респондентов⁶ [Табл. 1].

Таблица 1. Распределение первокурсников по форме финансирования обучения в зависимости от статуса участия в исследовании

	Платное место (N=154)	Бюджетное место (N=758)
Не заполнял ни одну анкету (N=323)	14,9%	85,1%
Заполнил только анкету первокурсника (N=305)	19,0%	81,0%
Участник лонгитюдного исследования (N=284)	16,9%	83,1%

Были выявлены значимые различия в желании участвовать в лонгитюдном исследовании в зависимости от направления подготовки первокурсника, на которое он поступил⁷. Большой интерес проявили студенты отделений социологии и прикладной математики и информатики, меньший — студенты экономических специальностей и программной инженерии [Табл. 2]. Скорее всего, выбор определенного направления подготовки для обучения имеет взаимосвязь с отношением студента к социологическим опросам, что отражается на различии в отклике для учащихся с разных специальностей. Таким образом, данные подобных обследований следует анализировать в отдельности по специальностям, или группам специальностей, принадлежащих к смежным профессиональным областям и имеющим примерно одинаковый уровень отклика.

Также участники нашего лонгитюдного исследования отличаются от тех, кто заполнял только короткую анкету, и тех, кто

⁶ Различия не значимы. Хи-квадрат = 1.930, df = 2, p=0.381

⁷ Различия значимы. Хи-квадрат = 43.774, df=14, p<0.001

Таблица 2. Распределение первокурсников по направлению подготовки в зависимости от статуса их участия в исследовании

	Не заполнял ни одну анкету (N=323)	Заполнил только анкету первокурсника (N=307)	Участник лонгитюдного исследования (N=284)
Политология (N=76)	32,9%	31,6%	35,5%
Прикладная математика и информатика (N=194)	25,7%	38,7%	35,6%
Программная инженерия (N=113)	38,0%	38,1%	23,9%
Психология (N=79)	43,1%	21,5%	35,4%
Совместная программа по экономике НИУ ВШЭ и РЭШ (N=65)	43,1%	33,8%	23,1%
Социология (N=118)	23,7%	29,7%	46,6%
Экономика (N=202)	43,6%	34,7%	21,7%
Экономика и статистика (N=65)	41,5%	29,2%	29,3%

не участвовал ни в одном опросе, в зависимости от основания зачисления в НИУ ВШЭ⁸. Как видно из таблицы 3, среди студентов, поступивших в НИУ ВШЭ по результатам вступительных испытаний (ЕГЭ), сравнительно больше доля тех, кто не заполнял ни одну анкету, в то время как среди первокурсников, поступивших как победители олимпиад, сравнительно больше тех, кто заполнил анкету первокурсника [Табл. 3]. При этом соотношение студентов с разными основаниями поступления среди участников лонгитюдного исследования значимо не отличается от соотношения в генеральной совокупности.

⁸ Различия значимы. Хи-квадрат = 17.834, df=8, p=0.023

Таблица 3. Распределение первокурсников по основанию поступления в вуз в зависимости от статуса участия в исследовании

	Не заполнял ни одну анкету (N=323)	Заполнил только анкету первокурсника (N=307)	Участник лонгитюдного исследования (N=284)
Победитель олимпиады школьников (N=391)	34,0%	36,1%	29,9%
По результатам вступительных испытаний (N=292)	42,8%	25,0%	32,2%
Вне конкурса по федеральным льготам (N=21)	23,8%	42,9%	33,3%
По целевому набору (N=22)	54,5%	36,4%	9,1%
На основании договора (N=133)	36,1%	35,3%	28,6%

Таким образом, эффект самоотбора по характеристикам обучения проявляется в том, что в онлайн-опросе в большей степени участвуют студенты определенных направлений подготовки, что не накладывает существенных ограничений на выводы, сделанные в результате исследования при анализе данных при контроле направления подготовки.

Социально-демографические характеристики

При сравнении распределений по полу среди участников исследования и тех, кто предпочел не участвовать, были найдены значимые различия: среди участников лонгитюдного обследования на 10% больше представительниц женского пола, в то время как среди первокурсников, имеющих другие статусы участия в опросе, имеет место обратное соотношение юношей и девушек⁹ [Табл. 4]. Поскольку соотношение девушек и юношей на разных направлениях подготовки

⁹ Различия значимы, хи-квадрат = 10.644, df=2, p = 0,005

Таблица 4. Распределение первокурсников по полу в зависимости от статуса их участия в исследовании

	Мужской пол (N=465)	Женский пол (N=447)
Не заполнял ни одну анкету (N=323)	37,8%	32,9%
Заполнил только анкету первокурсника (N=305)	35,9%	30,9%
Участник лонгитюдного исследования (N=284)	26,3%	36,2%

разное, а предыдущий анализ показал, что студенты определенных направлений подготовки в меньшей степени склонны участвовать в исследовании, было решено проанализировать взаимосвязь пола и статуса участия для каждого направления подготовки в отдельности. При анализе распределений респондентов по полу с учетом направления подготовки, обнаруженные значимые различия исчезают для всех специальностей, кроме социологии, что, скорее всего, связано с относительно небольшим количеством наблюдений [Табл. 5]. Однако выявленное соотношение, заключающееся в том, что доля первокурсниц, принявших участие в исследовании, больше, чем доля первокурсников примерно на 10%, и это соотношение сохраняется для всех специальностей, кроме «экономики» и «экономики и статистики». Полученные результаты позволяют нам говорить о том, что принадлежность к определенному полу влияет на принятие решения об участии в исследовании, что подтверждает результаты предыдущих исследований эффекта самоотбора в студенческих аудиториях [4; 18; 19].

При сравнении региона проживания до поступления в университет первокурсников, обладающих разным статусом участия в исследовании, было обнаружено, что более склонны участвовать в исследовании были студенты, проживающие в Московской области. При этом студенты, приехавшие из других регионов, были менее склонны участвовать в опросе по анкете первокурсника, что, возможно, связано с тем, что опрос проходил в первую неделю сентября, и часть иногородних студентов еще не успели устроиться в общежитии и получить доступ к интернету [Табл. 6]. Обнаружен-

Таблица 5. Распределение первокурсников по полу в зависимости от статуса их участия в исследовании, в разрезе направлений подготовки

		жен.	муж.			жен.	муж.
Политология ¹	Не заполнял ни одну анкету (N=25)	31,0%	35,3%	Прикладная математика и информатика ²	Не заполнял ни одну анкету (N=50)	17,3%	28,9%
	Заполнил только анкету первокурсника (N=24)	28,6%	35,3%		Заполнил только анкету первокурсника (N=75)	40,4%	38,0%
	Участник лонгитюдного исследования (N=27)	40,4%	29,4%		Участник лонгитюдного исследования (N=69)	42,3%	33,1%
Программная инженерия ³	Не заполнял ни одну анкету (N=43)	36,8%	38,3%	Психология ⁴	Не заполнял ни одну анкету (N=34)	40,0%	57,1%
	Заполнил только анкету первокурсника (N=43)	31,6%	39,4%		Заполнил только анкету первокурсника (N=17)	26,2%	0,0%
	Участник лонгитюдного исследования (N=27)	31,6%	22,3%		Участник лонгитюдного исследования (N=28)	33,8%	42,9%
Совместная программа по экономике НИУ ВШЭ и РЭШ ⁵	Не заполнял ни одну анкету (N=28)	37,5%	46,3%	Социология ⁶	Не заполнял ни одну анкету (N=28)	22,3%	29,2%
	Заполнил только анкету первокурсника (N=22)	33,3%	34,2%		Заполнил только анкету первокурсника (N=35)	25,5%	45,8%
	Участник лонгитюдного исследования (N=15)	29,2%	19,5%		Участник лонгитюдного исследования (N=55)	52,1%	25,0%
Экономика ⁷	Не заполнял ни одну анкету (N=88)	41,3%	45,9%	Экономика и статистика ⁸	Не заполнял ни одну анкету (N=27)	40,4%	44,4%
	Заполнил только анкету первокурсника (N=70)	34,6%	34,7%		Заполнил только анкету первокурсника (N=19)	29,8%	27,8%
	Участник лонгитюдного исследования (N=44)	24,0%	19,4%		Заполнил только анкету первокурсника (N=19)	29,8%	27,8%

¹ Хи-квадрат = 1.024, df=2, p=0.599

² Хи-квадрат = 2.938, df=2, p=0.230

³ Хи-квадрат = 0.825, df=2, p=0.662

⁴ Хи-квадрат = 4.712, df=2, p=0.095

⁵ Хи-квадрат = 0.889, df=2, p=0.641

⁶ Хи-квадрат = 6.051, df=2, p=0.049

⁷ Хи-квадрат = 0.743, df=2, p=0.690

⁸ Хи-квадрат = 0.087, df=2, p=0.958

Таблица 6. Распределение первокурсников по региону проживания в зависимости от статуса участия в исследовании

	Москва (N=386)	Московская область (N=119)	Другой регион (N=409)
Не заполнял ни одну анкету (N=323)	33,9%	26,9%	39,1%
Заполнил только анкету первокурсника (N=307)	36,5%	31,1%	31,5%
Участник лонгитюдного исследования (N=284)	29,6%	42,0%	29,4%

ные различия по региону значимы¹⁰. Однако нет значительных различий по такой характеристике, как тип населенного пункта¹¹.

В рамках опроса абитуриентов респондентов просили оценить материальное положение своей семьи. Мы использовали эти ответы для того, чтобы проверить, есть ли отличия в материальном положении первокурсников в зависимости от их решения принять участие в опросе первокурсника и в лонгитюдном исследовании. Проведенный анализ показал отсутствие значимых различий по данной характеристике¹².

Полученные результаты позволяют нам говорить о том, что с определенными погрешностями, выборка исследования соответствует генеральной совокупности по социально-демографическим показателям, за исключением соотношения по полу.

Уровень исходной подготовки и академическая успеваемость

Между респондентами, согласившимися на участие в лонгитюдном исследовании, теми, кто заполнил анкету первокурсника, а так-

¹⁰ Различия значимы. Хи-квадрат=11.274, df=4, p=0.024

¹¹ Различия не значимы. Хи-квадрат=9.836, df=10, p=0.455

¹² Различия не значимы. Хи-квадрат=14.977, df=8, p=0.060

же теми, кто не заполнил ни одну из анкет, не было обнаружено значимых различий, связанных с исходным уровнем подготовки, о чем свидетельствует отсутствие статистических отличий в баллах ЕГЭ студентов по следующим дисциплинам (указаны предметы, ЕГЭ по которым сдавали не менее 50 студентов):

- английский язык¹³;
- биология¹⁴;
- информатика¹⁵;
- история¹⁶;
- математика¹⁷;
- обществознание¹⁸;
- русский язык¹⁹;
- физика²⁰.

Между указанными группами первокурсников также нет значимых различий по такой характеристике, как наличие медали (диплома с отличием)²¹. Кроме того, студенты, участвовавшие в первой волне лонгитюдного исследования, заполнившие только анкету первокурсника и не участвовавшие ни в одном из опросов, не отли-

¹³ Различия не значимы (анализ проводился с помощью одномерного дисперсионного анализа, $N=588$, $F=1.037$, $p=0.411$)

¹⁴ Различия не значимы (анализ проводился с помощью одномерного дисперсионного анализа, $N=80$, $F=0.811$, $p=0.704$)

¹⁵ Различия не значимы (анализ проводился с помощью одномерного дисперсионного анализа, $N=314$, $F=1.450$, $p=0.094$)

¹⁶ Различия не значимы (анализ проводился с помощью одномерного дисперсионного анализа, $N=116$, $F=0.877$, $p=0.643$)

¹⁷ Различия не значимы (анализ проводился с помощью одномерного дисперсионного анализа, $N=897$, $F=1.191$, $p=0.201$)

¹⁸ Различия не значимы (анализ проводился с помощью одномерного дисперсионного анализа, $N=501$, $F=1.373$, $p=0.071$)

¹⁹ Различия не значимы (анализ проводился с помощью одномерного дисперсионного анализа, $N=898$, $F=0.900$, $p=0.603$)

²⁰ Различия не значимы (анализ проводился с помощью одномерного дисперсионного анализа, $N=235$, $F=1.350$, $p=0.107$)

²¹ Различия не значимы. Хи-квадрат = 1.633, $df=2$, $p=0.442$

чаются значимо по их академической успеваемости по результатам первого и второго семестра²².

Полученные результаты позволяют сделать предположение, что участники лонгитюдного исследования не отличаются существенно от своих одноклассников по успешности учебной деятельности в университете, а также уровню своей изначальной подготовки.

Лояльность к университету и к выбранному направлению подготовки

На основе данных опроса абитуриентов, который охватил 83% первокурсников, рассмотрим различия участников лонгитюдного исследования и других студентов по показателям, отражающим лояльность к НИУ ВШЭ и к выбранному направлению подготовки.

Во время подачи заявления в НИУ ВШЭ абитуриентов спрашивали, собираются ли они подавать (или уже подали) заявление на поступление в другие вузы. Мы сравнили ответы студентов на данный вопрос в зависимости от того, участвовали ли они в лонгитюдном исследовании, заполнили ли только анкету первокурсника, либо не участвовали ни в одном опросе. Проведенный анализ показал отсутствие значимых различий по данной переменной в зависимости от статуса участия в исследовании²³. Также нет статистически значимых отличий респондентов в зависимости от осознанности выбора направления подготовки²⁴, вовлеченности в поиск информации об обучении в НИУ ВШЭ на сайте²⁵, количества времени, которое прошло с момента принятия решения поступать в НИУ ВШЭ²⁶.

²² Анализ проводился с помощью однофакторного дисперсионного анализа, зависимая переменная — средний балл в первом семестре: $F=1.020$, $p=0.430$, зависимая переменная — средний балл во втором семестре: $F=1.021$, $p=0.424$

²³ Различия не значимы. Хи-квадрат=3.760, $df=2$, $p=0.153$

²⁴ Различия не значимы. Хи-квадрат=7.889, $df=12$, $p=0.794$

²⁵ Различия не значимы. Хи-квадрат=0.387, $df=2$, $p=0.824$

²⁶ Различия не значимы. Хи-квадрат=13.820, $df=12$, $p=0.312$

Таким образом, на основе имеющихся данных нельзя сказать, что участники лонгитюдного исследования отличаются от других студентов по степени их лояльности к НИУ ВШЭ.

Цели поступления в университет

Из опроса абитуриентов нам также были доступны данные о целях поступления изучаемых нами студентов в вуз. Вопрос задавался в виде двух противоположных высказываний, где абитуриенту необходимо было указать, к какому высказыванию ближе его убеждения относительно обучения в университете (указав цифру от 1 до 5). Нами не было обнаружено значимых различий в ответах респондентов для следующих оппозиций:

- «Скорее важно получить высокий уровень знаний» — «Скорее важно просто получить диплом об окончании вуза»²⁷;
- «Скорее важно получить разносторонние, широкие знания» — «Скорее важно получить конкретные навыки, которые могут сразу использоваться»²⁸;
- «Скорее важно, чтобы было легко учиться» — «Скорее важно получить большой багаж знаний, даже если будет тяжело учиться»²⁹.

Однако наблюдаются некоторые статистически значимые различия в оценках близости следующих противоположных по смыслу высказываний:

- «Скорее важно окончить престижное учебное заведение» — «Скорее важно получить высшее образование, престиж учебного заведения не имеет значения»³⁰;

²⁷ Различия не значимы. Хи-квадрат=7.960, df=8, p=0.437

²⁸ Различия не значимы. Хи-квадрат=6.853, df=8, p=0.553

²⁹ Различия не значимы. Хи-квадрат = 3.563, df=8, p=0.894

³⁰ Различия значимы. Хи-квадрат = 16.889, df=8, p=0.031

- «Скорее важно получить диплом с высокими оценками» — «Скорее важно уделять внимание только тем предметам, которые интересны»³¹.

В отношении первой оппозиции участники лонгитюдного исследования значительно меньше выбирали промежуточный вариант ответа по сравнению с участниками только опроса первокурсников и тех, кто не участвовал ни в одном исследовании. В отношении второй оппозиции по выбору промежуточного пункта они значимо отличаются только от тех, кто не участвовали ни в одном опросе.

Учитывая полученные результаты, можно сделать вывод, что участники лонгитюдного исследования не продемонстрировали отличий в целях поступления в университет, которые могли бы существенно повлиять на их ответы на анкету исследования.

Оценка эффекта самоотбора на основе данных опроса первокурсников

Далее проведем сравнение студентов, участвовавших только в опросе первокурсников и студентов, согласившихся участвовать в лонгитюдном исследовании, на основе данных, собранных в ходе опроса первокурсников.

Проведенный анализ не показал значимых отличий студентов, решивших участвовать в лонгитюдном исследовании, от тех, кто от такого участия отказался, по следующим показателям:

- проживание в общежитие или у родителей/родственников в этом учебном году³²;
- время, которое будет тратить первокурсник на дорогу до учебы³³;
- наличие друзей из НИУ ВШЭ, которых знали до поступления³⁴;
- важность получения только хороших и отличных оценок³⁵;

³¹Различия значимы. Хи-квадрат = 19.644, df=8, p=0.012

³² Различия не значимы. Хи-квадрат=1.422, df=4, p=0.840

³³ Различия не значимы. Хи-квадрат=3.848, df=6, p=0.697

³⁴Различия не значимы. Хи-квадрат=1.022, df=1, p=0.312

³⁵Различия не значимы. Хи-квадрат=0.795, df=4, p=0.939

- важность возможности во время обучения посещать спортивные секции, состоять в университетской спортивной команде³⁶;
- важность возможности во время обучения подружиться с другими студентами³⁷;
- важность участия в научных исследованиях и проектах³⁸;
- важность посещения студенческих вечеринок³⁹;
- важность возможности получить красный диплом (с отличием)⁴⁰;
- важность возможности повысить свою эрудированность по результатам обучения⁴¹;
- важность возможности стать специалистом в выбранной профессиональной области⁴²;
- важность возможности установить полезные связи, которые помогут при построении карьеры⁴³;
- важность возможности получения возможности продолжить обучение в зарубежном университете⁴⁴;
- важность получения знания и опыта, которые позволят найти работу за границей⁴⁵;
- важность возможности найти высокооплачиваемую работу по результатам обучения⁴⁶;
- важность развить свои организаторские способности⁴⁷.

³⁶Различия не значимы. Хи-квадрат=1.919, df=4, p=0.751

³⁷Различия не значимы. Хи-квадрат=2.055, df=4, p=0.726

³⁸Различия не значимы. Хи-квадрат=7.310, df=4, p=0.120

³⁹Различия не значимы. Хи-квадрат=4.776, df=4, p=0.311

⁴⁰Различия не значимы. Хи-квадрат=2.229, df=4, p=0.694

⁴¹Различия не значимы. Хи-квадрат=6.489, df=4, p=0.165

⁴²Различия не значимы. Хи-квадрат=1.406, df=4, p=0.843

⁴³Различия не значимы. Хи-квадрат=4.726, df=4, p=0.317

⁴⁴Различия не значимы. Хи-квадрат=4.875, df=4, p=0.300

⁴⁵Различия не значимы. Хи-квадрат=3.241, df=4, p=0.518

⁴⁶Различия не значимы. Хи-квадрат=2.008, df=4, p=0.734

⁴⁷Различия не значимы. Хи-квадрат=3.771, df=4, p=0.438

Наблюдаются значимые различия в оценке важности участия в творческих и культурных мероприятиях во время обучения в университете среди участников лонгитюдного исследования и теми первокурсниками, кто заполнил только анкету первокурсника⁴⁸. Первая группа студентов в среднем выше оценила важность данной активности для них во время обучения в университете [Табл. 7]. При этом похожее соотношение ответов сохраняется при анализе в разрезе направлений подготовки и в отдельности для женского и мужского пола.

Также участники лонгитюдного исследования значительно отличаются от других респондентов по их оценке важности возможности улучшить знания иностранного языка в результате обучения⁴⁹ [Табл. 8].

Таблица 7. Важность участия в творческих кружках и культурных мероприятиях в университете в зависимости от статуса участия в исследовании

	Заполнил только анкету первокурсника (N=307)	Участник лонгитюдного исследования (N=283)
Совсем не важно	8,5%	5,3%
Скорее не важно	33,6%	26,1%
Скорее важно	43,0%	47,0%
Очень важно	11,1%	20,5%
Затрудняюсь ответить	3,9%	1,1%

При этом по многим показателям культурного капитала студентов не было обнаружено значимых отличий. Так, не наблюдается различий в степени согласия со следующими утверждениями:

⁴⁸ Различия значимы. Хи-квадрат=18.421, df=4, p=0.001

⁴⁹ Различия значимы. Хи-квадрат=9.848, df=4, p=0.043

- «я люблю классическую музыку»⁵⁰;
- «я знаю большинство известных композиторов»⁵¹;
- «я хорошо разбираюсь в литературе»⁵²;
- «я часто посещаю музеи, театры, концерты»⁵³;
- «я часто читаю книги»⁵⁴;
- «когда я был (а) ребенком, близкие родственники поощряли мой интерес к чтению»⁵⁵;
- «у моей семьи много книг»⁵⁶;
- «когда я учился (ась) в школе, я часто посещал (а) музыкальную/ художественную школу/ творческие кружки»⁵⁷;

Таблица 8. Важность улучшить знания иностранного языка в зависимости от статуса участия в исследовании

	Заполнил только анкету первокурсника (N=307)	Участник лонгитюдного исследования (N=283)
Совсем не важно	0,3%	1,1%
Скорее не важно	4,6%	1,1%
Скорее важно	25,4%	21,2%
Очень важно	69,4%	76,0%
Затрудняюсь ответить	0,3%	0,7%

⁵⁰ Различия не значимы. Хи-квадрат=8.625, df=4, p=0.071

⁵¹ Различия не значимы. Хи-квадрат=4.691, df=4, p=0.321

⁵² Различия не значимы. Хи-квадрат=2.736, df=4, p=0,603

⁵³ Различия не значимы. Хи-квадрат=3.097, df=4, p=0.542

⁵⁴ Различия не значимы. Хи-квадрат=8.039, df=4, p=0.090

⁵⁵ Различия не значимы. Хи-квадрат=8.538, df=4, p=0.074

⁵⁶ Различия не значимы. Хи-квадрат=2.186, df=4, p=0,702

⁵⁷ Различия не значимы. Хи-квадрат=5.343, df=4, p=0.254

Отличия наблюдаются в оценке степени согласия с высказыванием «Я люблю читать художественные книги»⁵⁸. При этом участники лонгитюдного исследования в среднем склонны называть большее количество книг, прочитанных (прослушанных) за год [Табл. 9].

Таблица 9. Количество книг, прочитанных за год в зависимости от статуса участия в исследовании

	Заполнил только анкету первокурсника (N=307)	Участник лонгитюдного исследования (N=283)
Ни одной	0,7%	0,4%
Менее 5 книг	14,7%	11,0%
От 5 до 10 книг	34,5%	30,7%
Более 10 книг	43,6%	54,4%
Затрудняюсь ответить	6,5%	3,5%

Вопрос: Сколько примерно книг (художественных, научно-популярных и любых других, не считая учебников) Вы прочли (прослушали) в этом году?

По посещению социальных сетей участники лонгитюдного исследования отличаются только относительно сети «ВКонтакте»⁵⁹: они более склонны, чем другие указывать эту социальную сеть в качестве посещаемой. Кроме того, они также более склонны отмечать, что проводят более двух часов, просматривая страницы социальных сетей [Табл. 10⁶⁰].

При этом эффекта самоотбора не наблюдается в отношении образования матери⁶¹ и отца⁶² первокурсника.

⁵⁸Различия значимы. Хи-квадрат=10.510, df=4, p=0.033

⁵⁹Различия значимы. Хи-квадрат=4.977, df=1, p=0.026

⁶⁰Различия значимы. Хи-квадрат=7.979, df=2, p=0.019

⁶¹Различия не значимы. Хи-квадрат=7.319, df=6, p=0.292

⁶²Различия не значимы. Хи-квадрат=8.898, df=6, p=0.179

Таблица 10. Время, затрачиваемое на посещение социальных сетей в зависимости от статуса участия в исследовании

	Заполнил только анкету первокурсника (N=279)	Участник лонгитюдного исследования (N=264)
Менее часа	28,00%	23,50%
От 1 до 2 часов	44,4%	37,5%
Более 2 часов	27,6%	39,0%

Эффект самоотбора во второй волне исследования

Во второй волне исследования приняли участие 257 первокурсников, еще 6 начали заполнять анкету и не закончили, 1 студент отказался от участия во второй и последующих волнах, и еще 19 студентов были отчислены из университета к моменту второго опроса. Таким образом, не считая отчисленных студентов, отклик во второй волне составил 97%. Одной из причин такого высокого отклика стало использование напоминаний о возможности принять участие в исследовании по телефону наряду с рассылкой анкет на электронные адреса студентов. Эффект самоотбора во второй волне мог проявляться на этапе до телефонных напоминаний, поэтому в данном разделе мы предприняли попытку оценить различия респондентов по основным параметрам: исходная подготовка и академическая успеваемость, пол, направление подготовки и форма финансирования обучения.

Значимые различия в исходной подготовке респондентов в зависимости от периода их ответа на опрос были обнаружены только для результатов ЕГЭ по английскому языку: студенты, ответившие на вторую анкету до обзвона, в среднем имеют более высокие баллы ЕГЭ по английскому языку по сравнению со студентами, ответившими на анкету в результате обзвона [Табл. 11].

При этом различия в академической успеваемости между рассматриваемыми группами не значимы для первого семестра, но являются существенными для результатов второго семестра: более ранние респонденты имеют в среднем выше академическую успеваемость [Табл. 12].

Таблица 11. Средний балл ЕГЭ по дисциплинам в зависимости от периода ответа на опрос

	Ответили до обзвона	Ответили после обзвона
Английский язык ¹	84,63 (N=111)	77,96 (N=56)
Информатика ²	81,68 (N=47)	82,32 (N=37)
Математика ³	76,75 (N=161)	75,61 (N=89)
Обществознание ⁴	80,25 (N=92)	76,29 (N=49)
Русский язык ⁵	86,47 (N=161)	87,01 (N=89)
Физика ⁶	72,28 (N=29)	71,46 (N=28)

¹ Различия значимы. $F=8.552$, $p=0.003$ ² Различия не значимы. $F=1.933$, $p=0.151$ ³ Различия не значимы. $F=0.311$, $p=0.733$ ⁴ Различия не значимы. $F=2.897$, $p=0.058$ ⁵ Различия не значимы. $F=0.166$, $p=0.847$ ⁶ Различия не значимы. $F=1.621$, $p=0.207$

Таблица 12. Академическая успеваемость участников лонгитюдного исследования в зависимости от периода ответа на опрос

	Ответили до обзвона	Ответили после обзвона
Средний балл за 1-ый семестр (по 10-балльной шкале) ¹	7,41 (N=156)	7,29 (N=88)
Средний балл за 2-ой семестр (по 10-балльной шкале) ²	7,39 (N=153)	6,81 (N=87)

¹ Различия не значимы. $F=1.986$, $p=0.139$ ²¹ Различия значимы. $F=8.552$, $p<0.001$

Также ранние и более поздние респонденты второй волны различаются по полу⁶³: среди участников, ответивших на анкету до обзвона, на 12% больше девушек [Табл. 13].

Не наблюдается значимых различий по направлениям подготовки⁶⁴ и форме финансирования обучения⁶⁵.

⁶³ Различия значимы. Хи-квадрат=6.729, df=2, $p=0.035$ ⁶⁴ Различия не значимы. Хи-квадрат=13.571, df=14, $p=0.482$ ⁶⁵ Различия не значимы. Хи-квадрат=2.533, df=2, $p=0.282$

Таблица 13. Статус участия респондентов во второй волне в зависимости от пола

	Мужской пол (N=111)	Женский пол (N=150)
Отвечили до обзвона (N=166)	56,8%	68,7%
Отвечили после обзвона (N=89)	38,7%	30,7%
Не ответили после обзвона (N=6)	4,5%	0,7%

Обсуждение результатов

Проведенный анализ эффекта самоотбора студентов в результате реализации лонгитюдного исследования на первом году позволил еще раз показать, что студенты женского пола более склонны участвовать в социологических исследованиях по сравнению с одноклассниками мужского пола. Однако не было выявлено различий в уровне исходной подготовки студентов и в их последующей академической успеваемости в университете, что противоречит выводам, полученным в предыдущих исследованиях студентов, выполненных за рубежом. Однако можно высказать предположение о том, что эффект влияния успеваемости на желание принять участие в исследовании может проявиться в ходе дальнейшей реализации исследования, что находит подтверждение в некоторых результатах проведенного анализа. Так, при сравнении характеристик ранних и поздних респондентов, ответивших на анкету второй волны исследования, было выявлено, что группа поздних респондентов отличается более низкой успеваемостью по сравнению с группой ранних респондентов. В целом можно отметить, что статистически значимые различия наблюдаются только по небольшому набору переменных, что позволяет нам говорить о незначительности эффекта самоотбора и его влиянии на качество данных при условии, что исследователи учтут в своем анализе наличие смещений по полу. Второй значимый вывод данной работы состоит в том, что при лонгитюдном дизайне исследования сочетание e-mail рассылки со звонками респондентам и напоминанием им об исследовании позволяет не только добиться более высокого отклика на последующие волны исследований, но и предотвратить смещения,

вызванные отличием респондентов, заполнивших анкеты до напоминаний по телефону и после.

ЛИТЕРАТУРА

1. Couper M.P. Review: Web surveys: A review of issues and approaches // *Public opinion quarterly*. 2000. Vol. 64. №4. P. 464-494.
2. Bethlehem J. Selection bias in web surveys // *International Statistical Review*. 2010. Vol. 78. № 2. P. 161-188.
3. Hudson D., Seah L.H., Hite D., Haab T. Telephone presurveys, self-selection, and non-response bias to mail and internet surveys in economic research // *Applied Economics Letters*. 2004. Vol. 11. №4. P. 237-240.
4. Porter S. R., Whitcomb M. E. Non-response in student surveys: The role of demographics, engagement and personality // *Research in higher education*. 2005. Vol. 46. №. 2. P. 127-152.
5. Groves R.M., Peytcheva E. The impact of nonresponse rates on nonresponse bias a meta-analysis // *Public opinion quarterly*. 2008. Vol. 72. №. 2. P. 167-189.
6. Lahaut V. M. et al. Estimating non-response bias in a survey on alcohol consumption: comparison of response waves // *Alcohol and Alcoholism*. 2003. Vol. 38. №. 2. P. 128-134.
7. Whitehead J.C., Groothuis P.A., Blomquist G.C. Testing for non-response and sample selection bias in contingent valuation: analysis of a combination phone/mail survey // *Economics Letters*. 1993. Vol. 41. №2. P. 215-220.
8. Kojetin B.A., Borgida E., Snyder M. Survey topic involvement and nonresponse bias // *Proceeding of the Section on Survey Research Method*. 1993. Vol. 2. P. 838-843.
9. Groves R.M., Singer E., Corning A. Leverage-saliency theory of survey participation: Description and an illustration // *Public Opinion Quarterly*. 2000. P. 299-308.
10. Van Kenhove P., Wijnen K., De Wulf K. The influence of topic involvement on mail-survey response behavior // *Psychology & Marketing*. 2002. Vol. 19. №. 3. P. 293-301.
11. Groves R.M., Couper M. P. Contact-level influences on cooperation in face-to-face surveys // *Journal of Official Statistics*. 1996. Vol. 12. №. 1. P. 63.
12. Curtin R., Presser S., Singer E. The effects of response rate changes on the index of consumer sentiment // *Public Opinion Quarterly*. 2000. Vol. 64. №. 4. P. 413-428.
13. Singer E., Van Hoewyk J., Maher M. P. Experiments with incentives in telephone surveys // *Public Opinion Quarterly*. 2000. Vol. 64. №. 2. P. 171-188.
14. Moore D. L., Tarnai J. Evaluating nonresponse error in mail surveys // *Survey nonresponse*. 2002. P. 197-211.
15. Kandel D., Raveis V., Logan J. Sex differences in the characteristics of members lost to a longitudinal panel: A speculative research note // *Public Opinion Quarterly*. 1983. Vol. 47. №. 4. P. 567-575.
16. Singer E., Groves R. M., Corning A. D. Differential incentives: Beliefs about practices, perceptions of equity, and effects on survey participation // *Public Opinion Quarterly*. 1999. P. 251-260.
17. Goyder J., Warriner G.K., Miller S. Evaluating socio-economic status (SES) bias in survey nonresponse // *Journal of Official Statistics*. 2002. Vol. 18. №. 1. P. 1-11.
18. Dey E.L. Working with low survey response rates: The efficacy of weighting adjustments // *Research in higher education*. 1997. Vol. 38. №2. P. 215-227.
19. Sax L.J., Gilmartin S.K., Bryant A.N. Assessing response rates and nonresponse bias in web and paper surveys // *Research in higher education*. 2003. Vol. 44. №4. P. 409-432.

Раздел 5

Методология онлайн-исследований: геймификация

Использование исследовательских игр вместо геймифицированных опросов. Влияние метода исследовательских игр на вовлеченность респондентов и вероятность их будущего участия в подобных проектах

*Бетти Адаму
(Research Through Gaming Ltd),
доктор Дэвид Беркс
(Бизнес-школа при Университете Винчестера)*

1. Введение

1.1 В каких онлайн-исследованиях применялась геймификация — краткий исторический контекст

Метод геймификации использовался по всему миру в таких отраслях, как маркетинг, здравоохранение, управление и образование [16], для повышения уровня вовлеченности людей, занимающихся различными видами деятельности. Сам термин «геймификация» пока остается неустоявшимся [17, 18], однако существует его распространенное определение — «использование игровых элементов и игровой механики в неигровом контексте для повышения уровня вовлеченности людей и изменения их модели поведения». В работах [17,18,19] показано, как можно использовать игры для обеспечения высокой мотивации к участию в различных сферах деятельности. Для понимания данного определения геймификации необходимо пояснить, что понимается под игровой механикой и игровыми элементами. Игровая механика, считающаяся основой любой игры, — это правила, система/системы обратной связи, цель/цели и добровольный характер участия [20]. Игровые элементы

включают в себя прочие элементы игр — от нарративов, виртуальной среды и графики до аватаров и аудиоэффектов [21].

В маркетинговых исследованиях респонденты зачастую получают денежные вознаграждения и розыгрыши призов, что побуждает их принимать участие в опросах [22]. Однако материальное вознаграждение человека за выполнение поставленной задачи не всегда является надежным методом [23,24]. Поэтому, несмотря на его использование, в отрасли маркетинговых исследований наблюдается снижение уровня вовлеченности респондентов [25]. Это уже привело к снижению детализации ответов участников, повышению доли случаев слепого заполнения [26] и ухудшению качества и надежности получаемых данных [27]. Такие последствия, в свою очередь, могут отрицательно сказаться на качестве решений, принимаемых организациями на основе результатов маркетинговых исследований.

Другой серьезный недостаток низкого уровня вовлеченности респондентов — его негативное влияние на работу компаний, создающих панели для опросов и предоставляющих другим маркетинговым агентствам и корпорациям доступ к респондентам. Для получения необходимого количества заполненных анкет, уже гарантированных компанией своему клиенту, они постоянно покупают друг у друга панелистов (и при этом также могут конкурировать друг с другом). Проще говоря, компании, работающие с панелями для опросов, переживают довольно сложные времена, стараясь предоставить необходимое число респондентов в условиях низкого уровня их вовлеченности [28]. Практика приобретения панелистов у конкурентов уже стала обычной, и это привело к возникновению новых проблем в исследовательской отрасли, таких как сокращение прибыли от каждого проекта и повышение степени единообразия панелей разных компаний.

Проблема низкого уровня вовлеченности участников маркетинговых онлайн-исследований оказала существенное влияние на процессы в этой отрасли и ее дальнейшее развитие. Для ее решения в онлайн-опросы начали внедрять геймификацию [9], [29]. Это привело к целому ряду как положительных, так и отрицательных результатов. Среди положительных можно назвать, главным образом, повышение уровня вовлеченности [30]. Однако возможности при-

менения геймификации в онлайн-исследованиях также имеют свои ограничения, потому результаты наблюдались не только положительные, но и отрицательные [31].

Принято считать, что ограничения «геймифицированных» опросов связаны как с действиями разработчика, так и с особенностями геймификации как «инструмента вовлечения» (несмотря на то, что само слово «геймификация» остается неоднозначно понимаемым термином [32]), поскольку она также включает в себя элементы внешней мотивации участников. В большинстве случаев геймификация вводит такие внешние мотиваторы, как система очков и значков [33], применяя подход «сделай х, получи у», или метод кнута и пряника. Этот подход был представлен в теории оперантного обусловливания, предложенной Скиннером в 1937 г. [34]. Таким образом, можно утверждать, что в геймификации нет ничего «нового» [34] и она не может сформировать внутреннюю мотивированность и вовлеченность человека, в отличие от самой игры, которую она имитирует [11].

Дальнейшие ограничения при использовании геймификации в исследованиях обусловлены способом ее реализации. Большинство «геймифицированных» опросов, созданных до сегодняшнего дня, используют такие приемы, как, например, переформулировка вопросов, призванная сделать их более сложными [5], [35,36] в целях повышения вовлеченности респондентов. Слип (2012 [интернет-ресурс]) приводит следующий пример: «Вместо того чтобы спросить, какую одежду респондент любит носить, мы спрашиваем его, какую одежду он надел бы на первое свидание». Подобная постановка вопросов может привести к смещению в ответах.

Другой способ «геймификации» опросов — это внедрение графических элементов, не связанных с темой исследования [30, 31], [8]. Принято считать, что исследования на игровой основе должны быть четко связаны с темой исследования [37,38]. Без релевантных, целостных и близких к теме исследования игровой механики и игровых элементов, связанных с поставленными задачами, его создатели рискуют запутать участников, как отмечено в работе Доунса-Ле Гуина и других авторов [31], или отвлечь их от основной цели, как утверждает Мавлетова [8]. В действительности оба эти отрицательных

эффекта — лишь часть того, что может произойти при геймификации различных заданий и видов деятельности [12].

В данной работе рассматриваются ситуации, когда при проведении социальных и научных исследований в интернете для сбора и анализа данных вместо геймификации были использованы полноценные *игры*. Этот тип игр, предназначенный для проведения онлайн-опросов, получил название «исследовательской игры» [7], [37,38,39]. Исследовательские игры позволяют повысить уровень вовлеченности респондентов и эмоциональность их ответов, что необходимо для получения актуальной и полностью соответствующей контексту информации.

1.2 Разница между геймифицированными опросами и исследовательскими играми

Исследовательские игры — это особый метод разработки онлайн-исследований. В его рамках исследователь рассматривает процесс сбора данных с точки зрения дизайнера игр, однако руководствуясь целями исследования. Предпочтение отдается игровым элементам, которые соответствуют и согласуются с содержанием и задачами исследования. По сути, игровые элементы и игровая механика в исследовательских играх должны быть непосредственно соотнесены с содержанием исследования. В этом заключается существенное отличие от метода геймификации онлайн-исследований, который часто реализуется с позиций исследователя путем «добавления» элементов, считающихся «геймификацией», хотя простое их использование не соответствует приведенному выше определению геймификации.

Нарратив как игровой элемент играет все более важную роль в современных цифровых играх, он принципиален для вовлечения пользователей в игровой процесс. В видеоиграх и онлайн-играх, таких как *Grand Theft Auto 5*, *Max Payne*, *Gone Home*, *Unfinished Swan* и *World of Warcraft*, предыстории персонажей и повествованию в целом уделено не меньше внимания, чем другим элементам и игровой механике. Истории в играх или рассказы об отдельных персонажах сами по себе стали стимулом для вовлечения игроков и источником развлечения в современной поп-культуре [40], что до-

казывает широкое медийное покрытие построенных на нарративе цифровых игр, таких как Rapo and Yo (Minority Media, 2013).

Результативность геймификации и, следовательно, «геймифицированных» опросов может быть ограничена в плане вовлечения респондентов в тех случаях, когда разработчики исследования исключают повествование целиком или же вводят сюжет, зачастую не соотносящийся и не согласующийся с целями исследования. Соотносящиеся с нарративом системы вознаграждения (в которых вознаграждение и признание достижений участников соответствует содержанию игры) важны и необходимы в игровом процессе цифровых игр — и именно этого не хватает многим «геймифицированным» опросам. Отсутствие связного и релевантного нарратива разрушает жизненно важный элемент игрового дизайна — возможность временного принятия предложенного допущения, что является одним из ключевых факторов, способствующих погружению и вовлечению в игру [41].

Таким образом, нарратив как элемент исследовательской игры может стать одним из инструментов достижения целей исследования. В данной работе будет рассмотрено повествование, использованное в двух исследовательских игр, взятых в качестве примеров. Будут также даны необходимые пояснения о том, *почему* для этих исследовательских игр выбраны именно такие сюжеты, как они соотносятся с целями исследования и помогают их достичь, повышая уровень вовлеченности респондентов, преодолевая «эмпатический разрыв» и вызывая у участников соответствующие контексту эмоции.

2. Подход к исследованию и методология

2.1 История исследования

Две исследовательские игры, представленные в данной работе, называются T.E.S.S.A. Undercover Agents (далее просто «TESSA») и Dubious. Они были созданы в ходе решения исследовательских задач, поставленных профессором Лизбет ван Зуунен¹ (далее просто

¹ <http://imprintsutures.org/professor-liesbet-van-zoonen/>

«Лизбет») и ее командой из IMPRINTS Futures². В системе отношений между авторами данной работы и командой IMPRINTS Futures последняя выступала в качестве клиента, а Research Through Gaming Ltd — в качестве исполнителя исследования.

Проект IMPRINTS Futures представляет собой трехлетнее научное исследование³, завершившееся в конце 2014 г., которое финансировал Совет по исследованиям в области инженерии и физики (EPSRC⁴). Целью исследования было изучение опыта пользователей, их табу, культурные, гендерные и возрастные различия и ожиданий от *современных* форм идентификации и аутентификации, а также от *футуристических* форм идентификации и аутентификации.

Сведения, полученные в ходе исследования IMPRINTS Futures, уже были использованы рядом государственных организаций в качестве источника информации о восприятии и ожиданиях общества в области управления процессами идентификации и аутентификации (далее «ПИА»), а также в целях разработки государственной политики в области поддержки создания и управления будущими формами ПИА [42].

Руководитель проекта со стороны клиента, профессор Лизбет ван Зуунен, включила возможность применения игр для проведения онлайн-исследований в свое первоначальное предложение, представленное EPSRC. Однако она признает, что на тот момент ей было неизвестно, существует ли в принципе платформа «игр для проведения исследований». Использование игры в качестве исследовательского инструмента стало отличительной чертой предложения Лизбет среди прочих представленных вариантов, и в итоге именно она и ее команда были выбраны для проведения исследования IMPRINTS Futures. Она оформила презентацию своего предложения организаторам в стиле «Apprentice» (популярная на Западе теле-

² Другие члены рабочей команды проекта IMPRINTS перечислены на сайте IMPRINTS, на странице: <http://imprintsfutures.org/about/>

³ Спонсором которого выступила организация EPSRC в рамках своей программы Global Uncertainties Program, EP/J005037/1

⁴ Интернет-сайт EPSRC: <http://www.epsrc.ac.uk/Pages/default.aspx>, дата обращения: 10.01.2014

игра). Лизбет и ее коллеги поделились на команды и представили лучшие, по их мнению, исследовательские методики, эксперименты и исследования, которые можно было бы реализовать на практике в рамках трехлетнего проекта IMPRINTS Futures. В итоге Лизбет с командой были признаны лучшими и получили средства на реализацию IMPRINTS Futures.

У Лизбет было множество причин предпочесть в своем предложении игры формату традиционных онлайн-опросов. Она чувствовала, что использование игры позволит поддерживать необходимый высокий уровень вовлеченности участников, особенно учитывая большое количество утверждений и форматов идентификации, которые предстояло протестировать пользователям. Как опытный полевой исследователь, теоретик и статистик она *«знала, что такое усталость респондентов от участия в опросе и прочие ограничения традиционных онлайн-исследований»*.

В работе, написанной в соавторстве с коллегами на основе исследовательской игры TESSA, Лизбет обратила внимание, что одной из основных причин использования такого инструмента вовлечения участников исследования, как игры, для нее стало снижение доли откликов в обычных онлайн-опросах: *«И хотя поначалу этот феномен, казалось, ограничивался только почтовыми опросами, в Web-опросах теперь также заметны низкие показатели отклика респондентов. При этом зачастую отмечается, что причиной снижения активности, как бы парадоксально это ни звучало, является увеличение числа проводимых исследований»*. [43] У Лизбет были и другие причины прибегнуть к помощи игр. Она узнала, что *«Социальная сеть Facebook собирала данные о том, как пользователи играют в предложенные на ее страницах игры. В частности, она определяла стиль вашей игры — склонны ли вы рисковать или же предпочитаете тщательно планировать свои действия. Я подумала, что если это смогли реализовать в Facebook, то почему бы и IMPRINTS Futures не использовать эту возможность в своих играх»*.

Еще одной причиной для использования игр, по словам Лизбет, стало то, что *«В EPSRC хотели как раз такое исследование, выходящее за рамки представлений об «обычной науке», которое позволило бы впоследствии шаг за шагом привести к внедрению*

масштабных нововведений и, возможно, даже изменить саму парадигму управления идентификацией пользователей. Наша первоначальная идея состояла в том, чтобы использовать игру как средство распространения полученных результатов, однако организаторы захотели применить этот метод на более ранних этапах исследования».

Модерируемое Лизбет обсуждение целей исследования продолжалось несколько недель, и в конечном итоге были выработаны идеи по созданию исследовательской игры, впоследствии представленные в виде двух раскадровок (в качестве примера см. Рис. 1).

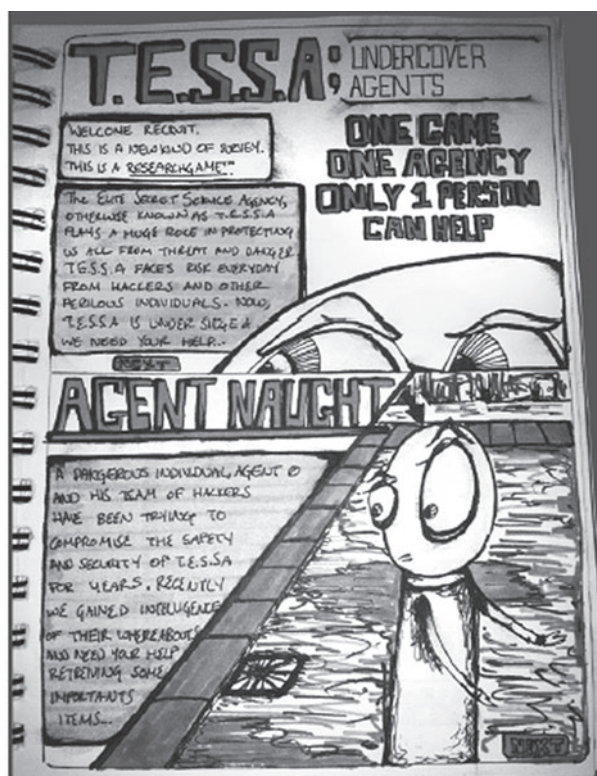


Рис. 1 Фрагмент раскадровки игры TESSA Undercover Agents перед началом ее разработки в цифровом виде.

2.2 Категоризация тематических направлений и задач исследования

В процессе обсуждения целей исследования и информации, которую необходимо было собрать посредством исследовательских игр, Лизбет также подготовила ряд утверждений. Впоследствии они были включены в игры для тестирования. В основе данных утверждений лежат существующие теории, которыми Лизбет решила воспользоваться в целях определения ключевых «типов личности» и профилирования участников (далее «игропдентов»⁵).

Первый этап процесса создания исследовательской игры включал в себя категоризацию задач исследования по содержанию и темам с использованием множества различных утверждений. Например, были установлены четкие разграничения между задачами и утверждениями, относящимися к уже существующим ПИА и тем, которые возможны в будущем. Впоследствии эти категории сформировали «уровни» исследовательских игр, при этом нарратив каждого из этих уровней напрямую соотносился с вопросами и утверждениями этого уровня, а также его общей тематикой.

Ввиду существования столь четкого разграничения в понимании существующих и будущих форматов ПИА, было решено создать не одну большую, а две исследовательские игры, охватывающие все задачи и темы проекта. Участие в одной большой исследовательской игре заняло бы у игропдента около полутора часов; были опасения, что участники исследования могут переутомиться, в результате качество данных оказалось бы под вопросом. В итоге нарратив исследовательской игры TESSA был выстроен таким образом, чтобы у игропдентов возникало ощущение настоящего времени, тогда как сюжет Dubious должен был мысленно перенести их в будущее — в 2030 г. Это было сделано для того, чтобы помочь игропдентам более связно и живо представить себе каждый из предложенных сценариев.

⁵ «Игропдент» (playspondent) — термин, придуманный Бетти Адаму и являющийся зарегистрированным товарным знаком Research Through Gaming Ltd, который используется для обозначения людей, являющихся в рамках исследовательских игр одновременно «игроками» и «респондентами».

2.3 Раскадровка

Следующим этапом процесса разработки двух исследовательских игр стали раскадровки. После создания они были представлены на согласование команде IMPRINTS Futures. Раскадровка помогла членам команды IMPRINTS Futures получить представление, как именно будут выглядеть исследовательские игры, каким образом они позволят решить поставленные задачи, собрать данные в рамках каждого из заданий, уровней или сценариев исследовательских игр. Эти раскадровки также демонстрировали, как именно нарративы, разработанные для исследовательских игр, помогут игродам лучше понять содержание последних, стимулировать собственные воспоминания о подобных ситуациях, а также имитировать впечатления, соответствующие задачам исследований.

Далее на основе обратной связи, данной командой IMPRINTS Futures, в раскадровки были внесены коррективы, направленные на более точное соответствие игр исследовательским вопросам. Затем последовали этап создания на основе согласованных раскадровок цифровой графики для TESSA и Dubious и, наконец, этап программирования графики.

2.4 Решения по использованию звукового сопровождения

В процессе разработки цифровой графики было выдвинуто предположение об использовании в игре музыки и звуковых эффектов. Клиент, то есть Лизбет, согласилась внести в проект соответствующее дополнение.

Роль музыки как крайне важного и благоприятно влияющего на игроков элемента доказана на примере множества игр. Музыкальное сопровождение таких игр, как World of Warcraft [44], оказало мощное эмоциональное воздействие на игроков. Доказана эффективность использования музыки не только в игровой индустрии. Так, например, музыкальное сопровождение помогает повысить производительность труда на рабочем месте [45], а также может служить средством сдерживания при проявлении антисоциальных моделей поведения [46], применяться для повышения продаж в ма-

газинах [47] и даже в танатологии⁶ [48]. Поэтому было принято решение использовать в играх TESSA и Dubious музыку и звуковые эффекты, а также выяснить, могут ли они оказать столь же позитивный эффект на вовлеченность игроков, изменение моделей их поведения и эмоциональную связь с контентом этих исследовательских игр. К примеру, в тех случаях, когда сценарий TESSA предполагал создание стрессовой ситуации, темп фоновой музыки нарастал, а мелодия приобретала более жесткое звучание для создания у игроков ощущения паники или напряжения. Принято считать, что эмоции влияют на процесс принятия решений [49]. В этой связи нас интересовало, сможем ли мы воссоздать в TESSA стрессовую ситуацию средствами нарратива, графического и звукового оформления, чтобы вызвать напряжение и увидеть, как участники принимают в предложенном им контексте решения под влиянием эмоций.

Звуковое оформление также было использовано для изменения скорости, с которой участники отвечали на вопросы и реагировали на различные игровые ситуации. Так, например, когда требовался спонтанный ответ игрока без размышлений над вопросом, звучала быстрая музыка, а на экране появлялся таймер с обратным отсчетом. В обеих исследовательских играх был создан целый ряд ситуаций со строгим ограничением по времени, при которых на экране демонстрировались таймеры с обратным отсчетом. Сочетание этого визуального элемента с соответствующей музыкой оказалось довольно мощным фактором воздействия на участников. Оно побуждало как можно быстрее ответить на вопрос или отреагировать на тот или иной поворот сюжета игры.

Поскольку во многих компьютерных играх звуковое сопровождение уже доказало свою способность обеспечивать более глубокое погружение игроков в игровой процесс, в данных исследовательских играх оно также было использовано для оказания на игроков подобного воздействия. За счет этого предполагалось усилить нарративы и обогатить предложенные сценарии.

⁶ Танатология занимается изучением возможностей использования музыки при уходе за людьми, находящимися при смерти.

Таким образом, звуковое сопровождение исследовательских игр было разработано для того, чтобы:

- 1) дополнить нарративы исследовательских игр;
- 2) содействовать имитации определенного опыта;
- 3) ускорить или замедлить реакцию игроков (где это необходимо);
- 4) содействовать проявлению эмоций, соответствующих целям исследования и побудить респондентов отвечать честно, так как их ответы будут основаны на реальных переживаниях;
- 5) привлечь внимание игроков к представленным демонстрируемым изображениям, дизайну игры и т.д.

3. Обзор процесса разработки исследовательских игр

3.1 Количество утверждений и форм идентификации для тестирования

Помимо различных задач, поставленных в рамках исследования, а также утверждений для проработки, Лизбет предоставила также формы идентификации, которые необходимо было протестировать в рамках исследовательских игр.

Исследовательская игра TESSA включала в себя:

- семь утверждений, с помощью которых необходимо было реагировать на все то, что могло произойти в ходе игры с банковскими картами игрока (например: «Я потерял (-а) свою карту» или «Моя карта была испорчена»);
- четырнадцать типов карт, которыми могли владеть игроки (например, кредитная карта, банковская карта, карта скидок);
- десять форм ПИА, существующих в наше время, которые необходимо было протестировать на основе предложенного контекста и задач игры;
- 26 утверждений на основе Шкалы культурных ценностей Хофстеде [50];
- четыре средства идентификации личности будущего, которые необходимо было протестировать по предпочтительности и вероятности внедрения;

- пятнадцать утверждений о пожеланиях в отношении будущих средств идентификации (например: *«Я хотел (-а) бы, чтобы моя карта издавала звуковой сигнал, когда рядом со мной оказываются преступники»*).

Всего в TESSA было представлено 48 утверждений, на которые должны были реагировать игроденты, а также 28 современных и будущих форм ПИА для тестирования по предпочтительности внедрения и использования.

Исследовательская игра Dubious включала в себя:

- три футуристических типа одежды для выбора и использования (они были представлены в первой части Dubious в разделе Avatar Creator Tool, где игроденты могли создать аватар для себя и еще одного человека, выбирая одежду и аксессуары из предложенных вариантов);
- девять футуристических аксессуаров, которые необходимо было протестировать (тестирование проводилось дважды в разделе Avatar Creator Tool: первый раз, когда игродент создавал аватар для себя, и второй раз, когда он создавал аватар для еще одного человека — своего друга, коллеги или учителя);
- четырнадцать футуристических форм ПИА, которые необходимо было протестировать по приоритетности их выбора и использования;
- пятнадцать существующих и футуристических форм ПИА, которые необходимо было протестировать по приоритетности их выбора и использования,
- двадцать утверждений, являвшихся частью теоретической структуры, которую Лизбет пожелала включить в исследование. Сюда входили утверждения, связанные с проблематикой социальной открытости, контроля, правительственных служб, ориентации на пользователя, коммерческой деятельности (в рабочей группе для их обозначения использовалась аббревиатура TSGUC);
- раздел в форме личного дневника для заполнения, описывающего «один день из жизни в 2030 году» с разбивкой на утро, день и вечер.

Всего в игре Dubious было представлено двадцать утверждений, на которые должны были ответить игроки, и 53 форм ПИА для тестирования. Помимо этого, участникам необходимо было принять ряд решений при создании своего персонажа — выбрать цвет волос, прическу, цвет кожи, фигуру, возраст, имя, пол, дату рождения и одежду для нижней части тела (все эти решения также необходимо было принять дважды — сначала для своего аватара, затем для аватара своего друга, врача, преподавателя или коллеги).

Все элементы обеих исследовательских игр, включая сюжет, уровни, миссии, характеристики, речевые интонации, лингвистические средства, семиотические конструкции, звуковое сопровождение (оригинальная музыка была приобретена на интернет-сайте свободного распространения материалов, изменена и отредактирована для лучшего соответствия созданным уровням и предложенным сюжетам) и даже шрифты, были созданы для поддержки утверждений, которые необходимо было включить в исследование, форм ПИА, которые требовалось протестировать, а также некоторых других задач исследования. Таким образом, тестируемые форматы ПИА и сбор данных в рамках исследования стали неотъемлемой частью геймплея. Например, некоторые повторяющиеся утверждения были представлены в виде части «личностного теста» или использовались для того, чтобы убедить определенных персонажей Dubious в том, что вы действительно прибыли из будущего.

Обе исследовательские игры были разработаны таким образом, чтобы вызывать определенные эмоции. Данный эффект был достигнут путем использования различных игровых средств и элементов; предполагалось, что игроки будут испытывать те же эмоции, как и в том случае, если бы события из исследовательских игр происходили с ними в реальной жизни. Например, в реальной жизни обнаружение потери документа, удостоверяющего личность (например, паспорта или свидетельства о рождении), в момент, когда этот документ необходимо предъявить, стало бы стрессовой ситуацией. Было крайне важно воспроизвести соответствующие ситуации и эмоции, чтобы узнать, как именно почувствовал бы себя игрок в стрессовой ситуации и какие действия бы предпринял. Это помогло Лизбет определить, какие модели поведения и соответствующие им реше-

ния выбирают люди, вынужденные выбирать одно из менее привычных им средств идентификации и аутентификации.

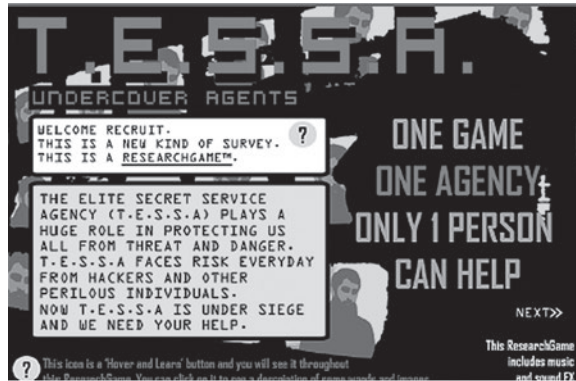


Рис. 2 Первая страница игры TESSA Undercover Agents с описанием сюжета.

4. Создание и разработка: анализ каждой исследовательской игры

4.1 T.E.S.S.A. Undercover Agents: краткое описание геймплея и сюжета

В игре TESSA игроки — секретные агенты на службе выдуманной организации T.E.S.S.A. (The Elite Secret Service Agency — Элитная секретная служба) (см. Рис. 2). Таким образом, с целью использования проективной методики игроки оказываются в контексте альтернативной личности. Был создан образ секретного агента и использован соответствующий стиль повествования. В кино, литературе, видеоиграх и СМИ секретные агенты часто фигурируют в ситуациях, так или иначе связанных с темой идентификации личности; данное исследование также посвящено идентификации и аутентификации. Поэтому образ героев и сюжет игры были раз-

работаны таким образом, чтобы для игроков содержание исследования имело какой-то смысл.

Чтобы поставить перед игроками игровую задачу, им сообщалось, что существованию организации TESSA в течение долгих лет угрожал вымышленный «Агент Ноль», враждебно настроенный хакер. Обязанностью игрока как секретного агента TESSA была поимка Агента Ноля. Эта сюжетная линия была введена с целью создания интриги, а также формирования основы для различных «миссий», которые игроки должны были выполнить в ходе исследовательской игры, при этом предоставляя данные, необходимые для проекта. Также она должна была сформировать у игроков внутреннюю мотивацию, поставив перед ними четкую цель, что является одной из базовых психологических потребностей человека, удовлетворяемых играми [51].

В первой миссии игроки должны были помешать Агенту Нолю покинуть страну, забрав у него средства идентификации личности, что сделало бы для него нормальное существование за границей невозможным. Благодаря использованию проективной методики эта миссия и развертывающийся нарратив помогли определить, какие средства идентификации и аутентификации (а также набор устройств) игроки считают неотъемлемой частью повседневной жизни. На последнем, четвертом уровне игры игроку, уже ставшему опытным и успешным секретным агентом, было поручено предоставить свои рекомендации TESSA относительно того, что ей следует предпринять в будущем, чтобы помешать хакерам взламывать ее системы безопасности. В эту часть вошли ответы на вопросы и выбор утверждений, соответствующих пожеланиям игроков, связанным со средствами идентификации и аутентификации в будущем. И вновь благодаря проективной методике мы смогли выявить те форматы идентификации и аутентификации, которым игроки доверяли, и которые, по их мнению, могли бы использоваться в будущем как достаточно безопасные. Это была одна из игровых задач, позволивших игрокам почувствовать значимость своего участия и сохранить дальнейший интерес к нарративу.

Целевой группой игр TESSA и Dubious стали мужчины и женщины в возрасте 18–65 лет и старше. Одним из способов эффективно

увлечь столь широкую аудиторию было использование знакомых, хорошо узнаваемых сюжетов, понятных проживающим в Великобритании представителям обоих полов и указанных возрастных категорий. Так, например, в игре TESSA можно найти множество сходств с известными сюжетами из фильмов о Джеймсе Бонде и Борне, а также «Шпион, выйди вон!», «На гребне волны» и «Серпико». Во всех перечисленных картинах обыгрывается тема сокрытия личности и работы секретных агентов.

4.1.1 Функция отображения подсказки при наведении курсора

В обеих исследовательских играх была реализована функция предоставления дополнительной справочной информации под названием «Отображение подсказки при наведении курсора» (Рис. 3). Она предназначена для снижения вероятности возникновения путаницы при использовании слов, имеющих двойной смысл. Так, например, было необходимо дать небольшое пояснение в отношении термина «имплантат» в ситуации, когда в качестве одного из средств идентификации и аутентификации тестировался имплантат радиочастотной идентификации. Функция отображения подсказки при наведении курсора также была использована для изображений, которые нуждались в комментариях (например, логотип Whatsapp⁷). Это было сделано для того, чтобы игроки максимально точно поняли наши задания и отреагировали на них соответствующим образом.

4.1.2 Вознаграждение, обусловленное нарративом

В TESSA был реализован принцип вознаграждения⁸, обусловленного нарративом, который усиливал эффект погружения участников

⁷ Whatsapp: <http://www.whatsapp.com/>

⁸ «Вознаграждение, обусловленное нарративом», — это предложенный мной термин для описания системы вознаграждения, признания достижений и стимулирования игроков, тесно связанной с игровым нарративом.

в игру и помогал им лучше справляться с заданиями на последующих уровнях. Так, например, код из цифр и букв, полученный игро-дентом на втором уровне, можно было использовать для открытия

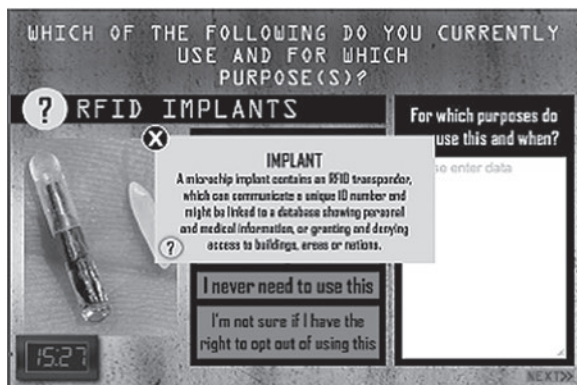


Рис. 3 Функция отображения подсказки при наведении курсора была использована в игре TESSA Undercover Agents для представления описания различных инструментов идентификации личности.

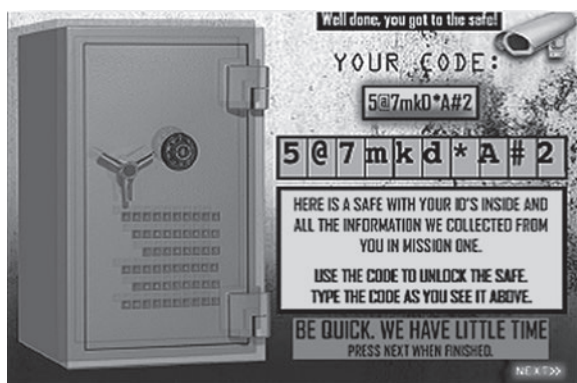


Рис. 4 Код, полученный в конце третьего уровня, оказывается полезным на четвертом уровне, когда с его помощью геймпондент может получить наиболее важные для себя средства идентификации личности.

сейфа на третьем уровне (Рис. 4). После завершения четвертого уровня также присуждалась награда за «особое усердие при выполнении своего служебного долга» в качестве секретного агента. Таким образом мы давали игроку понять, что его усилия будут цениться очень высоко даже после поимки Агента Ноль (Рис. 5).



Рис. 5 Медаль за особое усердие при выполнении своего служебного долга, изображение которой геймспонденты могут открыть в большом разрешении и затем сохранить на свой компьютер.

4.1.3 Уровни

Игра TESSA была поделена на четыре уровня или «миссии». Такая структура исследовательской игры помогала игрокам понять, на каком этапе они находятся, способствовала развитию сюжета (см. Рис. 6) и позволяла нам использовать каждый из уровней как средство достижения разных исследовательских целей. В ранее опубликованной статье, описывающей игру TESSA, Лизбет (как ведущий автор) представила в форме таблицы информацию о том, как уровни и сюжет игры помогали решать задачи исследования и собирать необходимые данные. Она также привела стандартные вопросы, которые могли бы стать их эквивалентом в том случае, если бы использовалась традиционная форма опроса респондентов (Таблица № 1).



Рис. 6 Список уровней исследовательской игры TESSA.

Таблица № 1 Цели исследования и эквивалентные вопросы
в традиционном формате онлайн-опроса,
соответствующие каждому из уровней игры TESSA
(Автор: профессор Лизбет ван Зуунен, 2013 г.)

Сюжет	Цели исследования	Эквивалентный вопрос
Введение. Необходимо про- браться в квартиру злодея, найти и уничтожить наиболее важные средства	Описать ситуацию и выяснить, какие три средства иденти- фикации личности респонденты считают неотъемлемой частью повседневной жизни	Выберите из представленного ниже списка средств иденти- фикации личности те, которые Вы считаете необходимыми для ведения нормальной жизни. Отметьте не более трех вариантов ответа
Злодей наносит ответный удар и пытается уничтожить сведе- ния о личности репондента. Он/она должен указать, какие из имеющихся у него/нее документов должны защитить агентство	Выяснить, какие средства идентификации и аутентифи- кации есть у людей	Что из следующего списка у Вас есть (Вы используете): – Паспорт, пинкод, пароль и т. д. – Паспорт, идентификацион- ная карта – Карта клиента – Биометрические средства идентификации – Средства радиочастотной идентификации

Сюжет	Цели исследования	Эквивалентный вопрос
Респонденту необходимо про- браться в офис TESSA и вы- яснить, были ли уничтожены сведения о его/ее личности. Поэтому он/она должен/ должна ответить на вопросы о своих ценностях, о которых знает только он сам/она сама и которые, предположительно, указаны в его/ее личном деле	Определить культурные цен- ности респондентов в соответ- ствии с моделью культурных ценностей Хофстеде [50]	Ниже представлен список раз- личных утверждений, не мог- ли бы Вы отметить степень своего согласия с каждым из них по следующей шкале: полностью согласен, согласен, не согласен, абсолютно не согласен. Например: <i>Люди, занимающие высокие посты, должны принимать большинство решений, не со- ветуясь с теми, кто занимает нижестоящие позиции</i>
Респонденту необходимо указать, как следует предот- вращать подобные ситуации в будущем, выбрав три более эффективных средства иденти- фикации личности	Получить представление о предпочтительных для обычных людей средствах идентификации личности	Ниже представлен список возможных средств иденти- фикации личности в будущем, не могли бы Вы выбрать три наиболее привлекательных для Вас варианта и расположить их по степени предпочтительности для Вас?

4.1.4 QR-код

QR-код в игре TESSA является частью вознаграждения за прохождение третьего уровня (Рис. 7). Игрок мог отсканировать этот QR-код с помощью своего смартфона. На экране ноутбука или настольного компьютера игрока QR-код был частью «Разрешения на выполнение миссии». При сканировании его телефоном можно было увидеть текст, сообщавший о новом задании. Подобный графический контент, нарратив и другие элементы, включая QR-код, были использованы для повышения увлекательности игры по сравнению с ее традиционным эквивалентом в форме онлайн-опроса. Мы хотели предоставить возможность получения дополнительной информации через QR-код по трем причинам: 1) чтобы помочь игрокам погрузиться в сюжет, позволив им при желании перенести что-то с экрана компьютера на свой телефон; 2) чтобы предложить игрокам более интересную форму вознаграждения по сравнению с теми, которые использовались ранее, таким образом внося в игру элемент

новизны; 3) чтобы избежать повторов при прохождении исследовательской игры. Мы хотели выяснить, сколько человек отсканируют QR-код, чтобы получить еще один показатель их вовлеченности, однако собрать эти данные оказалось невозможно.



Рис. 7 QR-код на Разрешении на выполнение миссии, выданном секретному агенту TESSA.

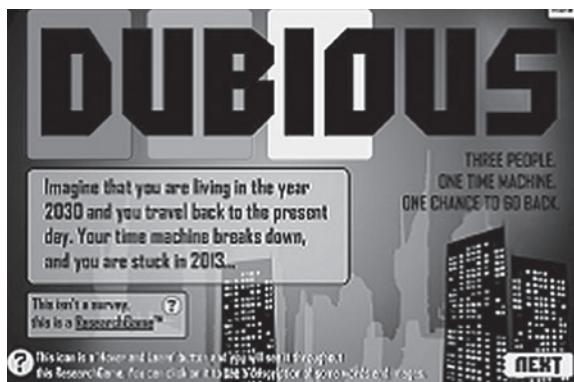


Рис. 8 Начальная заставка игры Dubious с описанием ее сюжета.

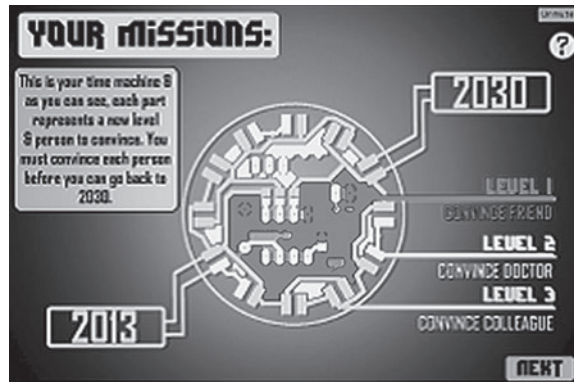


Рис. 9 Уровни игры Dubious с привязкой к сломанным деталям машины времени, которые необходимо починить.

4.2 Dubious: краткое описание геймплея и сюжета игры

Игроденту сообщают, что он живет в 2030 г.; отсюда он прибыл в прошлое — в 2013 г., после чего его машина времени сломалась (Рис. 8 и 9). Чтобы вернуться обратно, ему необходимо убедить трех знакомых в том, что он действительно прибыл из будущего. Это можно сделать, только одним способом — отвечая на вопросы этих людей о будущем.

После того, как игроденту удастся убедить всех троих собеседников, они помогут ему починить сломанную машину времени. До этого момента в историю игрока о его путешествии никто не верит, знакомые сомневаются, что он действительно прибыл из будущего.

В Dubious, как и в TESSA, нарратив должен был помочь решить сразу несколько задач: предоставить игродентам цель, которая стимулировала бы их внутреннюю мотивацию, создать общую структуру, в рамках которой можно было бы решить задачи исследования, а также глубже погрузить игроков в контент исследования, представленный в виде истории и различных игровых ситуаций.

Как и в TESSA, здесь был использован увлекательный и хорошо знакомый многим сюжет, соответствующий целям исследования.

Его задача состояла в том, чтобы обеспечить игродентам яркий игровой опыт, позволяющий представить такие малопонятные футуристические концепции, как «гаджеты будущего» и «будущие технологии идентификации личности». Была необходима история, достаточно узнаваемая и простая для понимания в целевой группе жителей Великобритании в возрасте 18–65 лет и старше. Фильмы и сериалы «Назад в будущее», франшиза «Доктор Кто» и «Машина времени» непосредственно связаны с темой путешествий во времени и представлениями о мире будущего. При этом все эти сюжеты хорошо известны данной аудитории. Поэтому для достижения поставленных перед исследованием задач была использована история, основанная именно на этих произведениях. Мы стремились понять, как люди представляют себе использование различных средств идентификации и аутентификации в будущем и как эти средства изменятся к 2030 г.

4.2.1 Персонализация и детали повествования

Три персонажа, которых игроденту необходимо убедить в том, что он прибыл из будущего,— это его коллега или преподаватель (в зависимости от того, кто из них присутствует в реальной жизни игрока), врач и друг. Участнику предоставляется возможность вписать реальные имена своего врача, друга и коллеги / преподавателя (Рис. 10). Персонализация исследовательской игры с помощью имен реальных людей из жизни игродентов (см. Рис. 10, 11 и 12) была использована для того, чтобы побудить их отвечать на вопросы честно. Было выдвинуто предположение о том, что если предоставить игроку возможность отвечать на вопросы о будущем так, будто они общаются с реальными знакомыми, то они ответят так же, как если бы говорили с этим человеком в реальности. Предполагалось, что такая персонализация сюжета благодаря ощущению тождественности персонажей игры с реальными людьми побудит игродентов давать более честные и прямые ответы, а также повысит уровень вовлеченности участников исследования. Коммерческие организации используют данный прием для повышения вовлеченности потребителей, особенно детей [52].

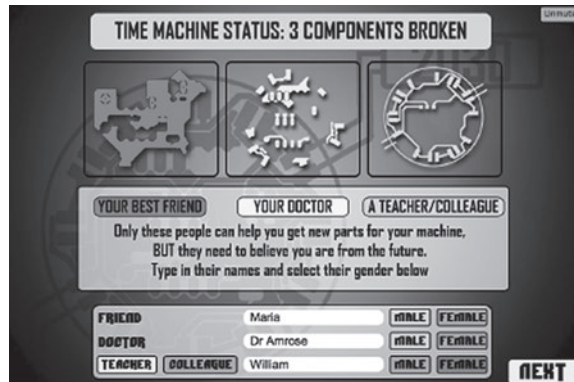


Рис. 10 Геймспондент вводит имена персонажей и выбирает их пол.



Рис. 11 Имя реального друга используется в описании задания первого уровня игры.

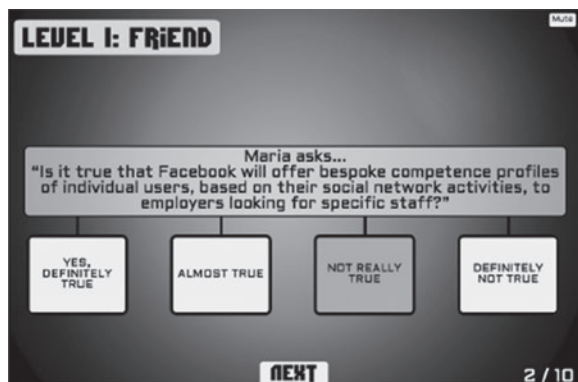


Рис. 12 Персонализация прослеживается и в вопросах, на которые геймспондент должен ответить.

4.2.2 Дальнейшее использование трех NPC (неигровых персонажей)

Нарратив было сформирован, в первую очередь, для включения в исследование разработанных Лизбет вопросов по тематике TSGUC. В каждой из групп было представлено по пять утверждений или вопросов. Категоризация TSGUC была проведена таким образом, что еще до разработки самой исследовательской игры были созданы три самостоятельные группы и три персонажа, каждый из которых в игре «отвечал» за одну из групп. Таким образом, персонажи игры «произносили» утверждения и вопросы в соответствии со своей областью интересов или знаний. Так, например, врач задавал вопросы о том, как в будущем будут хранить медицинские сведения о пациентах и предоставлять к ним доступ; друг задавал вопросы о том, какие сайты будут популярны в 2030 г., как изменится закон о хранении личных данных в Сети и о том, как к 2030 г. изменятся современные социальные сети.

4.2.3 Представление о будущем

Чтобы помочь игрокам представить себе 2030 г., в начале игры Dubious им было задано пять вопросов, связанных с «описанием ме-

ста действия». На экране появлялась надпись «Представьте себе 2030 год» в сопровождении успокаивающей музыки, которая должна была помочь игроку погрузиться в размышления о будущем. Затем следовали вопросы о том, сколько лет исполнится участнику исследования в 2030 г., как он будет жить в это время, как он представляет себе общество будущего и почему. Это было сделано для того, чтобы сфокусировать мысли игродента на определенном временном периоде и собрать информацию о том, как он «видит будущее».

4.2.4 Аватары

Чтобы собрать сведения о том, как игродент представляет свою одежду, а также одежду других людей в будущем, в игре Dubious была дважды реализована функция создания аватара: во-первых, при создании игроком собственного аватара, во-вторых, при создании аватара одного из трех упомянутых ранее реально существующих персонажей.

В распоряжении компании Research Through Gaming Ltd уже было необходимое для создания аватаров программное обеспечение под названием АСТ (Инструмент создания аватара), однако для игры Dubious его необходимо было модифицировать в соответствии с установленными целями исследования. Например, в качестве дополнительных вариантов предлагаемых аксессуаров и одежды мы добавили ассортимент существующих и футуристических технологий и устройств. Это было сделано для того, чтобы узнать, что игроденты выберут для себя в 2030 г. По сюжету игроку необходимо было «одеться» перед встречей с тремя знакомыми. Он получал следующие инструкции: *«Теперь Вам необходимо одеться для встречи с (имя друга), (имя врача) и (имя коллеги или преподавателя)! Создайте свой аватар здесь. Помните, что вы из 2030 г., поэтому должны одеться так, как будто это произошло в будущем»*.⁹

В одном из модифицированных нами компонентов АСТ мы убрали все варианты одежды, предназначенной для верхней части тела

⁹ Имена, указанные в скобках, игроденты вводят в начале исследовательской игры Dubious для своего врача, друга и коллеги / преподавателя.

аватара, и заменили их тремя футуристическими предметами одежды с длинным рукавом. Они выглядели совершенно одинаково, и мы были уверены, что игрок выбрал тот или иной вариант на основании его функциональности, а не внешнего вида. Эти данные были важны для Лизбет, чтобы разграничить понятия функциональности и внешнего вида для средств идентификации и аутентификации. О функциональности этих предметов одежды можно было узнать, получив подсказку при наведении курсора (Рис. 13). У этого программного обеспечения была еще одна функция — снизить повторяемость элементов этой части исследования. По сути, вместо того чтобы спросить человека, как его зовут, сколько ему лет, какого он пола, какую одежду он бы носил в будущем и т.д., как это было бы сделано в традиционном опросе, в нашей исследовательской игре надо было выполнить визуальное интерактивное задание. В результате проделанной работы игроки получали собственный графический образ.

Использование инструмента создания аватаров помогло нам избежать повторения в случае, когда требовалось выяснить, как



Рис. 13

Инструмент создания аватара был изменен и использован в исследовательской игре Dubious вместе с функцией отображения подсказки при наведении курсора для описания свойств предложенных предметов одежды и носимых устройств.

игроденты представляют себе одежду *других* людей в будущем. Мы использовали его для реализации проективной методики, попросив игрока одеть другого человека так, чтобы он был похож на обычного горожанина из 2030 г. Это позволило нам понять, какие варианты футуристических и существующих средств идентификации будут, по предположениям игродентов, применяться людьми будущего. Мы получили множество подтверждений того, что возможность самостоятельно создать свой аватар способна повысить уровень вовлеченности пользователя в игру или другую платформу, созданную с помощью программного обеспечения [53]. Поэтому у инструмента создания аватара была и третья функция — увлечь игрока и поддерживать это увлечение до самого конца игры.

После прохождения каждого из уровней появлялось сообщение о том, что игроку удалось убедить еще одного персонажа. В качестве вознаграждения за это он получал не только работающую деталь для своей машины времени (Рис. 9), но также и «сведения из будущего». Последние представляли собой краткое изложение фактов о будущем, основанных на ответах игродента на том уровне, который он только что прошел. Это было сделано для повышения вовлеченности участников, а также для предоставления им информации, которая могла показаться интересной и улучшила бы впечатление от участия в исследовании в целом.

4.2.5 Мини-дневник

В Dubious был заключительный, необязательный для прохождения уровень игры — раздел с мини-дневником (Рис. 14). Он был необходим для понимания того, как игроденты представляют себе один день обычной жизни в 2030 г. К моменту перехода на этот уровень игроки должны были вжиться в свою роль «человека из будущего».

В случае традиционных онлайн-опросов использование анкет в форме дневников зачастую отпугивает участников исследования, не позволяя им дойти до конца исследовательской процедуры. Поэтому в Dubious этот уровень был необязательным для прохождения. Однако в итоге мы отметили на удивление высокий уровень позитивных отзывов о дневнике.

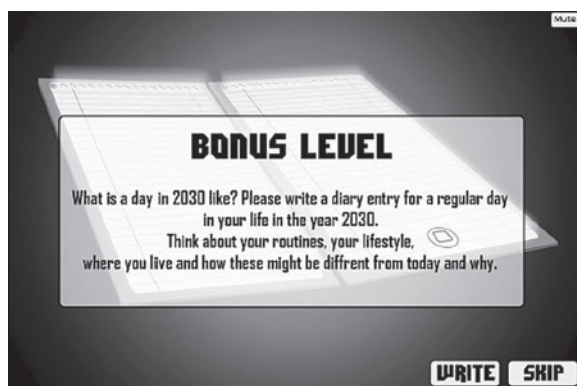


Рис. 14 После появления на экране надписи: «МИССИЯ ВЫПОЛНЕНА» геймспондент получает неожиданное предложение выполнить еще одно задание по заполнению дневника.

5. Методология

Лизбет было необходимо, чтобы в обеих исследовательских играх приняли участие одни и те же игроденты. Результаты, полученные в ходе обеих игр, включали полное демографическое описание участников, ответы на вопросы в TESSA и ответы на вопросы в Dubious. Перед участием в исследовательских играх игроденты заполняли так называемую «скрининговую анкету», что занимало у них около десяти минут. Эта анкета была представлена в традиционном для онлайн-опросов формате (белый фон, ячейки для отметок возле предложенных вариантов ответа) и включала в себя ряд демографических вопросов и вопросов об образе жизни респондентов, позволяющих провести профилирование и убедиться, что участники попадают в целевую группу по возрасту. Между выходом TESSA и Dubious прошло четыре месяца. За этот период со всеми участниками связывались дважды, высылая им приглашения к участию в каждой из исследовательских игр.

Этот процесс практически идентичен тому, что используется при проведении традиционных онлайн-опросов. Игроденты получили приглашение по электронной почте с URL-ссылкой на страни-

цу игры. Обычно респонденты таким же образом получают ссылки на традиционные онлайн-опросы в тексте электронного письма. После нажатия кнопки мыши открывалось новое окно с исследовательской игрой вместо обычной опросной анкеты.

Игра TESSA (N=1400) была запущена в марте 2013 г., Dubious — в июле 2013 г. Рассылка приглашений в Dubious была сделана всем игрокам, прошедшим до конца TESSA, что в итоге позволило получить N=902.

В обе игры нельзя было сыграть в браузере, не поддерживающем стандарт HTML5 (например, в Internet Explorer 8). Всем участникам, чей браузер не соответствовал необходимым требованиям к программным средствам, демонстрировалось соответствующее сообщение. В нем игроку было предложено пять URL-ссылок для загрузки последних версий самых популярных браузеров для настольного ПК, а именно: Opera, Google Chrome, Internet Explorer, Firefox и Safari.

5.1 Продолжительность исследования

Для прохождения Dubious игроку требовалось около 25 минут. Игра включала в себя 88 вопросов и утверждений, бонусный минигуровень с заполнением дневника и два задания с использованием Инструмента создания аватара. Прохождение TESSA занимало максимум 30 минут, включая заполнение скрининговой анкеты¹⁰. Всего в TESSA было 78 утверждений / вопросов.

Итак, каждый игрок должен был принять участие в одном скрининговом опросе и двух играх — TESSA и Dubious — с перерывом в четыре месяца, что должно было занять у него в целом один час времени.

В среднем игроки потратили на прохождение TESSA 21 минуту. Как было отмечено в работе 2013 г. под авторством Ван Зуунен и др., на основании выводов, сделанных Кроуфорд и др., это можно считать хорошей продолжительностью. Однако, по рекомендациям Фэна

¹⁰ Максимальное время рассчитывалось как время, которое потребуется 95% игрокам на прохождение игры от начала до конца.

и Янга, для удержания внимания респондентов максимальная длительность онлайн-опроса не должна превышать 13 минут.

5.2 Средства поощрения

Компания, предоставившая свои панели для рекрута игроков для TESSA и Dubious, использует систему баллов, которые можно обменивать на различные товары и услуги. За участие в TESSA и Dubious игроки получили 300 баллов, так как длительность обоих исследований составила около 30 минут. Для сравнения: средняя длительность онлайн-опросов, в которых обычно принимают участие панелисты этой компании, как нам сообщили, составляет от 10 до 15 минут. За участие в них респонденты получают 100–150 баллов.

6. Использованные методы измерения

6.1 Измерение вероятности повторного участия в подобной исследовательской игре

Чтобы понять, оказался ли формат исследовательской игры увлекательным для игроков, мы измеряли три показателя, определяющие вероятность их участия в подобных исследованиях в будущем, так как пришли к выводу, что именно вероятность будущего участия может служить индикатором вовлеченности. Мы использовали следующие показатели:

1. Доля игроков (в %), изъявивших желание принять участие в последующих исследовательских играх при предоставлении обратной связи по завершении игры (среди тех, кто предоставил обратную связь, 29 игроков изъявили желание принять участие в последующих исследовательских играх).
2. Доля игроков (в %), ответивших «Да» на вопрос «Хотели бы Вы снова сыграть в подобную игру?»
3. Доля игроков (в %), реально *принявших* участие в следующей исследовательской игре (этот показатель рассчитывался

на основании параметра возобновления участия, учитывавшего количество испытуемых, сыгравших в Dubious через четыре месяца после завершения TESSA).

С помощью третьего показателя мы смогли также оценить честность ответов на вопрос второго показателя.

6.2 Измерение вовлеченности

Для измерения вовлеченности в обеих исследовательских играх были использованы четыре подхода.

1. Вероятность принятия участия в подобных исследовательских играх в будущем по личной оценке игроков, а также доля (в %) тех, кто сыграл в Dubious после завершения TESSA. Согласно полученным результатам, показатель возобновления участия составил 81%.
2. Отношение игроков к играм, высказанное ими в процессе предоставления обратной связи в разделе комментариев после прохождения каждой из исследовательских игр. Согласно полученным результатам, более 90% всех отзывов оказались положительными.
3. Сведения, предоставленные в рамках обратной связи по игре Dubious, которые указывали на тот факт, что игроки помнили первую исследовательскую игру TESSA. По нашему мнению, подобные воспоминания также являются индикатором вовлеченности участников.
4. Процентная доля игроков, прошедших необязательный «Бонусный уровень» с заполнением дневника в Dubious, так как, по нашим предположениям, выполнение необязательных заданий также является признаком вовлеченности участников. Доля игроков, заполнивших обе части дневника, — утро / день и вечер — составила 19%.

Высказывания респондентов в рамках обратной связи, предоставленной после прохождения игр TESSA и Dubious, были поделены на категории, указанные ниже. Эти категории соответствуют общим темам, выявленным в процессе анализа полученной обратной связи.

1. Выражение интереса к участию в будущих исследовательских играх.
2. Музыка и звуковое сопровождение.
3. Отрицательное отношение.
4. Ошибки и сбои.
5. Выражение опасений и неуверенности.
6. Неопределенное отношение.
7. Положительное отношение.

7. Показатели вовлеченности участников (приведенные здесь результаты разбиты на следующие категории: результаты по игре TESSA, результаты по игре Dubious, общие результаты для двух игр)

7.1 Общие результаты по игре TESSA

В среднем на прохождение исследовательской игры TESSA игро-дентам потребовалась 21 минута. 27% испытуемых, принявших участие в TESSA, добровольно предоставили обратную связь (N=376). 54% из тех, кто предоставил обратную связь,— женщины.

В репликах, предоставленных после просьбы оставить комментарии или пожелания по игре TESSA, наиболее часто встречались прилагательные и фразы, относящиеся к положительным впечатлениям, что указывает на вовлеченность участников и полученное ими удовольствие. Слова, наиболее часто употреблявшиеся игродентами (по порядку от наиболее до наименее часто встречающихся):

1. Веселая.
2. Интересная.
3. Непохожая на другие.
4. Прошел/прошла с удовольствием.
5. Отличная.
6. Хорошая.
7. Приятная.
8. Необычная.
9. Нестандартная.

7.1.1 Выражение сомнений и неуверенности в рамках обратной связи

18 респондентов из числа тех, кто предоставил обратную связь, выразили сомнения и неуверенность в отношении двух аспектов игры TESSA.

- 1) Формат исследовательской игры. В комментариях, оставленных игроками, встречаются опасения, связанные с тем, что испытуемые никогда раньше не видели таких исследовательских игр. Проведенный анализ позволил заключить, что эти респонденты впервые столкнулись с исследованием в игровом формате.
- 2) Содержание исследования. На втором уровне игрокам был предложен ряд выпадающих меню, при этом их попросили заглянуть в свои бумажники или сумки и перечислить все средства идентификации личности, которые там находятся (Рис. 15).

Участников спросили, сколько у них банковских карт, карт постоянного клиента и т.д. Затем их попросили сообщить, как и где они обычно хранят свои пароли. По сюжету, игроков попроси-



Рис. 15 В TESSA геймспондентов попросили заглянуть в свой бумажник или сумку, чтобы указать имеющиеся в них средства идентификации личности.

ли предоставить все эти сведения для их «защиты силами агентства TESSA», потому что их персонажу как секретному агенту в этот момент угрожала опасность раскрытия. При этом мы учитывали тот факт, что предоставление информации о своих банковских картах и средствах идентификации личности считается небезопасным действием, которое не рекомендуется производить в сети даже в рамках онлайн-исследования. Полученные результаты показали, что 21% игроков, начавших игру TESSA, покинули ее между 11 и 12 страницами, как раз когда их спросили об используемых средствах идентификации и их количестве (Рис. 15). На Рис. 16 можно увидеть резкий скачок, отражающий долю игроков, покинувших исследовательскую игру в момент появления этих вопросов. Конфиденциальный характер запрашиваемых сведений в сочетании с сюжетом о секретном агенте / похищении личных данных вполне мог вызвать подозрения и сомнения, отмеченные несколькими участниками.



Рис. 16 График, на котором представлены показатели доли участников, покинувших исследовательскую игру TESSA до ее окончания, с заметным скачком в момент появления на экране личных вопросов
(Источник: Research Through Gaming Ltd., 2013 г.)

Игроденты не пожелали предоставить конфиденциальную информацию, что вполне понятно, учитывая ее персональный характер. Это нежелание / подозрительность получили подтверждение в комментариях участников, оставленных ими при ответе на вопросы в конце каждой исследовательской игры.

Один из игроков отметил: *«Это исследование сильно отличалось от всех, в которых я когда-либо принимала участие, но я так и не продвинулась дальше первых вопросов, потому что они вызвали подозрение, что меня пытаются обмануть мошенники»*. Женщина, 52 года.

Некоторые респонденты сообщили о своих подозрениях и сомнениях при предоставлении обратной связи, однако это не мешало им отметить свой интерес к участию в последующих исследовательских играх. Почти все игроки ответили «Да» на вопрос *«Хотели бы Вы снова сыграть в подобную игру?»* даже после того, как высказали в комментариях свои опасения.

7.1.2 Музыка / звуковое сопровождение

Мнения участников игры TESSA о представленной в ней музыке и звуковом сопровождении разделились. Четыре игрока прокомментировали звуковую дорожку словами *«превосходно»* или *«замечательно»*, в то время как другие тринадцать продемонстрировали ее резкое неприятие, назвав музыку *«мрачной»* или *«угнетающей»*. (В действительности, мы хотели создать у игроков ощущение напряжения, так как эта эмоция соответствовала сценарию, использованному нами для достижения поставленных целей исследования). Следует также отметить, что в разные моменты исследовательской игры TESSA воспроизводились разные музыкальные фрагменты, а именно: 1) музыка вступления, 2) разная фоновая музыка для первого, второго, третьего и четвертого уровней, 3) триумфальная музыка (при успешном прохождении уровня), 4) фоновая музыка при отображении таблицы прохождения игры (с представлением всех уровней и полученных на них достижений — см. Рис. 6), 5) музыкальное оформление окончания игры.

В TESSA не было кнопки выключения звука, поэтому некоторые из отрицательных отзывов о музыкальном сопровождении могли быть вызваны тем фактом, что игроки не могли его отключить. Кнопка выключения звука была добавлена в интерфейс второй исследовательской игры — Dubious.

7.1.3 Сообщения о технических проблемах

Четыре игрока сообщили о небольших технических проблемах, возникших в игре TESSA, например, о необходимости обновить страницу перед началом игры. Кроме того, на некоторых мониторах участники не смогли увидеть все изображения целиком. Высказывая эти критические замечания, они предоставили положительную обратную связь по исследованию в целом. Так, например, одна из игроков, прошедшая TESSA до конца, потратила время на то, чтобы выяснить мое имя и адрес электронной почты; 29 марта 2013 г. она прислала мне письмо, текст которого разрешила опубликовать.

«Дорогая Бетти, я только что завершила участие в опросе и подумала, что следует Вам сказать о том, что этот опрос получился просто изумительным. Он — один из лучших, в которых я когда-либо принимала участие, а я довольно часто и с раннего возраста принимаю участие в различных исследованиях, включая личные и онлайн-интервью. Думаю, Вам также следует узнать о том, что многие страницы вашего опроса не загрузились у меня полностью — у нас обычный широкополосный доступ в Интернет без кабеля — на что, наверное, довольно часто жалуются и другие люди. Поэтому я не смогла заполнить страницу с опросной анкетой. И, наконец, следующую страницу мне пришлось перезагружать дважды, прежде чем я смогла нажать «Да». Наверное, следует принять во внимание тот факт, что у многих пользователей недостаточно высокая скорость подключения в районе проживания?»

Как бы то ни было, спасибо за самый нескучный опрос в моей жизни!

XXXX

P. S. На прохождение у меня ушло больше часа времени...»

Во втором электронном письме, написанном в ответ на мое сообщение, участница написала: «Надеюсь, благодаря столь све-

му взгляду на вещи Вам удастся добиться еще больших успехов». Вполне логично заключить, что усилия по поиску моего электронного адреса и написанию письма в сочетании с положительной обратной связью и конструктивной критикой могут служить признаком вовлеченности игродента в исследовательскую игру, даже несмотря на возникшие технические проблемы и большую продолжительность опроса.

7.1.4 Неопределенное отношение / смысл слов неясен

Некоторые игроденты оставили сообщения, по которым крайне сложно понять, какое отношение к опросу они хотели выразить. Так, например, женщина 43 лет написала об игре TESSA, что это *«очень необычный метод исследования»*. Без дополнительных сведений невозможно понять, служит ли эта фраза показателем положительной или отрицательной оценки или же является поверхностным комментарием без каких-либо эмоций.

7.1.5 Крайне отрицательное отношение к игре TESSA или предпочтение традиционных опросов

Из 1400 игродентов, сыгравших в TESSA, 24 высказались об этой исследовательской игре отрицательно.

Восемь из этих 24 человек высказали пожелание или предпочтение принять участие в традиционном опросе. Например, одна из респонденток отметила: *«Честно говоря, мне больше нравятся обычные опросы»* (женщина 60 лет). Другой высказал просьбу: *«Просто задавайте вопросы!»* (мужчина 36 лет), третья недоумевает: *«Почему вы не можете просто задавать людям вопросы?»* (женщина 55 лет). Двое из этих восьми (оба старше 60 лет) при описании формата исследовательской игры использовали слово *«детский»*.

Некоторые игроденты объяснили, что причиной отрицательного отношения к исследовательской игре стала, например, сама *«структура / стиль игры»* (женщина 21 года) или же *«очень слабая идея»* (женщина 55 лет). Три из восьми игродентов, отдавших предпочтение традиционным форматам опросов, ответили «Нет» на вопрос

«Хотели бы Вы снова сыграть в подобную игру?». Четыре участника исследования ответили на этот вопрос «Да», а один игрок оставил этот вопрос без ответа. Несмотря на все замечания, высказанные в качестве обратной связи, 50% участников, которым не понравился формат исследовательской игры (т.е. четыре из восьми игроков), впоследствии сыграли во вторую игру — Dubious.

7.1.6 Положительные отзывы

91% от общего числа игроков в TESSA оставили о ней положительные отзывы. Частое использование слова «опыт» при предоставлении обратной связи позволяет предположить, что игроки настолько погрузились в геймплей / сюжет игры, что начали воспринимать ее как реальное переживание. Двадцать девять игроков спросили, есть ли возможность сыграть в другие подобные игры, или попросили выслать им *«больше подобных приглашений на игру, пожалуйста»*. 40-летний мужчина оставил следующий комментарий: *«Это, вероятно, одно из лучших и наиболее изобретательных исследований, в которых я когда-либо принимал участие, хотелось бы, чтобы таких было больше. В нем интересно и приятно принимать участие. Я бы поучаствовал в подобном исследовании снова»*.

Женщина 67 лет так прокомментировала игру: *«Это был очень приятный опыт. Было бы здорово, если бы таких исследований было больше!»*

7.1.7 Репрезентативная выборка по стране

Одним из факторов, вызывавших беспокойство при подготовке к данному исследованию, стало предположение о том, что в целевой аудитории окажутся подгруппы, представителей которых будет крайне сложно привлечь к участию. Например, это жители Великобритании в возрасте старше 60 лет и молодые мужчины. Самому старшему из игроков, сыгравших в TESSA, было 89 лет, а полученные в итоге результаты показали равное распределение между мужчинами и женщинами во всех возрастных группах.

Участники исследования, представляющие возрастную группу старше 60 лет, также дали обратную связь. Мужчина 72 лет оставил следующий отзыв: *«Мне очень понравилось играть в эту игру. Я не геймер и не пользуюсь ни смартфоном, ни мобильными приложениями, но этот игровой опыт помог мне узнать нечто новое»*.

В отчете, подготовленном Лизбет и ее коллегой, доктором Джорджиной Тернер, было отмечено: «Мы не выявили никакой взаимосвязи...».

7.1.7 Анализ результатов игры TESSA

Исследовательская игра TESSA определенно понравилась большинству игроков. С помощью обратной связи — положительных комментариев и мнений — они предоставили нам достаточно оснований полагать, что этот формат оказался настолько удачным, что способен повысить вероятность повторного участия пользователей в подобных исследованиях с четко выраженной мотивацией. Проанализировав впечатления игроков от участия в исследовательской игре, мы сделали вывод о том, что оценку восприятия можно повысить, добавив кнопку выключения звука, что и было реализовано в следующей игре — Dubious.

Как разработчику мне хотелось, чтобы в определенные моменты игроки оказались в стрессовой ситуации, что позволило бы повысить реалистичность игровой истории. Однако в те моменты, когда напряжение *не было* так необходимо, но были необходимы быстрые ответы без долгих раздумий, можно было выбрать музыку в быстром темпе с более позитивными модуляциями и без нарастания интенсивности звучания. В те моменты, когда темп музыки замедляется, давая возможность респонденту подумать, вероятно, следовало сделать ее менее напряженной по интонации.

7.2 Общие результаты по игре Dubious

В среднем на прохождение исследовательской игры Dubious у игроков ушло 28 минут. Дополнительные 7 минут можно объяснить тем фактом, что участников эксперимента попросили создать два аватара.

115 игроков, сыграв в Dubious, предоставили обратную связь по игре (12,7%). Прилагательные и фразы, наиболее часто употреблявшиеся игроками в комментариях, представлены ниже в порядке от наиболее до наименее часто встречающихся.

1. Веселая.
2. Хорошая.
3. Интересная и отличная — с одинаковой частотой.
4. Приятная.
5. Непохожая на другие и прошел/прошла с удовольствием — с одинаковой частотой.
6. Замечательная.

19% игроков, сыгравших в Dubious, также выполнили предложенное нами бонусное задание «Дневник: один день из жизни в 2030 году», заполнив обе части — «утро» и «вечер». Это задание было не обязательным. Кнопки «НАПИСАТЬ» и «ПРОПУСТИТЬ» были четко видны на экране (Рис. 14). 21% участников заполнили только одну часть — «утро». 19% участников заполнили обе части дневника.

7.2.1 Отсутствие сомнений и неуверенности

Ни один из игроков не упомянул наличие каких-либо сомнений или неуверенности в правильности своих действий в рамках игры Dubious. Вероятно, это можно объяснить тем фактом, что все опрошенные ранее прошли игру TESSA и были знакомы с исследованиями «в игровом стиле». Кроме того, снижение уровня подозрительности и неуверенности также могло быть вызвано изменениями в содержании игры. В TESSA были представлены вопросы, связанные с конфиденциальной информацией, тогда как в Dubious таких вопросов не было.

7.2.2 Показатель доли участников исследования, прошедших игру частично

В случае Dubious мы наблюдаем два скачка на графике выходов игроков из игры, отмеченные на Рис. 17. Первый (линия красного цвета) связан с вопросом, предвещающим работу с инструмен-

том создания аватара. На этой странице мы попросили участников ввести такие данные, как имя своего друга, врача и преподавателя или коллеги, а также их пол. На этой же странице они должны были выбрать, кто именно будет фигурировать в игре — их преподаватель или коллега. Эта часть показалась непонятной некоторым игроде-там, принимавшим участие в исследовании в первые несколько дней после его запуска. Они не знали, какую кнопку следует нажать в дан-ном случае — с надписью «Преподаватель» или «Коллега». Это было ошибкой дизайнера данной страницы. Не сделав выбор между пре-подавателем и коллегой, геймспонденты не могли двигаться по игре дальше. Поэтому был подготовлен «патч» (небольшое изменение программного обеспечения), прояснивший смысл данного выбора и объяснивший участникам, каких действий от них ожидают. Про-анализировав данные, полученные до и после выхода патча, мы уви-дели, что вышеупомянутый скачок был определенно вызван именно этим недопониманием, и изменение интерфейса на этой странице

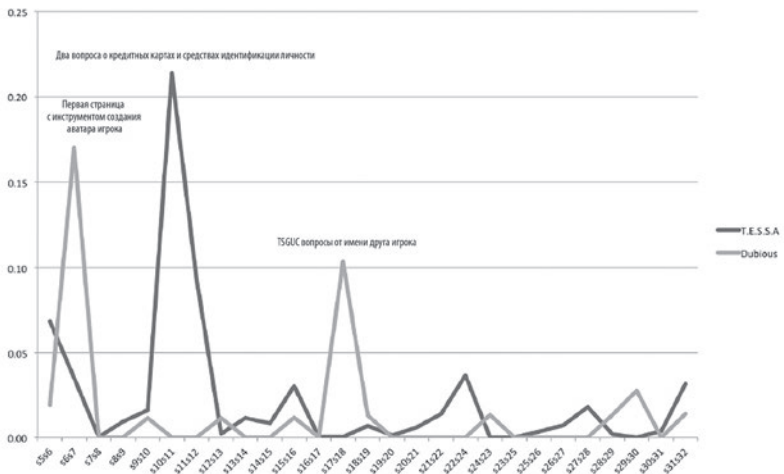


Рис. 17 График, на котором представлены показатели доли участников, покинувших исследовательскую игру до ее окончания, с разбивкой по играм – TESSA и Dubious (Источник: Research Through Gaming Ltd., 2013 г.)

привело к снижению показателя доли геймспондентов, покинувших исследование на этом этапе (скриншот страницы с этим вопросом представлен на Рис. 8).

Второй скачок на графике, по нашему мнению, является результатом применения методики Research Through Gaming для отслеживания достижений игроков. Он отражает почти десятикратное превышение стандартного показателя доли покинувших исследование участников на странице с вопросами TSGUC (о последних было упомянуто выше) от имени друга игрока. На этой странице было представлено десять вопросов, поэтому мы не считаем этот факт признаком затруднений, вызванных непосредственно исследовательской игрой.

7.2.3 Музыка / звуковое сопровождение

При создании Dubious мы пришли к мнению о том, что медленная музыка в стиле эмбиент будет наиболее эффективным средством помощи игрокам представить себе жизнь в 2030 г. В тех ситуациях, когда это было необходимо, мы использовали более отрывистую, ускоренную мелодию с «позитивными» интонациями. Не забывая об обратной связи, полученной по итогам проведения игры TESSA в отношении музыкального сопровождения, в Dubious мы также добавили кнопку выключения звука. Только один игрок прокомментировал музыкальное оформление игры при предоставлении обратной связи. Мужчина 40 лет оставил комментарий: *«Спасибо за кнопку выключения звука»*. Мы полагаем, что отсутствие обратной связи в отношении звукового сопровождения в данном случае объясняется тем фактом, что игроки, игравшие в Dubious, ранее уже приняли участие в TESSA и оставили соответствующие комментарии в рамках обратной связи по первой игре.

7.2.4 Другое

Два игрока при предоставлении обратной связи также указали две орфографические ошибки, замеченные ими в данной исследовательской игре.

7.2.5 Положительные отзывы

Как и в случае с TESSA, большинство игроков отметили, что получили удовольствие от Dubious, и выразили желание принять участие в других подобных исследованиях. Это, однако, говорит о наличии определенной предвзятости, так как тот факт, что игроки, сыгравшие в TESSA, сыграли и в Dubious, подтверждает их положительное отношение к исследованиям в подобном формате. Поэтому неудивительно, что отзывы этих респондентов о Dubious также оказались положительными.

7.2.6 Анализ результатов игры Dubious

Исследовательская игра Dubious существенно отличалась от TESSA. Музыка, звуковые эффекты, использование аватаров, кнопка выключения звука, нарратив и т.д. были направлены на решение существенно отличающихся исследовательских задач. Благодаря советам и конструктивной критике респондентов, прошедших TESSA (например, добавление кнопки выключения звука), разработчики получили возможность улучшить Dubious, в результате игра доставила участникам исследования больше удовольствия. Мы сделали такой вывод еще и потому, что многие респонденты оставили свою обратную связь только после первой исследовательской игры TESSA. Вероятно, они не сочли необходимым высказать свое мнение второй раз, после прохождения Dubious.

8. Общая вероятность принятия участия в подобных исследовательских играх в будущем

Используя три выбранных метода оценки вероятности повторного участия игроков в исследовательских играх, мы выяснили следующее:

- 1) Из 491 сообщения, полученного в рамках обратной связи после прохождения обеих игр — TESSA и Dubious, — в 29 было выражено желание сыграть снова. Следовательно, таково было содержание 6% откликов.

- 2) После прохождения TESSA 91% игроков ответили «Да» на вопрос: *«Хотели бы Вы снова сыграть в подобную игру?»* К сожалению, мы не получили аналогичные данные по Dubious; вопрос в конце игры был задан, однако из-за технического сбоя ответы не сохранились. Впрочем, у нас был и третий способ оценки вероятности повторного участия.
- 3) 81% игроков, прошедших TESSA, четыре месяца спустя также сыграли в Dubious. Это показатель фактического продолжения участия в проекте¹¹. Следовательно, из 91% респондентов, высказавших желание снова принять участие в подобном исследовании, 81% поступили соответствующим образом. Этот показатель доли респондентов, продолживших участие в проекте, представляется весьма внушительным. Он оказался на 11% выше по сравнению с прогнозом компании, предоставившей нам свои панели респондентов. Основываясь на показателях, достигнутых при использовании традиционных форм онлайн-опросов, проводившихся ранее, её специалисты предположили, что максимальный возможный уровень повторного участия составит 70%.

9. Общие показатели вовлеченности участников исследования

При проведении обеих исследовательских игр — TESSA и Dubious — были использованы четыре подхода к измерению вовлеченности.

- 1) **Отношение участников исследования к TESSA и Dubious вместе и по отдельности, полученное посредством обратной связи.** Отношение к игре TESSA оказалось преимущественно положительным — из 376 игроков, предоставивших обратную связь, 344 человека — или 91% — высказались о ней соответствующим образом.

Из 902 игроков, прошедших Dubious (а также TESSA за четыре месяца до этого), 115 человек (13%) предоставили по ней

¹¹ Показатель фактического продолжения участия в проекте определяется по числу участников, успешно завершивших опрос А и перешедших к опросу В.

обратную связь. 96 из 115 полученных сообщений оказались положительными. Таким образом, 83% от всех сообщений обратной связи, полученных по исследовательской игре Dubious, оказались положительными.

Из 2302 игроков, прошедших обе исследовательские игры (TESSA N=1400 + Dubious N=902), обратную связь по ним предоставили 491 человек (21% для обеих исследовательских игр).

Из 491 полученного нами сообщения 440 оказались полностью положительными.

Таким образом, положительными были 90% из 491 сообщения, полученного в рамках предоставления обратной связи.

Однако вовлеченность участников оценивается не только по их положительным впечатлениям, но также и по их усилиям, направленным на предоставление конструктивной критики и предложение улучшений, которые можно было бы реализовать в исследовательских играх, а также по любопытству, проявленному некоторыми игроками касательно цели проекта. Именно поэтому мы оценивали вовлеченность на основании двух дополнительных параметров.

2) **Наличие воспоминаний о первой исследовательской игре.** Из 115 сообщений, оставленных игроками после прохождения Dubious, в 19% были упоминания предыдущей игры TESSA. Хотя мы не смогли прийти к какому-либо определенному выводу, основанному на этом результате, мы полагаем, что вероятность наличия воспоминаний об участии в традиционном опросе через четыре месяца после его завершения должна быть довольно низкой. Поэтому мы предполагаем, что такой показатель для нашей исследовательской игры можно оценить как выше среднего уровня.

3) **Процентная доля игроков, выполнивших необязательное бонусное задание по заполнению дневника в игре Dubious.** Из 902 игроков, прошедших Dubious, 19% выполнили обязательное задание по заполнению дневника в конце игры. При этом 21% из них заполнили только части «утро» и «день». Этот результат оказался выше, чем ожидали мы с Лизбет, в частности, потому, что это было дополнительное задание, представленное по окончании игры,

на прохождение которой у участников ушло около 25–30 минут. Другим признаком вовлеченности стали большие объемы текстов, написанных игродентами при заполнении дневника в ответ на просьбу «представить один день из своей жизни в 2030 году». При этом сами тексты оказались очень откровенными и описательными, а в одном-двух случаях участники отвечали настолько искренне, что описали даже такие стороны своей жизни, которые обычно относят к числу табуированных.

10. Рассуждения, предложения и идеи по поводу дальнейшего проведения исследовательских онлайн-игр

- В Dubious не были собраны данные о нажатии кнопки отключения звука. Это могло бы дать информацию о влиянии звукового оформления на полученные сведения и вовлеченность участников в исследовательскую игру. В будущем данный вопрос может быть изучен дополнительно; было бы интересно выяснить, существуют ли какие-либо различия в данных от респондентов, которые слушали и не слушали музыку и звуковые эффекты.
- В будущих исследовательских играх будет проводиться тщательное тестирование механизма сбора данных о желании игродентов сыграть в подобные игры снова, с возможностью выбора из вариантов ответа «Да» и «Нет». Как было сказано выше, соответствующая информация не была получена после прохождения участниками игры Dubious из-за досадного технического сбоя, однако она была собрана после прохождения игры TESSA.
- Представляется полезным в последующих играх ввести возможность записи голоса игродентов в момент предоставления обратной связи. Это помогло бы лучше понять их мнение об игре и устранить недопонимание при анализе полученных результатов.
- Представляется полезным провести сравнительный анализ традиционных онлайн-опросов с исследовательскими играми

для сравнения результатов по количеству откликов, средней длительности опроса и качеству полученных данных с целью понять различия в степени вовлеченности участников исследования.

- Представляется полезным провести сравнительный анализ игровых методов с традиционными онлайн-опросами в развивающихся странах, поскольку подобные исследования на основе сравнения геймифицированных и обычных опросов до сих пор проводились преимущественно в западных странах.
- В ходе разработки рассматриваемых исследовательских игр было реализовано множество различных нестандартных решений, например, в области музыкального сопровождения, всплывающих подсказок, аватаров, нарратива, графического контента, системы уровней и вознаграждений и т.д. Поэтому невозможно сделать выводы о влиянии на вовлеченность респондентов каждого из этих элементов в отдельности. Будущие исследования должны рассмотреть все эти функции более подробно и/или сосредоточиться на влиянии и ценности отдельных функций или компонентов игры.
- В будущем раздел для «комментариев и пожеланий» в конце исследовательской игры можно сделать обязательным для заполнения с тем, чтобы получить обратную связь от каждого игрока. Однако у такого подхода есть и свои недостатки, так как добровольно предоставленная обратная связь сама по себе может служить признаком вовлеченности участников. Отказ от системы добровольного комментирования повысит степень сомнения в том, что игроки предоставляют обратную связь, потому что действительно хотят сделать это, а не потому, что их к этому принуждают.

11. Выводы

Рассмотрение всех полученных в ходе исследования результатов убедительно доказывает, что формат исследовательской игры действительно повышает уровень вовлеченности участников в онлайн-опрос, а также вероятность участия в подобных исследовательских

играх в будущем. В рамках углубленного анализа Лизбет пришла к выводу, что представители всех возрастных групп были в равной мере вовлечены в игровой процесс. Более 90% положительных отзывов, оставленных в рамках добровольной обратной связи, показатель численности продолжающих участие в проекте на уровне 81%, 21% выполнивших необязательное задание по заполнению дневника в Dubious и 19% игроков, упомянувших первую исследовательскую игру (TESSA) в ходе комментирования второй игры, указывают на достижение высокого уровня вовлеченности участников.

Ни одну из предложенных исследовательских игр нельзя назвать образцом для создания любых последующих игровых онлайн-опросов. Однако данные, полученные от игроков, указывают на то, что такие исследовательские решения меняют представления о том, как можно разрабатывать и развивать онлайн-опросы, повышая уровень вовлеченности участников исследования. TESSA и Dubious не были первыми исследовательскими играми, созданными Research Through Gaming Ltd, однако они стали первым проектом в области научных социальных исследований. Также это был первый случай применения исследовательских игр от Research Through Gaming для такой широкой возрастной аудитории (18–65+ лет). Полученные в итоге результаты и собранные данные могут убедить специалистов в возможности использования исследовательских игр в качестве методики проведения научных и коммерческих исследований. При этом возможности для совершенствования исследовательских игр, создания их более удачных и полезных версий весьма велики. В частности, для этого необходимо учитывать конструктивную критику, полученную от игроков.

ЛИТЕРАТУРА

1. GRIT Report Spring (2011). URL: <http://issuu.com/researchshare/docs/grit-spring-2011-full?e=2864560/2653077>
2. GRIT Report Spring (2012) URL: <http://issuu.com/researchshare/docs/grit-spring-2012-full-report>
3. GRIT Report Winter (2013) URL: <http://issuu.com/researchshare/docs/grit-winter-2013>
4. GRIT Report (2014) URL: <http://www.greenbook.org/grit>
5. Sleep D. (2012) Improving Online Market Research Through Gamification [Онлайн], Digital Marketing Hub. The Guardian, 15 августа 2012 г. URL: <http://www.theguardian.com>

- ian.com/media-network/media-networkblog/2012/aug/15/online-market-research-gamification. [Дата обращения: 20/12/14].
6. Koenig-Lewis N. Marquet M., & Palmer A. (2013) The Effects of Gamification on Market Research Engagement and Response, In: Doherty, A. M. (ed.): Marketing Relevance -Academy of Marketing Conference, 8–11 июля 2013 г., Pontypridd: University of South Wales, ISBN: 978–1–909838–00–0. URL: http://marketing.conference-services.net/resources/327/3554/pdf/AM2013_0291_paper.pdf [Дата обращения: 11/12/14].
 7. Adamou B. (2011) The Future of Research Through Gaming, *Proceedings of CASRO Online conference, Las Vegas, USA*. [Онлайн] URL: <http://www.researchthroughgaming.com/publication/view/the-future-ofresearch-through-gaming/> [Дата обращения: 29/12/14].
 8. Mavletova A. (2014) *Web Surveys Among Children and Adolescents: Is There a Gamification Effect?* Social Science Computer Review pages 1–27. SAGE Publications. URL: <http://ssc.sagepub.com/content/early/2014/08/01/0894439314545316.abstract> [Дата обращения: 11/12/14].
 9. Harrison P (2011) The Researchification of Games, *Proceedings of ESOMAR 3D Conference, Miami, United States*. URL: http://media.brainjuicer.com/media/files/The_Researchification_of_Games_-_ESOMAR_3D_2011.pdf [Дата обращения: 28/12/14].
 10. Bogost I. (2011a) *Persuasive Games: Exploitationware*, Gamasutra [Онлайн]. Опубликовано в мае 2011 г. URL: http://www.gamasutra.com/view/feature/6366/persuasive_games_exploitationware.php [Дата обращения: 22/12/14].
 11. Bogost I. (2011b) *Gamification Is Bullshit* [Онлайн] The Atlantic. URL: <http://www.theatlantic.com/technology/archive/2011/08/gamification-is-bullshit/243338/> [Дата обращения: 08/10/14].
 12. Driver P. (2011) *The Irony Of Gamification*. Issue 3, British Council IED Magazine. URL: <http://issuu.com/britishcouncilportugal/docs/inenglishdigital3/2>. p.22, 23, 24. и <http://ayogo.com/blog/ayogo-qa-what-is-the-difference-between-a-game-and-gamification/> [Дата обращения: 28/05/14].
 13. Csikszentmihalyi M. (1997) *Finding Flow: The Psychology of Engagement With Everyday Life*. New York: BasicBooks.
 14. Ravaja N., Saari T., Salminen M., Laarni J. & Kallinen K. (2006) *Phasic Emotional Reactions to Video Game Events: A Psychophysiological Investigation*, Media Psychology. 8:4. 343–367. DOI: 10.1207/s1532785xmep0804_2
 15. Frome J. (2007) *Eight Ways Videogames Generate Emotion*. Proceedings of DiGRA 2007 Conference. URL: <http://www.digra.org/wp-content/uploads/digital-library/07311.25139.pdf> [Дата обращения: 30/12/14]
 16. Soman D. & Huang HY, W. (2013) *A Practitioners Guide to Gamification in Education*. URL: <http://inside.rotman.utoronto.ca/behaviouraleconomicsinaction/files/2013/09/GuideGamificationEducationDec2013.pdf> Rotman School of Management, University of Toronto.
 17. Deterding S., Dixon D., Khaled R. and Nacke L. (2011a). *Gamification: Toward A Definition*. CHI 2011, Vancouver, Canada. URL: <http://gamification-research.org/wp-content/uploads/2011/04/02-Deterding-Khaled-Nacke-Dixon.pdf> [Дата обращения: 30/12/14].
 18. Deterding S., Dixon D., Khaled R., & Nacke L.E. (2011b) *From Game Design Elements To Gamefulness: Defining 'Gamification'*. Proceedings of Mindtrek Conference, Tampere, Finland. URL: <http://www.slideshare.net/dings/from-game-design-elements-to-gamefulness-defininggamification> [Дата обращения: 25/11/14].

19. Zichermann G. and Cunningham C. (2011) *Gamification by Design*, Sebastopol, Calif: O'Reilly Media.
20. McGonigal J. (2011) *Reality Is Broken, Why Games Make Us Better And How They Can Change The World*. New York: Penguin Paperbacks. 2011. Печатная версия.
21. Costikyan G. (2002) I Have No Words & I Must Design: Toward a Critical Vocabulary for Games, Computer Games and Digital Cultures Conference, ed. Frans Mäyrä. Tampere: Tampere University Press. 2002.
22. ESOMAR (Eds.) York, S., and Poynter, R. (2013) *Answers To Contemporary Market Research Questions* (3rd Ed.) Издатель неизвестен. Место публикации неизвестно.
23. Ariely D., Kamenica E. & Prelec D. (2008) *Man's Search For Meaning: The Case Of Legos*, Journal of Economic Behaviour and Organization [Онлайн], URL: <http://people.duke.edu/~dandan/Papers/UpSide/meaning.pdf> [Дата обращения: 29/12/14].
24. Kamenica E. (2012) *Behavioural Economics and Psychology of Incentives*. Doi: 10.1146/annurev-economics-080511-110909. Annu. Rev. Econ. 2012. 4:13.1–13.26. URL: <http://faculty.chicagobooth.edu/emir.kamenica/documents/behavioralIncentives.pdf>
25. Kaden R. J., Linda G., & Prince M. (2011) *Leading Edge Marketing Research: 21st-century Tools and Practices*. U.S.A.: SAGE Publications
26. Pallini K. Verhaeghe A. Langer B. (2014) *Collaborating Beyond Responding: Creating Engaging Surveys Which Capture The Complex Consumer Reality*. Proceedings of Market Research Society Annual Conference, 2014. Page 2. WARC Publications. [Дата обращения: 20/12/14].
27. True Sample (2012) *Online Research Quality: The Next Frontier*. URL: <https://truesample.com/white-paper/online-research-quality-the-next-frontier/>. [Дата обращения: 09/05/15].
28. Nulty D. (2008) *The Adequacy of Response Rates to Online and Paper Surveys: What Can Be Done?*. Assessment & Evaluation in Higher Education. 33 (3), p301–314. URL: <http://www.ua.edu/files/uafgov/fsadmin-nulty5-19-10.pdf>
29. Kaul A., Narang R. & Shant M., (2014) *Gamijoint: Improving Conjoint Through Gamification*, Proceedings of CASRO Digital Research Conference conference, San Antonio, TX, URL: http://cymcdn.com/sites/www.casro.org/resource/collection/E0F10496-BE87-48E8-8746-521D403EE4A2/Paper_-_Anil_Kaul_and_Rajat_Narang_-_AbsoluteData_Analytics.pdf [Дата обращения: 30/12/14].
30. Malinoff B., and Puleston J., (2011) *How Far Is Too Far? Traditional, Flash and Gamification Interfaces, and Implications for the Future of Market Research Online Survey Design*. Proceedings of ESOMAR 3D Digital Dimensions conference. Miami. U.S.A.: WARC
31. Downes-Le Guin T., Baker R., Mechling J., & Ruylea E. (2012) *Myths and Realities of Respondent Engagement in Online Surveys*. International Journal of Market Research, 54 (5), 1–21. doi: 10.2501. URL: http://www.marketstrategies.com/user_area/content_media/Online_Survey_Engagement.pdf. [Дата обращения: 28/05/14].
32. Juul J. (2003) *The Game, the Player, the World: Looking for a Heart of Gameness*. Level Up conference in Utrecht, NL. URL: http://ocw.metu.edu.tr/pluginfile.php/4471/mod_resource/content/0/ceit706/week3_new/JesperJuul_GamePlayerWorld.pdf [Дата обращения: 23/11/14].
33. McGonigal J. (2011b) *We Don't Need No Stinkin' Badges: How To Re-invent Reality Without Gamification*. Proceedings of Serious Games Summit, U.S.A. URL: <http://www.gdcvault.com/play/1014576/We-Don-t-Need-No> [Дата обращения: 30/12/14].
34. Fleming N. (2012) *Gamification: Is It Game Over?* [Онлайн] URL: <http://www.bbc.com/future/story/20121204-can-gaming-transform-your-life> [Дата обращения: 27/05/14].

35. Pettit A. (2012) Gamification of Surveys in the Real World. URL: <http://www.business2community.com/marketing/gamification-of-surveys-in-the-real-world-0301831>. [Дата обращения: 01/10/14].
36. Survey Analytics: How to Design a Gamified Survey. URL: <http://www.slideshare.net/surveyanalytics/how-to-design-a-gamified-survey> Дата обращения: 10 декабря 2014 г.
37. Adamou B. (2013a) ResearchGames As A Methodology: The impact Of Online ResearchGames Upon Participant Engagement And Future ResearchGame Participation, Association of Survey Computing Conference, University of Winchester, UK. URL: <http://www.researchthroughgaming.com/publication/view/researchgames-as-a-methodology-the-impact-of-online-researchgames-and-game-components-upon-participant-engagement-and-future-researchgame-participationby-betty-adamou/> [Дата обращения: 29/12/14].
38. Adamou B. (2013b) T.E.S.S.A. Undercover Agents And Dubious: Two ResearchGames to Revolutionize Market Research, User Centered Design (UCD) Conference. London, UK.
39. Adamou B. (2012) [Интервью] *Jugar, Una Herramienta Para Conocer Tendencias Del Mercado* (перевод на английский: *Play, A Tool To Meet Market Trends*). URL: http://www.portafolio.co/portafolio-plus/jugar_una-herramienta-conocer-tendencias-del-mercado [Дата обращения: 09/02/15].
40. Newman J. (2008) *Playing with Videogames*. New York: Routledge. pvii-viii.
41. Brown D. (2007) *Gaming DNA — On Narrative and Gameplay Gestalts*. Proceedings of DiGRA 2007 Conference. Authors & Digital Games Research Association (DiGRA) URL: <http://www.digra.org/wpcontent/uploads/digital-library/07311.40380.pdf>
42. Van Zoonen L., Briggs P., Norval A., Wilson S., Flores L.G., Harvey J., Prasopoulou E., Thomas L. & Walker S. (2012) *Scenarios of Identity Management in the Future*. IMPRINTS First Year Report. URL: http://www.imprintsfutures.org/assets/images/pdfs/Scenarios_of_identity_management_in_the_future_Report.pdf. [Дата обращения: 09/02/15].
43. Turner G., Van Zoonen L., and Adamou B. (2013) *Research through Gaming: Public Perceptions of (the Future of) Identity Management*. SAGE Book of Methods. URL: <http://dx.doi.org/10.4135/978144627305013496519>
44. Parsons G. Connor M. James M. (2010) *Music and the World of Warcraft*. URL: <http://www.richardcolby.net/files/Parsonsetal.pdf>. Дата обращения: 30 сентября 2013 г.
45. Автор неизвестен. (2013). *Silence is not golden in the workplace, say business owners*. URL: <http://www.musicworksforyou.com/index.php/news-and-charts/news/207-silence-is-not-golden-in-the-workplacesay-business-owners>. Дата обращения: 20 августа 2013 г.
46. "Tube Heads Metro's Classical Tune" URL: <http://news.bbc.co.uk/1/hi/england/4710426.stm> Дата последнего обновления: 13 февраля 2006 г. Дата обращения: 15 ноября 2013 г.
47. "How can in store music increase sales?" URL: <http://www.retail-week.com/how-can-in-store-music-increasesales/5029706.article> Дата последнего обновления: 30 сентября 2011 г. Дата обращения: 15 ноября 2013 г.
48. Music Thanatology Association International. URL: http://www.mtai.org/index.php/what_is
49. Naqvi N., Shiv B., & Bechara A. (2006). The role of emotion in decision making a cognitive neuroscience perspective. *Current Directions in Psychological Science*, 15 (5), 260–264.

50. Yoo B., Donthu N., & Lenartowicz T. (2011) “*Values at the Individual Level: Development and Validation of CVSCALE*”, Journal of International Consumer Marketing 23:3–4, 193–210. URL: http://people.hofstra.edu/Boonghee_Yoo/cvscale.pdf Дата обращения: сентябрь 2013 г.

51. Ryan R. M. & Rigby S. (2011) *Glued To Games: How Video Games Draw Us In And Hold Us Spellbound*. U.S.A.: Praeger

52. Kucirkova N et al. (2010). “*Sharing Personalised Books: A Practical Solution to the Challenges Posed by Home Book Reading Interventions*”. Literacy Information and Computer Education Journal. 1 (3). p187–191.

53. Jørgensen K. (2009) “*I’m overburdened!*” *An Empirical Study of the Player, the Avatar, and the Gameworld*”. Bergen: Department of Information Science and Media Studies. p1–8.

Приглашаем вас принять участие в исследовательских играх, описанных в данной работе:

Исследовательская игра TESSA Undercover Agents™:

<http://imprintsutures.org/tessa/>

Исследовательская игра Dubious™:

<http://imprintsutures.org/dubious/>

Предлагаем вам посмотреть трейлеры представленных исследовательских игр:

Трейлер исследовательской игры TESSA Undercover Agents™:

<https://www.youtube.com/watch?v=nU1jZcSQxmg>

Трейлер исследовательской игры Dubious™:

<https://www.youtube.com/watch?v=mmo7UxE3cw>

Благодарность: Выражаю огромную благодарность профессору Лизбет ван Зуунен, доктору Джорджине Тернер и всей команде проекта IMPRINTS Futures, а также всем нашим геймспондентам, без которых у меня не было бы возможности изучить исследовательские игры с совершенно новой стороны. Также хотелось бы от всего сердца поблагодарить наших поставщиков и мою команду из Research Through Gaming: Дэвида Визниовски, Уильяма Пойнтера и Пола МакГи, без которых мне бы не удалось реализовать свои идеи и написать эту работу. Благодарю профессора Дэвида Биркса за всю ту помощь, что он оказал мне в процессе подготовки данной работы и ее переработке из моих обычных записей в стиле «говорю и пишу ненаучным языком». Также хочется обратиться к тем, кто прочел

эту работу: благодарю вас за потраченное на это время и надеюсь, что вы также воспользуетесь возможностью прочесть мою будущую работу, посвященную исследовательским играм как новой методике опроса, и мою докторскую диссертацию, которая будет опубликована на нашем партнерском сайте: www.researchgames.org.uk.

Предсказывая будущее: начала научного прогнозирования

*Джон Пулестон
(Global Market Insite, Inc),
Александр Уитли
(Lightspeed GMI);
Хубертус Хофкирхнер
(Prediki Prognosedienste GmbH)*

Введение

Данная работа посвящена прогнозированию. В ней говорится о том, какие прогнозы у нас получаются хорошо, а какие — не слишком, что влияет на нашу способность сделать надежный прогноз и каким образом меняется собираемая нами информация, если, задавая вопрос, мы просим высказать предположение о будущем. Предложены рекомендации относительно того, как сформулировать подобные вопросы и провести прогностическое исследование. Статья основана на результатах более тридцати исследовательских экспериментов, проведенных более чем за три года, мета-анализе более 500 прогнозов и широком обзоре научной литературы на данную тему. Мы надеемся, что наша работа послужит руководством для специалистов по маркетинговым исследованиям, желающим использовать в своих проектах прогностические методики.

История создания: от геймификации к науке прогнозирования

В 2010 г. мы начали изучать роль геймификации в маркетинговых исследованиях, в частности, попробовали использовать в опро-

сах игры на угадывание. В ходе последних респонденты должны были угадывать правильные ответы, делая виртуальные ставки на различные варианты. Вскоре мы выяснили, что респондентам нравится играть в подобные игры, особенно если впоследствии они узнавали правильный ответ.



Увлечательность



Мы заметили, что участники исследований, отвечая на такие вопросы, из-за стремления победить прикладывают гораздо больше усилий. Помимо этого, используя методики такого рода, мы часто отмечали существенное повышение качества данных.

Например, решая задачу по оценке рекламы, мы выяснили, что отказ от простого рейтингования в пользу прогностической игры

со ставками позволяет получить данные, коррелирующие с результатами традиционных рейтингов на уровне в 0,89, но при уровне дифференциации данных повышался более чем в пять раз.

Сколько акций Вы хотели бы купить или продать?



Осталось
акций:
12

-3

-2

-1

+1

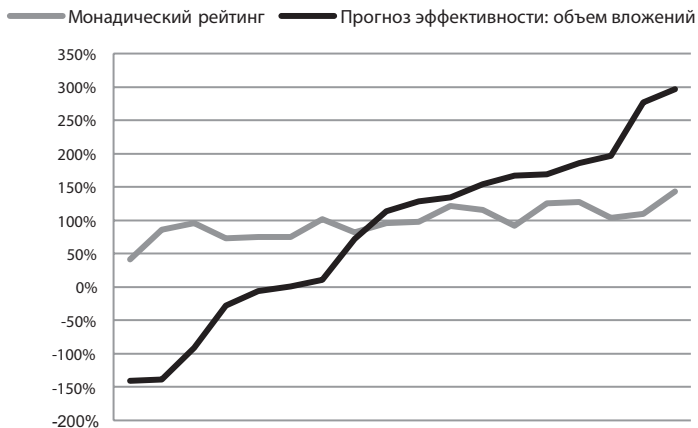
+2

+3

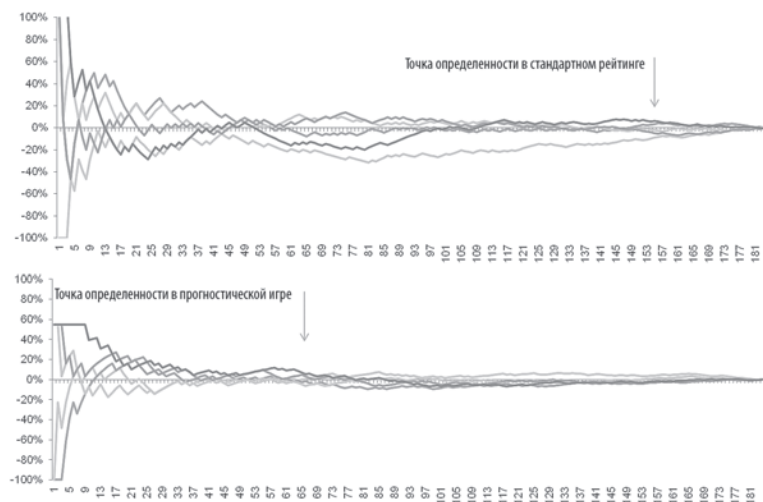
ПРОДАТЬ

КУПИТЬ

Оценка двадцати различных рекламных материалов



Мы также выяснили, что благодаря этому подходу можно с высокой степенью достоверности получить те же ответы, уменьшив выборку более чем в два раза.



Поэтому в ходе экспериментов мы стали переводить вопросы всех типов в прогностические задания. Например, вместо того, чтобы поинтересоваться мнением участников исследования об определенном бренде, мы просили их угадать мнение других людей о нем. Вместо оценки поступков, совершенных в прошлом, мы просили респондентов предсказать свои будущие поступки. Вместо вопроса о том, сколько они готовы заплатить за перечисленные продукты, мы просили спрогнозировать стоимость последних, а вместо вопроса о знании брендов мы показывали часть логотипов, слоганы или рекламу без указания названий и просили их угадать бренд.

Используя прогностические типы вопросов, мы начали собирать довольно большие объемы данных для того, чтобы провести прямое сравнение с результатами использования более традиционных формулировок. В процессе работы появилось множество вопросов о том, как отличаются полученные таким образом данные от данных, собранных в рамках традиционных опросов. Так началось наше путешествие в мир науки прогнозирования.

Для более тщательного изучения темы прогнозирования мы решили провести ряд масштабных экспериментов. В ходе последних мы просили участников опросных групп делать разнообразные про-

гнозы: результатов футбольных матчей, погоды, стоимости различных потребительских товаров, победителей телевизионных игровых шоу, объемов продаж отдельных продуктов, а также прогнозы в области политики и экономики. За двенадцать недельных экспериментов мы собрали более 500 прогнозов. Нашей целью было понять, что мы можем прогнозировать, а что не можем, и что влияет на нашу способность делать хорошие прогнозы.

Такого рода заигрывания с прогнозной тематикой не могли не привести нас в удивительный мир рынков предсказаний, находящийся на вершине достижений в области прогнозирования. Мы начали размышлять о том, как методы рынка предсказаний можно интегрировать в обычные исследования. Это привело нас к сотрудничеству с компанией *Prediki*, поставляющей программное обеспечение для рынков предсказаний, а также к проведению серии микроэкспериментов (более 50), в ходе которых прогнозы традиционных маркетинговых исследований сравнивали с результатами, полученными на рынках предсказаний.

Настоящая работа представляет собой краткую версию наших путевых заметок, сделанных в ходе этого путешествия. Мы провели довольно масштабное исследование на данную тему, однако проделанную работу следует рассматривать лишь в качестве предварительного исследования. Можно сказать, что мы рассмотрели предметную область предсказаний, не углубляясь в отдельные вопросы. В результате проведенного исследования мы сформулировали столько же новых вопросов, сколько ответов нам удалось найти на поставленные ранее.

Ненадолго погрузившись в мир научного прогнозирования, мы лишь можем сравнить себя с небольшой лодкой, приблизившейся к огромному нефтеналивному судну, нагруженному результатами множества исследований других, более известных специалистов, представляющих как коммерческие, так и научные организации.

Кто бы ни приступал к занятиям по данной теме, он не сможет не отдать должное превосходным исследованиям агентства *Brainjuicer*, посвященным рынкам прогнозов, изумительному труду Филипа Тетлока, изучавшему этот вопрос в течение пары десятков лет с научной точки зрения и собравшему данные более 15 000 гео-

политических прогнозов, а также работе Нейта Силвера, в одиночку превратившему науку прогнозирования в отличную тему для разговоров на любой модной вечеринке.

Мы очень надеемся, что те, кто уже знаком с данной темой, все же смогут найти в нашей работе что-то новое. Тем же, кто еще не знаком с миром прогнозов, могу сказать: пристегните ремни! По моим прогнозам, вас ожидает невероятно увлекательное путешествие.

Итак, с чего начнем?

Часть 1: Понимание полученных данных

Для начала сделаем шаг назад и обобщим все, что мы узнали в результате проведенных прогностических экспериментов. Вот что у нас получилось.

Математика прогнозирования

Качество прогноза основывается на нескольких зависимых переменных, в числе которых качество доступной исходной информации; усилия, затраченные на ее анализ; сложность интерпретации этой информации; объективность человека, проводящего анализ; наконец, присущий данной ситуации элемент случайности. Формула оценки выглядит следующим образом:

Качество прогноза = информация $\times$ усилия $\times$ объективность $\times$ (1 – сложность) $\times$ R

Специалисты по маркетинговым исследованиям обратят внимание на тот факт, что в этом уравнении нет прямой зависимости результата от размера выборки. То есть превосходный прогноз способен в одиночку сделать хорошо осведомленный, объективный человек с доступом к необходимой информации, обладающий навыками её интерпретации. (Например, во время последних президентских выборов именно так удалось предсказать их результаты во всех 50 штатах США — эта история описана Нетом Силвером (http://en.wikipedia.org/wiki/Nate_Silver)).

Какие прогнозы нам удаются?

Мы проанализировали способность наших респондентов делать прогнозы в самых разных областях — от стоимости потребительских товаров, поведения и мнений других людей, погоды и будущих покупок других потребителей и заканчивая ситуациями, в которых можно только угадать, — например, определяя лот, который на предстоящем аукционе будет продан по самой высокой цене. В итоге мы собрали около 400 различных прогнозов по семи категориям.



Можно с уверенностью сказать, что некоторые прогнозы мы делаем лучше, чем другие. В частности, — и это неудивительно — пройдя долгую эволюцию социального взаимодействия друг с другом, мы лучше всего можем предсказывать поступки и мнения других людей, хотя и в этом не достигли совершенства. Гораздо худшие результаты мы демонстрируем в ситуациях, когда необходимо провести более сложную оценку, требующую аналитического мышления, например, при прогнозировании покупательского поведения и стоимости товаров.

Прогнозы составляются на основе знаний и информации

Проблема большинства прогнозов состоит в нехватке и разрозненности сведений, необходимых для их составления. Зачастую человек, делающий прогноз самостоятельно, не обладает достаточной

информацией для обеспечения его точности. При этом у некоторых людей объем знаний больше, чем у других. Например, наша способность предугадывать стоимость бытовых товаров из предложенного списка в среднем оказывается лишь на 8% лучше, чем случайный выбор одного из вариантов. (В ряде тестов мы показывали участникам четыре различных варианта возможной стоимости для различных бытовых товаров и просили их сделать прогноз, выбрав один из них.) С другой стороны, группы потенциально способны делать более точные прогнозы, чем отдельные индивидуумы, а при наличии корректных механизмов агрегирования имеющихся у членов группы данных прогнозы могут стать гораздо точнее благодаря так называемой коллективной мудрости.

Этот феномен был открыт Фрэнсисом Гальтоном в начале 20 века и получил известность благодаря книге Джеймса Шуровески «Мудрость толпы». Автор последней попросил посетителей деревенской ярмарки угадать вес выставленного на всеобщее обозрение быка и обнаружил, что медианное значение их догадок отличается от реального веса животного в пределах 1%.

Ученые провели множество исследований, посвященных изучению процессов проявления коллективной мудрости (например, [22, 21], которые помогли нам понять, что прогноз каждого человека состоит из двух элементов — информации и ошибки. Если суждение каждого индивидуума будет независимым и непредвзятым, доля ошибки сократится, а процесс агрегации данных позволит извлечь необходимые знания из имеющегося массива.

Филип Тетлок смог измерить коллективную мудрость групп экспертов, проанализировав более 15 000 среднесрочных и долгосрочных политических и экономических прогнозов. По его оценкам, байесовский метод агрегации мнений в группе, в среднем показывает результаты по меньшей мере на 15% лучше, чем в случае индивидуальных прогнозов. Его книга «Экспертное политическое мнение» — одна из обязательных для прочтения всем желающим разобраться в научном прогнозировании.

В попытке повторить этот оригинальный эксперимент по оценке коллективной мудрости в XXI веке мы попросили респондентов попытаться спрогнозировать стоимость продажи нового устрой-

ства Apple iPad mini в Великобритании за месяц до его выхода. Мы сообщили им цену предыдущей модели и попросили предсказать, сколько будет стоить новая модель в розничных магазинах, введя соответствующее число в специальную ячейку¹. Только 10% респондентов смогли дать ответ в пределах 5%-го отклонения от реальной стоимости, однако коллективная мудрость предоставила прогноз в пределах всего 1%.



Медианное значение предполагаемой
стоимости = 316 фунтов
Реальная стоимость = 319 фунтов
Ошибка = 1%

Не забывайте, что толпа может оказаться глупой

Из чистого любопытства мы также попросили наших респондентов угадать вес быка по тому же методу, что и в оригинальном эксперименте. Сегодня мы не обладаем реальными знаниями о весе коров на коллективном уровне, а многие панелисты даже никогда не видели их вблизи, поэтому прогноз, сделанный их коллективной мудростью, оказался прискорбно ошибочным — отклонение составило 36%.

Ярмарка в графстве Плимут в 1906 г.



Медианное значение предполагаемого
веса = 1207 фунтов
Реальный вес = 1198 фунтов
Ошибка = <1%

¹ чтобы избежать эффектов якорения

Онлайн-опрос в 2014 г.



Медианное значение предполагаемого веса = 350 кг
Реальный вес = 550 кг
Ошибка = 36%

Это очень важное наблюдение. Мудрость толпы проявляется только в том случае, когда существует общая база знаний и информация, на основе которой можно сделать прогноз. В оригинальном эксперименте Гальтона принимали участие фермеры, у каждого из которых был релевантный жизненный опыт и знания о весе коров. В отсутствие подобных коллективных знаний толпа может оказаться весьма глупой.

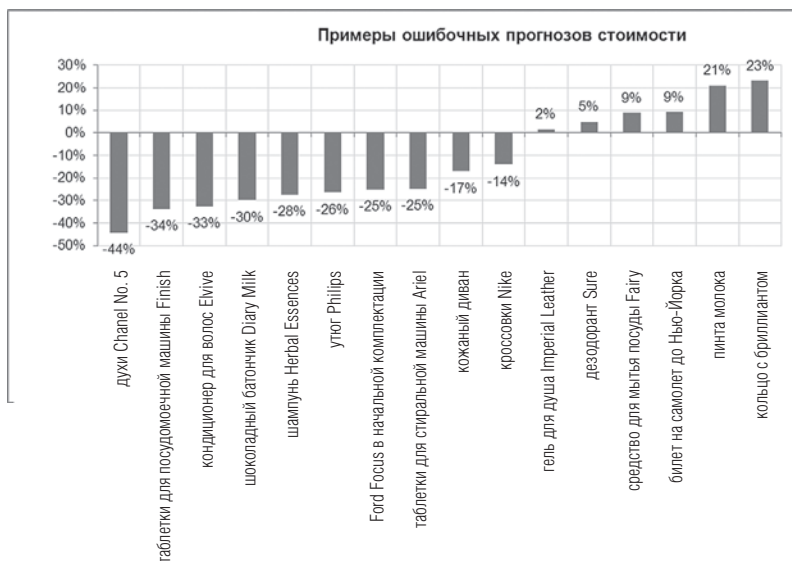
Вероятно, самый важный вопрос, который необходимо задавать самому себе, опрашивая группы респондентов в рамках прогнозного маркетингового исследования: «Какие знания о предложенной теме они могут привнести в проект, или какую информацию мы могли бы им предоставить и таким образом помочь им сделать верный прогноз?»

Прогнозы толпы могут немного отставать от времени

Когда просишь людей угадать стоимость товаров, имеющихся в продаже, их предположения зачастую основаны на прошлом покупательском опыте. Поэтому высказанные предположения, как правило, оказываются устаревшими на несколько лет. Кроме того, у нас остаются воспоминания о скидках на различные категории товаров, и эта информация также способна вытеснить из нашей памяти сведения об их реальной стоимости.

В своих экспериментах мы просили участников угадать стоимость целого ряда бытовых товаров. Были собраны данные в рамках более 30 тестов. Основываясь на них, мы пришли к выводу, что в среднем стоимость, названная респондентами, на 9% меньше реальной.

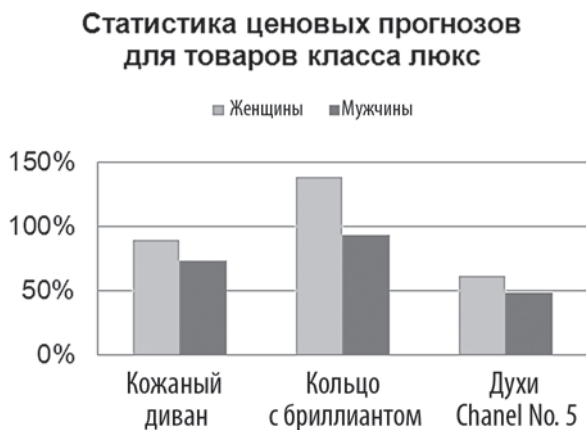
Ниже приведена диаграмма, отражающая результаты различных ценовых прогнозов. Она позволяет составить представление о том, в каком диапазоне они могут варьироваться.



Мы также отметили существенные различия между ценовыми прогнозами в различных гендерных группах. Прогнозы мужчин по стоимости бытовых товаров в среднем на 6% ниже, чем прогнозы женщин. Вероятная причина состоит в том, что мужчины реже совершают покупки. В сегменте товаров класса люкс различия могут оказаться даже более существенными. Например, предполагаемая женщинами цена кольца с бриллиантом на 40% выше цены, предполагаемой мужчинами, и даже выше его реальной стоимости. Можно предположить, что помимо накопленного опыта в данном случае влияние на респондентов оказывают и другие эмоциональные факторы. Эти различия заметили не только мы. Недавнее исследование BDRC Continental² показало, что женщины неизменно указывают более высокую предположительную стоимость брендовых то-

² Источник: <http://www.marketingweek.co.uk/analysis/essential-reads/how-to-stop-your-brand-becoming-a-commodity/4010334.article>

варов различных категорий по сравнению с мужчинами.



Также было выявлено несколько незначительных расхождений в ответах среди представителей разных возрастных групп. Молодежь — вероятно, в силу недостатка жизненного опыта — и пожилые люди — в силу наличия более долгой истории приобретения

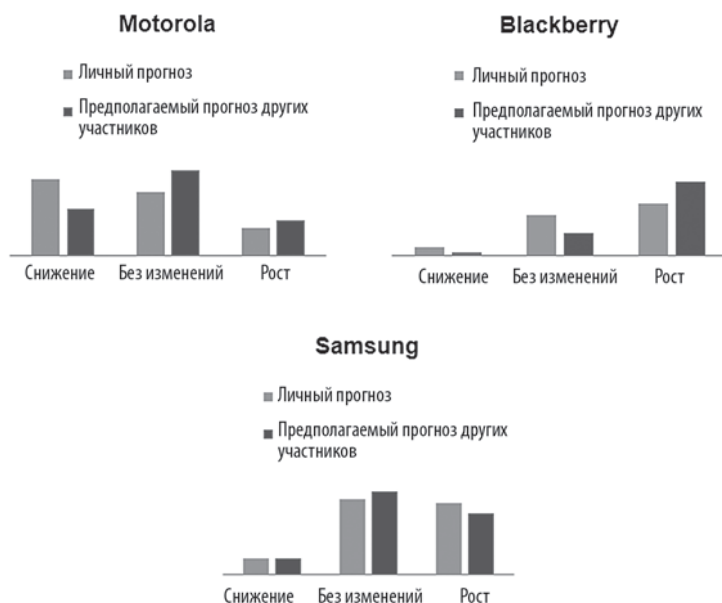
товаров по более низкой стоимости — называют немного более низкие цены. Однако данные различия не столь значимые.

Как посчитать мнение, подкрепленное коллективной мудростью: по медиане, моде или среднему значению?

В оригинальных экспериментах Фрэнсиса Гальтона мнение толпы определялось с использованием медианного значения прогноза, чтобы избежать искажений, вызванных иррациональными выбросами (например, если кто-то высказал предположение, что корова легче перышка). Сегодня процесс работы с выбросами стал проще благодаря Excel. Мы выяснили, что используя средние значения — предварительно отбросив верхние и нижние 10% выбросов, — можно получить чуть более точные данные, чем при учете медианных значений.

Мы думаем, что при прогнозировании толпа окажется не умнее нас

В 2010 г., непосредственно перед скоротечным закатом BlackBerry, подъемом Samsung и во время заметного ухудшения положения Motorola, мы попросили людей спрогнозировать будущее ряда брендов мобильных устройств. Многие респонденты смогли правильно предсказать будущий успех и неудачи этих трех компаний. В ответ на просьбу угадать ответы других людей респонденты высказывали предположения, что толпа менее информирована по сравнению с ними. Поэтому другие ответят, что в будущем с Motorola и Samsung существенных изменений не произойдет, а положение BlackBerry скорее улучшится.



Это очень важный вывод, так как в наиболее успешных из изученных нами методов геймификации опросов используются принципы проективных прогнозов. Однако мы нашли решение этой проблемы, изменив формулировку вопроса. Вместо того чтобы просить участников исследования угадать, что думает толпа, необходимо предложить им угадать мнение группы «экспертов».

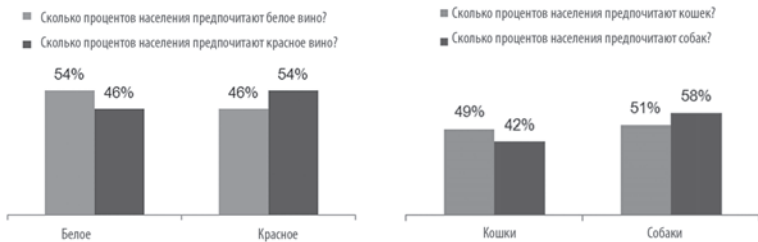
В отсутствие релевантных коллективных знаний или необходимой информации прогнозы можно с легкостью изменить посредством подсказок и эффектов якорения

Делая прогноз, мы используем практически любую доступную информацию, зачастую неосознанно. При отсутствии необходимых сведений на наши прогнозы можно повлиять с помощью намеков и подсказок, бессознательно обрабатываемых мозгом и незаметно меняющих наше мнение. Этот феномен был хорошо описан в бихевиоризме как когнитивное искажение: человеку свойственно при

принятии решений полагаться в большей степени на первые полученные сведения («якорь»).

Ярким примером такого рода влияния может служить следующий эксперимент. Попросите собравшихся в помещении людей угадать, как упадет монета — орлом или решкой. Только потому, что «орел» был упомянут первым и стал «якорем» для последующего когнитивного искажения, 70% присутствующих назовут именно его. Таким образом, прогноз, основанный на коллективных предположениях групп людей, содержит существенное системное искажение в ответах. Он равен 70:30 в ситуации, когда, как мы все знаем, вероятность получения того или иного результата в действительности составляет 50:50.

Для выявления зависимости прогнозов от эффекта якорения и наличия подсказок мы провели серию экспериментов, в ходе которых меняли формулировку вопросов, делая акцент на тот или иной вариант ответа. Таким образом, мы выяснили, что можем изменить процентное соотношение вплоть до 20%, а это гораздо больше, чем мы ожидали получить.



Какое влияние оказывают на прогнозы мнения других людей

Далее мы решили выяснить, как на прогнозы респондентов могут повлиять мнения других людей. Как мнение одного человека воздействует на мнение другого? Допустим, мы просим участников исследования спрогнозировать, какая реклама станет более популярной — А или В? При этом мы сообщаем, что другие люди

посчитали рекламу А лучше. Сколько людей изменят свое мнение в результате внушения³?

Чтобы решить эту задачу, мы провели очень простой эксперимент. Предложив участникам контрольной группы сделать ряд прогнозов самостоятельно, в тестовой группе мы различными способами «внушали» участникам мнения других людей, сообщив им уже сделанный прогноз или показав комментарии в поддержку того или иного варианта ответа.

Всего мы протестировали нашу технику внушения по двадцати различным темам и обнаружили, что изменить ответ респондентов в среднем можно в 9% случаев. Однако результаты заметно варьировались в зависимости от объема доступной информации, полученных знаний для формирования суждения по тому или иному вопросу, а также степени уверенности в них. Ниже приведено несколько примеров.

Тип прогноза	Вопрос	Распределение ответов	Уверенность	Эффект внушения
Сложный расчет	Больше ли людей посетили Америку, чем посмотрели сериал «Жители Ист-Энда»?	30/70	Низкая	20%
Личная оценка	Людям больше понравится реклама А или В?	58/42	Низкая	9%
Личная оценка	Людям больше понравится реклама С или D?	70/30	Высокая	2%
Представление о мнении людей	Какую знаменитость предпочтет большее число людей — Кларксона или Коуэлла?	53/47	Низкая	18%
Представление о мнении людей	Какую знаменитость предпочтет большее число людей — Гагу или Болта?	52/48	Высокая	2%
Экспертный	Мы сможем вылечить рак к 2050 г.?	56/44	Низкая	9%
Экспертный	Мы не высадимся на Марсе до 2050 г.?	70/30	Низкая	6%

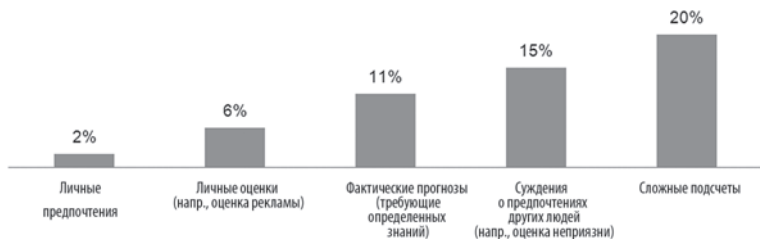
³ Мы определяем «внушение» как изменение мнения в результате представления некоторой информации или внешнего влияния.

Тип прогноза	Вопрос	Распределение ответов	Уверенность	Эффект внушения
Представление о поведении людей	Больше людей выберут для себя автомобиль А или В?	54/46	Низкая	27%
Представление о поведении людей	Больше людей выберут для себя банк А или В?	65/35	Высокая	0%

О тенденциях, отметить на этапе предварительного анализа, можно будет говорить всерьез только после проведения дальнейших исследований. Однако уже сейчас можно зафиксировать несколько наблюдений, на которые в дальнейшем имеет смысл обратить пристальное внимание.

Внушение сильнее влияет на ответы, если вопрос сложный. Прогнозы, основанные на экстраполяции собственных ощущений, гораздо менее подвержены влиянию внушений, чем прогнозы, касающиеся предпочтений других людей. Как в ситуации, когда респондентов попросили угадать, кого из двух британских знаменитостей, вызывающих неприязнь аудитории, — Саймона Коуэлла или Джереми Кларксона — другие люди не любят больше. Когда предмет вопроса связан скорее с фактами, чем с эмоциями, мы более внушаемы, чем в том случае, когда выносим личные суждения, например, о том, какую рекламу предпочтет большее число людей. Примером в данном случае может служить вопрос о том, какой автомобиль является более распространенным — Ford Focus или VW Golf.

Влияние внушения на прогнозы различных типов



Это явление можно сравнить с гравитацией. В отсутствие последней достаточно небольшого воздействия, чтобы отправить

объект в космос. Чем больше у нас информации и знаний, тем мощнее должна быть сила, которая заставит нас отклониться с нашей орбиты.

Совет: внушение можно рассматривать в качестве способа проверки надежности прогноза. Выяснив, насколько сильно меняется мнение респондентов под влиянием простых средств внушения, вы получаете оценку устойчивости полученного прогноза.

В этом эксперименте мы рассматривали только влияние мнения одного человека на другого. Можете себе представить, насколько мощным будет воздействие на человека целой толпы? Мы затронем эту тему далее.

Наше личное мнение может играть доминирующую роль, существенно искажая наши прогнозы

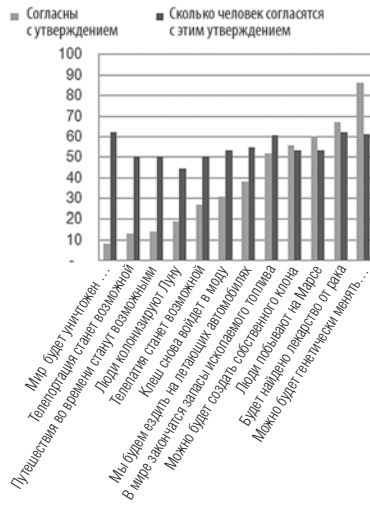
Мы склонны полагать, что большинство людей согласны с нами. На этом предположении основываются многие прогнозы.

Чтобы проиллюстрировать данный тезис, в другом эксперименте мы попросили его участников сделать ряд прогнозов о том, какой будет жизнь в 2050 г. Респондентам было предложено несколько утверждений, с которыми надо было выразить согласие или несогласие. Затем мы попросили этих респондентов угадать, сколько человек согласилось бы с их собственным мнением. По нашим оценкам, человек в среднем полагает, что с его мнением согласятся 60% других людей.

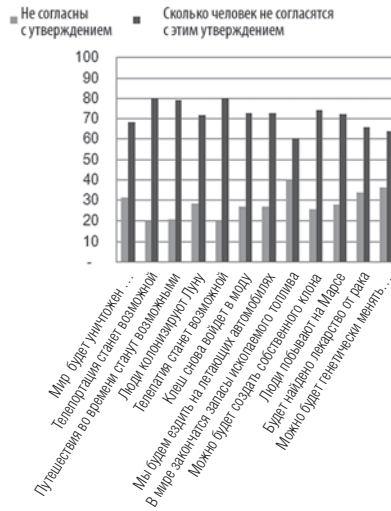
Любопытно, что участники эксперимента, не согласившиеся с предложенными утверждениями и оставшиеся таким образом в меньшинстве, также были склонны считать, что их мнение разделяет большинство (75% среди несогласившихся и 55% среди согласившихся). Филипу Тетлоку наверняка было бы что сказать об этом, так как он установил, что люди со стойкими убеждениями реже дают правильные прогнозы. Также интересно отметить, насколько устойчивой оказывается эта тенденция при ответе на разные вопросы по выбранной тематике.

Другой яркий пример этого явления — политические опросы. Мы попросили 500 респондентов предсказать, кто сформиру-

Выразили согласие



Выразили несогласие



ет следующее Правительство Великобритании. Ответы оказались настолько сильно смещены в сторону собственных политических предпочтений, что мы с равным успехом могли бы просто спросить, за кого респонденты собираются отдать свой голос.



Таким образом, одна из проблем проведения прогностических опросов заключается в том, как сделать ответы респондентов более объективными. Одно из лучших средств достижения этой цели, как мы увидим далее, — дать возможность получить деньги за объективный прогноз.

Толпа также подвержена воздействию когнитивных искажений

Коллективная мудрость также может подпасть под негативное влияние когнитивных искажений. В частности, нашей объективности при составлении прогноза могут существенно помешать наши эмоции.

Прекрасной иллюстрацией этого утверждения служит сравнение прогнозов результатов футбольных матчей, сделанных, с одной стороны, болельщиками играющих команд и, с другой стороны, людьми, не отдающими предпочтение ни одной из команд.

Например, когда мы попросили англичан угадать результаты матчей между командами Англии и Черногории, Германии и Ирландии, респонденты не смогли устоять перед соблазном предска-

зять Англии более масштабную победу, чем Германии. Однако если взглянуть на рейтинги команд, статистически у Германии было больше шансов на победу, учитывая тот факт, что ее результаты в прошлых матчах были лучше, чем у Англии. (Я уверен, что этот эксперимент можно с равным успехом повторить в любой другой футбольной стране мира. Мы везде обнаружим столь же оптимистичный взгляд на игру своей национальной сборной).

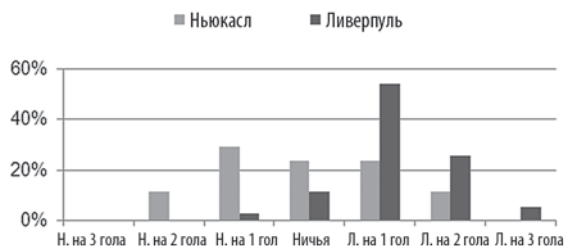
Наши британские респонденты предсказали следующее:

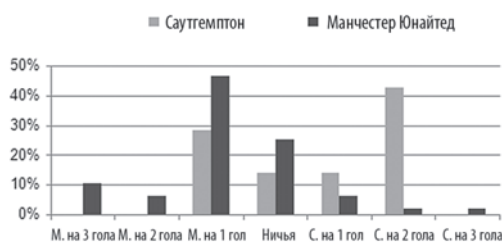
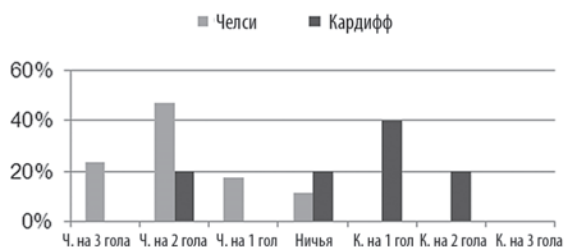
Команда Англии выиграет у команды Черногории с преимуществом в три гола.

Команда Германии выиграет у команды Ирландии с преимуществом в один гол.

Изучая прогнозы результатов матчей местных футбольных команд, мы обнаружили искажение под влиянием эмоций, проиллюстрированное в следующих диаграммах. Похоже, ни один болельщик не может спрогнозировать возможности проигрыша своей команды с отрывом более чем в два гола независимо от того, насколько хороша команда противника — даже если в турнирной таблице национального чемпионата она находится на 20 мест выше. Наоборот, очень немногие болельщики лучших команд страны могут себе представить возможность проигрыша командам, расположенным ниже в турнирной таблице.

Прогнозы результатов матчей, сделанные болельщиками разных команд:





Зачастую эти когнитивные искажения не учитываются при анализе. Однако если все члены группы подвержены влиянию одного и того же когнитивного искажения, могут возникнуть масштабные системные ошибки, негативно влияющие на прогноз группы в целом.

Наша самооценка — область наибольшего влияния эмоциональных искажений

На собрании группе студентов был задан вопрос, примут ли они участие в уборке помещения по окончании мероприятия. Положи-

тельный ответ дали 50%. Студентов другой группы попросили спрогнозировать, сколько всего человек примут участие в уборке. Среднее значение их прогнозов оказалось на уровне 15%. В итоге, подсчет реального числа студентов, оставшихся для уборки помещения, показал около 15%⁴.

Эти результаты демонстрируют мощное когнитивное искажение, вызванное склонностью человека думать о себе хорошо. Всем нам нравится относить себя к тому типу людей, которые обязательно сделают уборку помещения по окончании собрания в нем.

Подтверждение влияния этого искажения мы можем найти во множестве различных маркетинговых опросов. Будете ли Вы голосовать? Будете ли Вы сортировать мусор для переработки? Как часто Вы употребляете алкоголь? и т.п.

Существует также простое искажение уступки, влияющее на прогнозы респондентов при ответе на большинство вопросов категории «Купите ли Вы этот товар?».

Мы можем так запутаться в собственных эмоциях, что будем испытывать трудности с объективной оценкой собственного поведения. Действительно, родители и друзья влюбленной пары, очевидно, способны предсказать ее будущее в долгосрочной перспективе с большей точностью, чем сами влюбленные⁵.

Также мы склонны не видеть за деревьями леса. В одном из экспериментов группе мужчин был задан вопрос, почему им нравится футбол, и получили вполне рациональные ответы. Например, это увлекательная и волнующая игра, нравится наблюдать за мастерством игроков, нравится азарт победы. Когда мы спросили группу женщин, почему мужчинам нравится футбол, одна из них ответила: «Благодаря футболу им есть о чем поговорить в пабе». Пожалуй, последнее утверждение ближе к истине, чем все остальные.

⁴ Примечание от Джона Пулстона: Я неоднократно приводил этот эксперимент в пример, однако, боюсь, сведения об исходном источнике у меня не сохранились, поэтому вы можете рассматривать его как анекдот. Если кто-то может помочь установить этот источник, пожалуйста, свяжитесь со мной.

⁵ Источник: <http://psp.sagepub.com/content/25/11/1417.short>

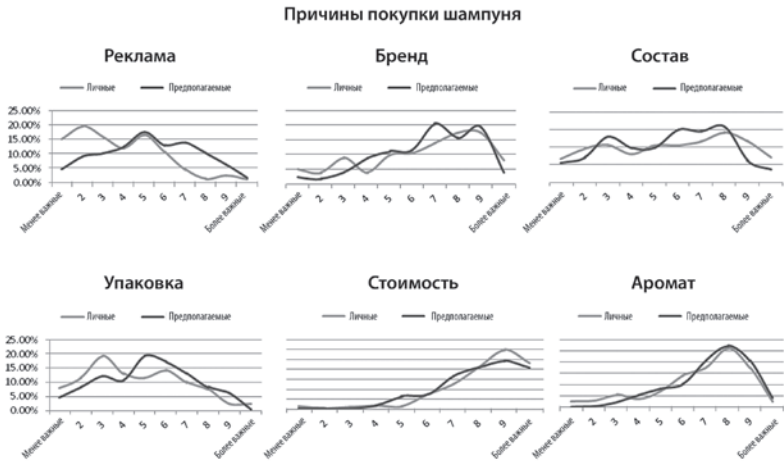
Дело не только в эмоциях, но и в умении выделить из шума важную информацию

Это основная идея книги Нейта Силвера «Сигнал и шум», также обязательной для прочтения любому специалисту в области маркетинговых исследований.

Силвер рассматривает несколько кейсов, включая знаменитый эксперимент Герда Гигеренцера в области бихевиористской экономики. В ходе последнего было установлено, что жители Германии точнее угадывают относительные размеры двух расположенных рядом калифорнийских городов, чем жители Калифорнии.

Наши мнения, мотивация и поведение зачастую очевидны людям, находящимся за пределами воздействия фонового шума. Однако мы редко можем объективно посмотреть на самих себя, так как неспособны отстраниться от шума, создаваемого неосознанными чувствами, которые и являются главными движущими факторами поведения. Калифорнийцам сложнее определить разницу между двумя соседними городами, поскольку они постоянно слышат множество «шумообразующих» упоминаний и сведений об обоих городах, находящихся в пределах одного штата. А жители Германии могут услышать лишь сигнал от более «громкого» города.

Ниже приведены результаты одного из экспериментов, в ходе которого мы попросили участников оценить важность различных факторов, повлиявших на их выбор шампуня. Затем попросили другую группу угадать, насколько сильное влияние оказывают различные факторы на покупателей в целом. Например, можно заметить, что респонденты недооценивают влияние на собственное поведение рекламы, упаковки и бренда, придавая слишком большое значение стоимости и составу. Поэтому определить, насколько в действительности важны эти факторы, довольно сложно. Очевидно, однако, что, задавая вопросы только о личном мнении респондентов, невозможно получить ответы, объективные на 100%. К примеру, все мы склонны преуменьшать степень влияния рекламы на нас лично.



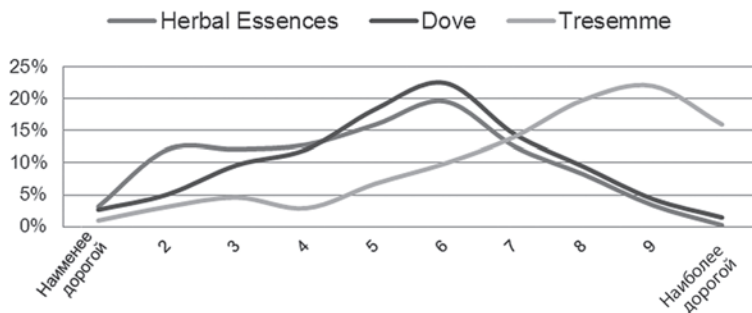
Наши прогнозы могут показать, что мы подсознательно думаем и чувствуем

Прогнозы способны «вытащить» наши неосознанные эмоции для повышения объективности оценки брендов. Ниже приведен пример из эксперимента, в ходе которого мы попросили респондентов оценить ряд банков, а затем спросили, услугами каких из них пользуются они лично. На графиках видно, что результаты оказались довольно тесно взаимосвязаны. Полученный в итоге рейтинг банков вполне соответствует показателям рыночной доли каждого из них и потому не представляет особой ценности. Однако затем мы попросили респондентов спрогнозировать, какой из банков с наибольшей вероятностью продолжит свое существование в 2050 г., что позволило нам выявить более интересные, подсознательные представления о преимуществах каждого из банков, то есть, получить принципиально иной фокус зрения.



Использование принципов прогнозирования для выявления сравнительной ценности различных товаров оказывается крайне полезным на практике и во многих отношениях гораздо более содержательным по сравнению со стандартными ценовыми тестами.

Угадайте стоимость этих шампуней



Проще спрогнозировать наличие леса, чем деревьев

Which instant coffee would you buy?



Прогнозирование может помочь нам более объективно оценить поведение других людей и выявить, что мы думаем о различных вещах (лесах). При этом у нас гораздо хуже получается угадать все множество

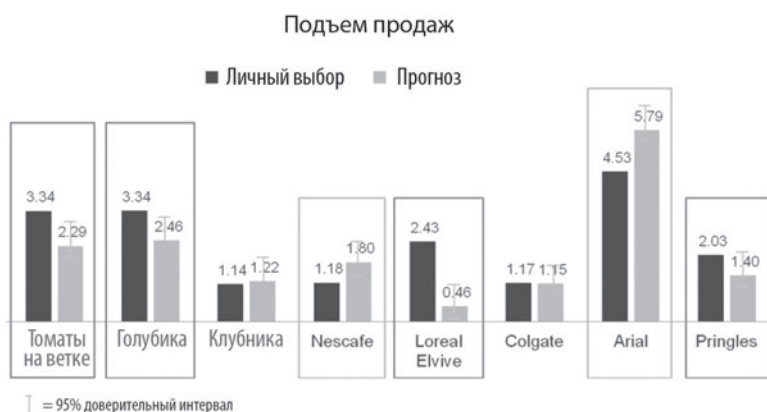
деталей и тонкостей, связанных с личным вкусом и образом мыслей человека (деревьями).

Этот феномен превосходно иллюстрируют данные, полученные нами в ходе следующего эксперимента. Первую группу людей мы отправили за виртуальными покупками, а вторую попросили спрогнозировать, что купят эти люди. Если точнее, задача была в том, чтобы попытаться предсказать выбор своего близкого друга или члена семьи.

В ходе исследования мы демонстрировали участникам различные выгодные предложения товаров и выясняли, какие из них стимулируют наибольший всплеск объемов продаж. Затем мы сравнили результаты тестовых покупок и прогнозов чужого выбора.

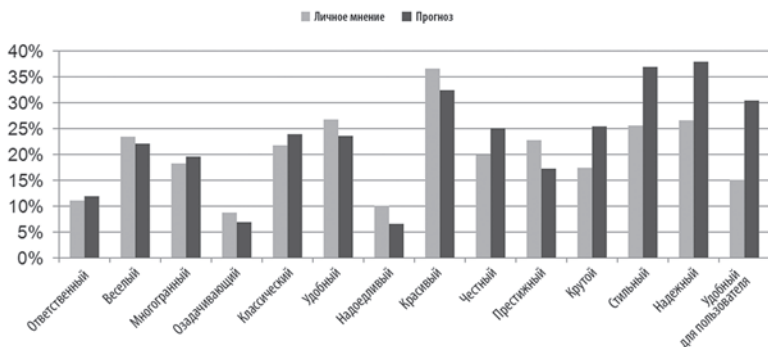
Тестовые покупки и прогнозируемый выбор совпали в девяти случаях из десяти, однако ситуация существенно менялась при пред-

ложении скидки. В группе, участники которой совершали покупки, наблюдались существенные подъемы продаж при наличии скидки на более дорогие товары, обычно приобретаемые, чтобы себя побаловать. В то же время респонденты из группы прогнозирования чаще реагировали на выгодные предложения, сделанные крупными брендами.



Если у вас есть свободное время, можете поразмышлять над тем, почему результаты этого исследования оказались именно такими. Ясно одно: мы не способны спрогнозировать эмоциональные решения других людей, потому очень полезно видеть совместно ответы на оба вопроса. Такая дифференциация окажется ценной для многих исследований.

Подобные различия обычно проявляются при детальном сравнении личной точки зрения с прогнозом. Ниже приведены результаты другого эксперимента, в ходе которого мы просили участников выбрать слова, ассоциирующиеся у них с определенными брендами мобильных телефонов. Затем мы сравнили их ответы с ответами респондентов из другой группы, которых просили спрогнозировать, какие слова выберут другие люди.



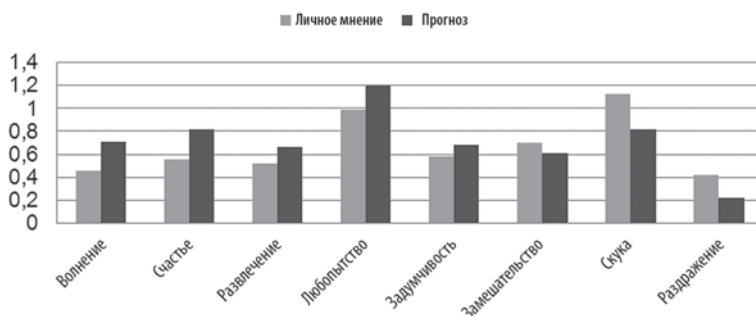
Во многих случаях результаты оказались довольно близки, однако те слова, которые были выбраны для прогноза, в большей степени взяты из «языка маркетинга», то есть во многом навязаны нам специалистами по продажам: крутой, стильный, удобный для пользователя...

Выражая личное мнение, мы склонны высказываться более негативно

Мы попросили респондентов из первой группы выбрать слова, ассоциирующиеся у них с определенными рекламными текстами, а из второй — спрогнозировать, какие слова будут выбраны другими. Выяснилось, что, выражая личное мнение, люди склонны высказываться гораздо более негативно. В первом случае чаще фигурировали слова «скука» и «раздражение», однако при прогнозировании участники исследования чаще отдавали предпочтение положительно окрашенным «любопытству» и «волнению».

(Это может стать нам всем хорошим уроком на тот случай, если придется прогнозировать мысли людей о том, что мы делаем.)

Слова, выбранные для описания рекламы



Часть 2: Как работать с толпой

Во второй части нашей статьи будут рассмотрены лучшие подходы к получению надежных прогнозов от «толпы».

Три главных проблемы при разработке прогнозного опроса: как мотивировать, как передать необходимые знания и как оценить обратную связь, предоставленную толпой. Все начинается с того, как правильно сформулировать вопрос для получения прогноза.

Как правильно сформулировать вопрос для получения прогноза

Секрет правильной формулировки прогнозного вопроса состоит в том, чтобы представить задачу объективно, обезличив ее.

Задайте толпе глупый вопрос, и вы получите глупый ответ

Проблема качества прогнозов, полученных при ответе на множество вопросов, задаваемых в рамках традиционных маркетинговых исследований, в действительности сводится к тому, как они сформулированы. Вопросы о том, кто победит на выборах, уберут ли респонденты помещение после собрания, приобретут ли потребители те или иные товары, как мы видим, в той или иной форме приводят к эмоциональному, когнитивному или системному искажению результатов.

Трансформируйте вопрос в задачу, которую вы просите решить участников исследования

Многие из основных прогностических искажений можно устранить, просто переформулировав вопрос. Вместо того чтобы спросить респондента о чем-либо, поставьте задачу, которую ему необходимо решить.

Вы можете спросить: «Орел или решка?» Возможна и другая формулировка: «Если подбросить монету 100 раз, сколько из них, по Вашему мнению, она упадет орлом?». Теперь это выглядит как задача, решить которую может каждый. Когнитивное искажение, возникающее при прогнозе уборки помещения после собрания, также можно устранить, сформулировав вопрос как задачу: «Спрогнозируйте, сколько человек из присутствующих на этом собрании останутся для уборки помещения?»

Это решение выглядит вполне очевидным. Однако если вы просмотрите множество анкет, разработанных разными людьми для разных целей, то увидите, как редко этот прием используется на практике. Секрет в том, чтобы смотреть на респондентов не как на носителей информации, а как на группу объединенных в сеть интеллектуалов.

Совет: Никогда не задавайте вопрос, не подумав о том, какую задачу вы пытаетесь решить с его помощью. Какую задачу позволяет решить вопрос: «Как бы Вы оценили этот бренд?»

Информация — это важно

Меняя формулировку вопроса с «Вы купите этот товар?» на «Спрогнозируйте, купят ли этот товар потребители?», вы потенциально можете устранить личные когнитивные искажения, однако при этом не восполните нехватку информации, необходимой для составления подобного прогноза. Повысить точность прогноза помогут дополнительные сведения. Для кого предназначен этот товар? Что эти люди покупают в настоящее время? Сколько этот товар будет стоить в сравнении с другими? Что отличает его от прочих или почему он лучше, чем конкурирующие предложения? и т.д. Без этой

информации мы вынуждены полагаться исключительно на свою интуицию, которая может как помочь, так и помешать нам сделать точный прогноз.

Взгляните на этот пример из проведенного нами эксперимента. Мы предложили опрашиваемым угадать, какая из этих четырех кружек продается лучше всего.

Угадайте, какая из этих кружек продается лучше всего?



Респонденты предположили, что лучше всего должна продаваться синяя кружка с рисунками, посвященными отдыху на природе. Этот прогноз оказался далек от истины. Анализ результатов продаж производителя показал, что синяя кружка продается хуже всех, а лучше всех — зеленая, с рисунками на тему садоводства, и красная, с изображениями рабочих инструментов. Мы решили выяснить, почему так происходит, и провели второй эксперимент. На этот раз респондентам был задан вопрос, какую кружку они бы купили для себя. Однако и в этом случае синяя кружка оказалась на первом месте.

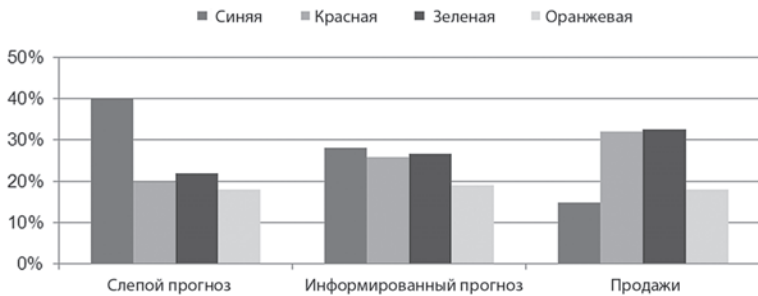
Чем же вызвана эта ошибка? Мы поговорили с человеком, занимающимся продажей таких кружек, и выяснили, в чем причина. Такие кружки чаще всего покупают или как подарок друзьям, которые любят повозиться в саду, или мужчинам в качестве благодарности за проделанные ремонтные работы по дому.

Выбирая ответ на вопрос, мало кто подумал о таком контексте покупки.

Поэтому мы провели третий эксперимент, добавив эту информацию в конце вопроса: «Примите, пожалуйста, во внимание тот факт, что такие кружки часто покупают в качестве подарка».

На нижеприведенной диаграмме вы можете видеть результат эксперимента. После добавления примечания прогноз участников

по-прежнему не был идеальным, однако он очевидно стал ближе к действительности, так как участники исследования задумались о том, какие из этих кружек стали бы хорошим подарком. Не зная об этом, они основывали свои суждения в основном на собственных чувствах. Информация помогла направить респондентов в сторону более объективного прогноза.



Хочется надеяться, что этот пример станет убедительной иллюстрацией утверждения, что без доступа к релевантной информации точность прогноза может существенно снизиться, а сама информация является ключевым условием получения достоверных прогнозов. Кроме того, участвуя в опросах, людям приходится думать гораздо больше обычного, что подчеркивает значимость вознаграждения, которое им за это предлагается.

Найти «мудрость» в толпе и поделиться со всеми

Предположим, вы тестируете новый товар, и у вас нет никакой информации, которая помогла бы определить, насколько этот товар будет популярен. Если нет проверенных фактов и доказательств, просветить толпу помогут содержательное обсуждение и дискуссии. В этом фактически заключена суть фокус-групп. Перекрестное общение участников друг с другом помогает всей группе получить полезные сведения и вынести более точные коллективные суждения.

Поэтому многие из наиболее эффективных методов принятия решений основаны на обсуждении и дискуссии. Именно так, к при-

меру, принимаются решения в парламенте и проводятся выборы. (Можно рассматривать выборы как коллективный прогноз жителей страны о том, кто сможет управлять страной наиболее эффективно). Представьте себе членов правления крупной компании, принимающих решения в полном молчании. Это невозможно! Все они активно участвуют в обсуждении, делятся своими мыслями до голосования.

Однако часто при проведении количественных исследований мы лишены возможности организовать подобный обмен информацией между испытуемыми. Это можно назвать одним из реальных недостатков традиционных маркетинговых исследований в той части, которая связана с получением точных прогнозов.

Необходимо побудить толпу сгенерировать сведения, важные для точности прогнозов

Точность большинства прогнозов страдает из-за того, что каждый из респондентов размышляет над ответом самостоятельно. В большой группе знания часто распределены крайне неравномерно. Отдельные респонденты могут обладать множеством полезных и никому не известных сведений, которые могут быть вытащены наружу в процессе агрегации данных. Представьте, насколько более эффективным стал бы процесс прогнозирования, если бы эти люди поделились своей информацией с другими участниками опроса. Распространение сведений — возможность сделать толпу «мудрее».

По изложенной выше причине потенциально весьма значимой представляется роль модератора, вовлеченного в процесс выработки точных прогнозов, который мог бы стимулировать обсуждение. Это еще один важный фактор, о котором мы поговорим ниже, когда перейдем к обсуждению рынков предсказаний.

Диалектическая перезагрузка

Херцог и Хертвиг [8] провели серьезные научные изыскания, изучая возможность улучшить прогноз толпы за счет аргументированных споров и самообразования. Указанные авторы продемонстрировали, что эти инструменты способны улучшить качество

сделанных прогнозов. Данный процесс получил прекрасное название — «диалектическая перезагрузка». По сути, имеется в виду самообразование в процессе диалога.

О важности мотивации и вознаграждения в форме обратной связи

Деля прогностические исследования, мы просим респондентов подумать. Мы хотим, чтобы они максимально тщательно обдумали предложенную им проблему или вопрос. Ключевую роль здесь играет мотивация. Один из важнейших инструментов нашего арсенала, способный мотивировать человека на составление точного прогноза,— последующее информирование его о том, оказался ли он прав в своем предположении. Нижеприведенный эксперимент может послужить превосходным доказательством этого тезиса.

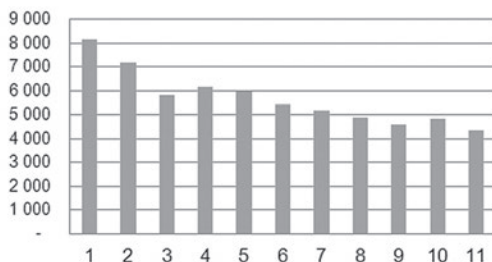
Мы разработали способ измерить знание бренда в форме прогноза-викторины с возможностью выбора нескольких вариантов ответа. Респондент играет в игру, угадывая бренд по визуальной подсказке. Чтобы доказать выдвинутое выше предположение, одной группе эти вопросы были заданы без предоставления какой-либо обратной связи по точности ответов. Во второй группе после каждого вопроса сообщалось, правильный или неправильный был получен ответ. Сам же опрос был построен в форме соревнования. Для усиления фактора состязательности мы также сообщили респондентам, что от скорости их ответов зависит, сколько очков они наберут в итоге.



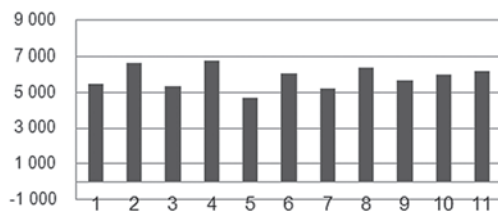
На диаграмме видно, что испытуемым, которые не получали обратной связи по своим ответам, выполнять предложенное задание надоело, время обдумывания каждого последующего вопроса постепенно сокращалось, и это отразилось на качестве собранной

информации. Те, кто получал обратную связь и дополнительную мотивацию быстрее принимать решения, потратили немного больше времени на обдумывание и отвечали точнее.

Стандартный опрос без обратной связи (время ответа в миллисекундах)



Геймифицированный опрос: ограничение по времени + обратная связь (время ответа в миллисекундах)



Доля правильных ответов



Примечание: временные ограничения способствуют проявлению интуитивных реакций системы № 1

В некоторых прогнозах, например, связанных с оценкой рекламы или узнаваемостью бренда, особо ценными оказываются быстрые интуитивные предположения. Высказывая их, респонденты используют мыслительные процессы своей системы № 1. Вышеприведенный пример такой методики предполагает вознаграждение респондентов за более быстрые ответы путем присуждения большего количества очков или же постановку задачи в форме соревнования: сколько ответов успеет дать участник за ограниченное время, например, за 30 секунд. По нашим данным, обе эти методики очень эффективны.

Геймификация всего процесса составления прогноза

Мы использовали этот прием во многих опросах. Простая фраза «приняв участие в следующем опросе, вы сможете узнать, насколько хорошо справились с этим» оказалась невероятно эффективной для получения серии не связанных друг с другом прогнозов от группы респондентов, а также для поддержания их высокой мотивации на тщательное обдумывание своих ответов. Именно эта методика была использована нами для получения многих прогнозов, представленных в этой работе. Мы называли ее «surveytainment» («опрос-развлечение»).



Если вы хотите узнать больше об этой методике, просмотрите презентацию, представленную на мероприятии Amsterdam ИЕХ.

Она доступна по следующей ссылке: <http://www.insightinnovation.org/2014/04/23/surveytainment-how-can-you-make-surveys-more-addictive/>

Повышение уровня коллективной мудрости путем тщательной оценки индивидуальных прогнозов

Как известно, Нейт Силвер доказал: если пойти дальше и не просто агрегировать прогнозы толпы, но оценивать индивидуальные источники прогнозов по степени их надежности в прошлом, в результате можно получить гораздо более точные прогнозы. Например, прогнозируя результаты выборов, он собрал целый ряд прогнозов, сделанных другими специалистами в этой области, проанализировал, насколько надежными оказался каждый из них, а затем оценил каждый прогноз, основываясь на точности предположений данного специалиста в прошлом. И все! Таким образом, ему удавалось делать превосходные прогнозы, гораздо более точные по сравнению с результатами любого отдельно взятого прогноза. Это было извлечение коллективной мудрости в чистом виде. По его собственным оценкам, результаты оказались точнее на 20%.

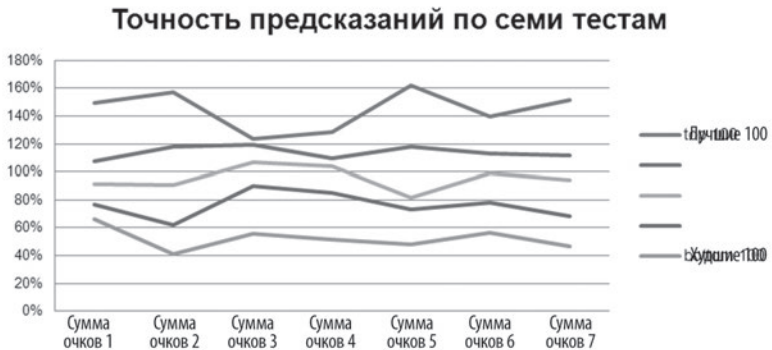
В научной среде на рынках предсказаний эта методика была известна уже довольно давно, на данную тему было написано несколько работ, одна-две из которых приведены в нижеприведенном списке цитируемой литературы. Одна из наиболее впечатляющих статей была опубликована совсем недавно под названием «Мудрость небольших умных толп» («The Wisdom of Smaller, Smarter Crowds») (GOLDSTEIN, MCAFEE & SIDDHARTH SURI⁶). В этой работе рассматриваются тысячи прогнозов результатов игр выдуманной футбольной лиги. Эта история наглядно показывает, что уровень коллективной мудрости можно существенно повысить, проанализировав прошлые успехи в прогнозировании и собрав прогнозы, сделанные испытуемыми, которые в прошлом проявили себя наиболее успешными предсказателями.

⁶ <http://dl.acm.org/citation.cfm?doid=2600057.2602886>

Теоретически мы можем взять эту методику для повышения точности прогнозов и в рамках традиционных маркетинговых исследований. Однако для этого нам необходимо выявить, кто из респондентов хорош в прогнозах, а кто — нет.

Выявление хороших прогнозистов

Мы приступили к решению этой задачи с вопроса: «Существуют ли вообще хорошие прогнозисты в таких исследованиях, которыми мы занимаемся?». Была проведена серия еженедельных прогностических экспериментов с одной и той же группой респондентов. Это позволило нам проследить их успехи в прогнозировании на протяжении всего этапа исследования. В итоге мы четко разделили респондентов на отдельные группы по качеству сделанных ими прогнозов. При этом лучшая из этих групп неизменно показывала более высокие результаты в сравнении с остальными.



Как определить хорошего прогнозиста?

Мы начали решать задачу по выявлению хороших и плохих прогнозистов, используя личностные тесты. Мы планировали отбирать лучших вручную или взвешивать индивидуальные прогнозы. Несомненно, Филипу Тетлоку удалось определить тип личности, способной делать более точные долгосрочные политические и эко-

номические прогнозы⁷. Он называет их «лисами». Это люди, мыслящие более объективно, они более гибкие по сравнению с «ежами», которые откровеннее в выражении своего мнения и подвержены влиянию собственных непоколебимых убеждений. Мы задались вопросом, можно ли категоризировать своих панелистов подобным образом?

Сразу же огорчу вас известием: ответ был получен отрицательный. В рамках пары экспериментов, проведенных за это время, нам не удалось категоризировать панелистов по типам личности. Это не означает, что задача не может быть решена в принципе, однако до сих пор мы так и не смогли прийти к приемлемому практическому решению.

В результате мы выяснили, что мужчины в целом делают немного более точные прогнозы, чем женщины — различие составляет около 10%. Мы бы определенно не хотели, чтобы эти данные цитировались где-либо без указания контекста. Проведенные нами эксперименты предусматривали составление прогнозов по ограниченному набору тем, которые, возможно, ближе мужчинам, чем женщинам. Кроме того, как мы выяснили, способности к точному прогнозированию варьируются в зависимости от предложенной темы. Так, например, женщины показывают более высокие результаты, когда речь идет о прогнозировании стоимости товаров массового потребления.



⁷ Источник: <http://press.princeton.edu/titles/7959.html>

Также мы выяснили, что быстрые и интуитивные ответы респондентов оказываются точнее, например, при оценке рекламы. Однако для многих сложных типов прогнозов увеличение времени на обдумывание приводит к получению гораздо более точных прогнозов. В качестве примера в данном случае можно привести предположения о поведении и мнениях других людей.

Составление правильных прогнозов требует применения целого набора различных интеллектуальных процедур. Мы часто используем эвристический прогноз как более легкий вариант в ситуациях, когда вопрос представляется слишком сложным для аналитической оценки. Это явление подробно рассмотрено в работе Даниэля Канемана. Последняя написана по результатам экспериментов в области биохевиористской экономики, сводимым к описанию роли и важности системы № 1 и системы № 2, отвечающих за интуитивные и рациональные суждения, а также к зависимости от объема информации, необходимой для правильной оценки. Например, в ней указано, что при угадывании стоимости биты и мяча⁸ система № 2 — рациональные вычисления — может показать лучшие результаты, чем система № 1 — интуитивное угадывание. Однако в другой ситуации, когда слишком сложно провести какие-либо вычисления, система № 1 становится более полезной. К примеру, если надо угадать количество монеток в банке.

Для изучения этого явления мы провели эксперимент и, оценив эффективность различных способов мышления отдельных участников, выяснили, насколько лучше или хуже они справляются с заданиями по прогнозированию того, какие товары будут покупать потребители. Мы попросили наших респондентов пройти стандартный тест для оценки персональной эффективности, являющийся одним из способов определения стиля мышления, и проанализировали, соответствуют ли высокие или низкие оценки по этому тесту более высокой точности сделанных прогнозов. К сожалению, этот

⁸ Известный эксперимент, описанный в отмеченной наградами работе ESOMAR под авторством Стивена Филиппа и Абилай Хилл «Исследования в мире иррациональных ожиданий» («RESEARCH IN A WORLD OF IRRATIONAL EXPECTATION»)

эксперимент не выявил сколько-нибудь существенных различий, вероятно, в силу ограниченности поставленной задачи. Поэтому для более глубокого понимания этого явления представляется необходимым провести более масштабную исследовательскую работу.

Уверенность в собственных прогнозах

Мы выяснили, однако, что хороший способ выявить успешных прогнозистов все-таки существует. Достаточно просто спросить человека, насколько он сам уверен в сделанном им прогнозе! Наиболее уверенные в себе участники исследования в итоге оказались на треть точнее в своих прогнозах.



Единственная проблема, связанная с использованием данного подхода, состоит в том, что полную уверенность в своих ответах обычно демонстрируют менее 10% респондентов в выборке. Поэтому на практике этот метод не слишком полезен.

Мы, впрочем, нашли более удачный и практичный способ оценки уверенности в своих прогнозах. Необходимо спросить респондента, сколько он готов поставить на свой ответ в виртуальном пари, которое мы уже упоминали в начале этой работы. Мы провели эксперимент, в рамках которого предоставили испытуемым определенную сумму денег, попросив их делать ставки по своим оценкам популярности или непопулярности демонстрируемой рекламы. Получив

данные по более чем 100 рекламным объявлениям, мы провели первичный анализ взаимозависимости между тем, сколько ставят участники, и точностью их прогнозов.

Полученные в итоге результаты оказались довольно интересными, с ними можно ознакомиться на нижеприведенной диаграмме. Мы вынуждали участников сделать хотя бы минимальную ставку, поэтому люди, не слишком уверенные в своем ответе, указывали минимально возможные суммы. Как можно заметить на диаграмме, их прогнозы гораздо менее точные, чем у тех, кто ставил большие виртуальные суммы. Любопытно, что небольшие отрицательные ставки дают наиболее высокую корреляцию, а слишком высокие ставки соответствуют менее точным результатам, чем умеренные.



Ценность этих выводов станет понятна несколько позже, когда мы перейдем к обсуждению рынков предсказаний.

Какие методы прогнозирования можно назвать лучшими?

Следующий важный вопрос — какой метод получения прогнозов можно назвать лучшим? Чтобы ответить на него, отправимся в мир «рынков предсказаний».

Бесспорно, рынки предсказаний находятся на переднем крае технологий прогнозирования. Ставшие известными благодаря Iowa

Electronic Markets⁹, они представляют собой, по сути, азартную игру в прогнозы, посвященные различным событиям.

Представление о методах рынков предсказаний в традиционных исследованиях

Попробуем объяснить суть этих методов для тех, кто с ними не знаком. Мы склонны использовать термин «рынки предсказаний» без какой-либо конкретизации, хотя между методами, применяемыми на рынках различных типов, существуют определенные отличия, которые стоит прояснить.

Основы рынков предсказаний в традиционных маркетинговых исследованиях. Рынки предсказаний, методы которых впервые были опробованы в исследованиях Brainjuicer и другими компаниями, представляют собой процесс, в рамках которого предлагается купить или продать виртуальные «акции» различных идей или целых пакетов идей. Затем чистое количество покупок или продаж акций по каждой идее сравнивается с общим объемом торговой активности. Подобный подход к процессу принятия решений показал убедительные результаты в плане усиления дифференциации предложенных вариантов, чем при использовании стандартных подходов к рейтингованию (источник: работы Brainjuicer, посвященные рынкам предсказаний).

Геймифицированные рынки предсказаний. Первый метод можно расширить, предоставив участникам исследования виртуальную сумму денег для ведения торговли и возможность решать, сколько они готовы инвестировать в каждую из предложенных идей. Затем процесс можно геймифицировать, сообщив по окончании торгов, сколько денег они в итоге выиграли или проиграли. Элемент состязательности, как мы уже продемонстрировали ранее, оказался крайне эффективным средством повышения активности участников при выполнении заданий.

⁹ http://en.wikipedia.org/wiki/Iowa_Electronic_Markets

Торги на рынке предсказаний. Ключевое различие между стандартными вопросами рынков предсказаний и правильно организованными торгами состоит в следующем. Во-первых, в процессе торгов обычно используются реальные деньги, которые можно получить за правильные прогнозы; во-вторых, стоимость представленных для покупки акций повышается или понижается в зависимости от того, сколько человек выбирают тот или иной вариант — так же, как на фондовой бирже. Если участники «инвестируют» в какую-либо идею, ее ценность возрастает. Важность этого процесса в том, что он помогает автоматически корректировать прогнозы. Например, если вы видите, что ставки на «орла» и «решку» распределились в соотношении 70:30 в пользу решки, вы получите стимул поставить на орла. Таким образом, рынок приводит ставки к соотношению 50:50.

Открытые и темные рынки. Некоторые рынки предсказаний предоставляют своим участникам возможность наблюдать за тенденциями в изменении стоимости различных предложений, свободно читать комментарии, обсуждения и представленные сведения, которые помогают им принимать более информированные решения, что — как мы объясняли выше — крайне важно. Участников стимулируют комментировать сделанный выбор.

Если пользователи не могут видеть покупки других, такой рынок называется «темным». Следовательно, методы рынка предсказаний, используемые в традиционных маркетинговых исследованиях, также можно описать как методы «темного рынка» в отличие от методов «открытой» торговли на рынке предсказаний.

Существует несколько проблем, связанных с особенностями торговли на открытых рынках. Очевидно влияние мнений отдельных участников на тенденции рынка, этот феномен известен под названием стадного эффекта. Как мы уже отмечали ранее, когда участникам приходится делать сложные прогнозы в условиях ограниченной исходной информации, они склонны больше полагаться на мнение других людей. А это может привести к тому, что мнение нескольких человек возобладает над мнением многих других, которые предпочтут к присоединиться к первым. Назовем это кумулятивным внушением: мнение одного человека влияет на мнение другого, и это,

в свою очередь, провоцирует целый каскад изменений мнений последующих участников исследования.

Все дело в деньгах!

Важное различие между традиционными исследованиями с элементами рынков предсказаний и реальными рынками предсказаний состоит в вознаграждении. Когда я принимаю участие в опросе, мне совершенно не важно, какой именно ответ я дам на тот или иной вопрос. Я получаю вознаграждение просто за заполнение опросной анкеты. Из-за этого, как мы уже продемонстрировали ранее, прогнозы некоторых типов (например: «Кто, по вашему мнению, сформирует следующее правительство») становятся практически бессмысленными. Мы обнаружили, что изменение системы вознаграждения путем геймификации процесса может резко изменить степень приложения усилий для ответа на вопросы, а также качество этих ответов. Однако при использовании методов рынков предсказаний можно пойти еще дальше и предложить реальное финансовое вознаграждение — не только собственно за участие в исследовании, но и за точность сделанных прогнозов.

Торги на рынке предсказаний vs правила составления прогнозов в рамках традиционных маркетинговых исследований. В чем разница?

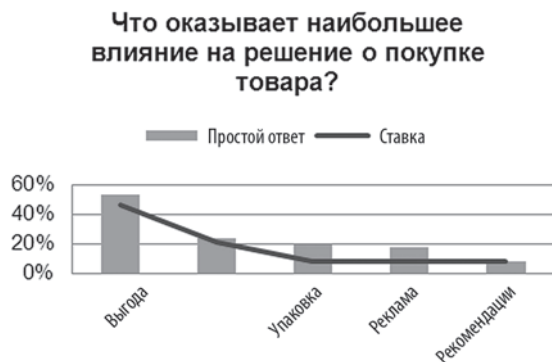
Мы решили составить более полное представление о сравнительных преимуществах рынков предсказаний при составлении прогнозов перед традиционными методами проведения маркетинговых исследований.

Для этого была проведена уникальная серия экспериментов в сотрудничестве с компанией Prediki, предоставляющей открытую платформу для ведения торговли в формате рынка предсказаний.

Мы начали с небольших экспериментов, в ходе которых просили опрашиваемых сделать несколько прогнозов. Затем мы перенаправляли их на рынок предсказаний, где просили сделать тот же прогноз за вознаграждение. При этом мы протестировали целый ряд различ-

ных тем — от прогнозов по доле рынка различных товаров до причин приобретения того или иного товара.

Полученные данные позволяют сделать вывод, что распределение ответов в обоих случаях получилось одинаковым, это можно увидеть на приведенной ниже диаграмме.

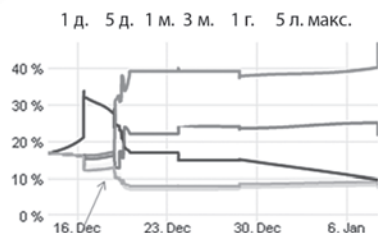


Однако мы также отметили несколько важных исключений. Например, мы провели семь тестов рекламы в формате прямого сравнения. В пяти случаях оба результата совпали, однако в двух других обращение к рынку предсказаний привело к различающимся ответам.

Единственное существенное различие между двумя подходами заключалось во времени и усилиях, затраченных на составление прогнозов, а также в числе участников, вовлеченных в процесс. В опросе обдумывание занимало в среднем около трех секунд, а всего в прогнозировании приняли участие более 150 человек. В исследовании по методу рынка предсказаний оказалось гораздо меньше участников — в среднем около 30 человек, однако довольно часто они несколько раз возвращались к заданиям, пересматривая свою ставку и улучшая прогноз. Средний респондент пересмотрел свои прогнозы трижды, основываясь на изменении ситуации на рынке и полученных сообщениях, и очень быстро реагировал на любую дополнительную информацию, которую мы предоставляли (см. пример ниже).

Реакция участников на новую информацию

Изменение стоимости (операции)



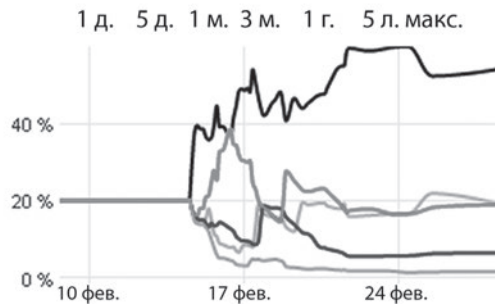
Момент представления новой информации, которая привела к незамедлительному изменению ситуации на рынке (мы добавили несколько комментариев)

Варианты ответа

< > WIDGET

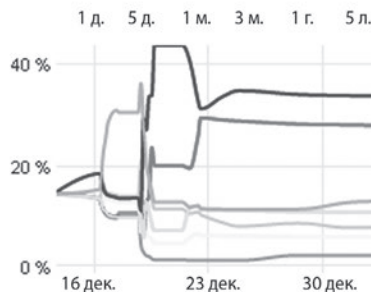
Аромат	46.79 %
Выгода	21.41 %
Другая причина	8.49 %
Дизайн упаковки	7.83 %
Реклама	7.81 %
Рекомендации	7.66 %

Выбор участников торгов на рынке предсказаний в большинстве случаев менялся очень быстро (смотри приведенную ниже диаграмму).



В первые несколько дней торгов некоторые варианты ответа были подвержены довольно существенным колебаниям показателей и вызывали ожесточенные споры (см. второй пример). После достижения консенсуса между участниками устойчивость сохранялась до тех пор, пока мы не добавляли больше информации. После этого мнения зачастую менялись довольно быстро. Это, скорее всего, можно объяснить тем, что мы задавали вопросы о брендах. Отличие, например, политических рынков предсказаний в том, что на них каждый день могут возникать новые внешние факторы, влияющие на ситуацию.

**Пример оживленных споров,
вызванных представленными колебаниями на рынке:**





Чтобы получить более глубокие результаты, мы решили провести масштабный эксперимент, в ходе которого попросили участников сделать ряд прогнозов в рамках традиционного маркетингового исследования (онлайн-опроса), а затем, как и в прошлый раз, перейти на платформу рынка предсказаний. Однако на этот раз мы привлекли к участию несколько ритейлеров. Респондентам предлагалось предсказать, у какого из представленных потребительских товаров объем продаж оказался самым высоким. То есть мы могли сравнить прогнозы участников исследования с реальными данными. Для сравнения мы также проанализировали статистическую надежность сделанных прогнозов на сопоставимых выборках.

Далее приведены полученные нами результаты.

Чудесное преимущество рынка предсказаний на первом этапе исследования

Используя метод прямого сравнения, мы проанализировали точность ответов на микровыборках из пятнадцати случайным образом отобранных респондентов, сделавших прогнозы в рамках традиционных опросов. Мы сопоставили их с прогнозами первых пятнадцати участников торгов в ходе 32 отдельных экспериментов с использованием рынков предсказаний. Причина, по которой мы выбрали такие маленькие выборки, состоит в том, что многие из наиболее успешных рынков предсказаний обычно работают с небольшим количеством пользователей. Так, в случае известного рынка Iowa Electronic Market многие прогнозы были сделаны группами, состоящими менее чем из двадцати человек.

Средняя точность прогнозов пятнадцати случайно отобранных респондентов¹⁰ в традиционном опросе составила 55%. При этом существовала 37%-я вероятность, что эти предсказания окажутся правильными при выборе ответов наугад. Средняя точность прогнозов первых пятнадцати участников торгов на рынках предсказаний составила 65%.

В данной ситуации проявилось неотъемлемое преимущество метода рынка предсказаний: люди, наиболее уверенные в правильности своего мнения, ставили на него больше, и это помогало им объективно оценить собственный ответ. Размер вашей ставки является показателем вашей уверенности; как мы уже видели ранее, уверенность напрямую соотносится с точностью прогнозов, стремлением участников приложить больше усилий для подготовки ответа, тщательнее его обдумать.

¹⁰ Примечание: для этого мы случайным образом выбрали пятнадцать человек из выборки в 400 участников 20 000 раз, анализируя среднюю точность прогнозов в каждой из групп.

Преимущество традиционных маркетинговых исследований на втором этапе эксперимента

Однако когда мы начали анализировать данные по более крупным выборкам и большему количеству торгов, показатели точности прогнозов на рынках предсказаний начали отставать от показателей традиционных маркетинговых исследований. После подключения большей выборки точность предсказаний в рамках традиционных маркетинговых исследований выросла, достигнув показателя в 75% верных прогнозов.

Мы обнаружили, что с того момента, как число участников исследования становится более двадцати человек, и в отсутствие какой-либо дополнительной информации на рынках предсказаний возникает тенденция к закреплению результатов на определенных показателях. Расширение выборки уже не приводило к их изменению, поэтому точность прогнозов не улучшалась. Соответствующий показатель так и остался ниже 70%.

Это можно считать непосредственным примером воздействия стадного эффекта. В своих экспериментах мы не применяли методы темных рынков, поэтому то, что мы увидели, стало результатом влияния мнения определенной части людей на других, более «слабых» прогнозистов.

Стадный инстинкт вот уже несколько лет называют одним из главных недостатков рынков предсказаний. В большинстве этих случаев у участников было очень мало информации, на которую они могли бы опереться в своем прогнозе, помимо собственных внутренних убеждений. Поэтому наиболее ценными данными, на которые приходилось полагаться, было мнение остальной части толпы. После того, как коллективное мнение сформировано, изменить его крайне сложно. Даже при том, что можно было делать ставки вразрез с тенденциями на рынке, при достижении определенной суммы от игрока требовались смелость и азарт, чтобы поставить на непопулярный вариант.

Увеличение объема доступной информации

Из научных работ, посвященных рынкам предсказаний, известно, что повышение объема изначально доступной информации и введение правил темного рынка на начальном этапе работы может смягчить действие стадного эффекта. Побуждение большего числа игроков объяснять свои решения с помощью активного модерирования обсуждений помогает приходить к более «рациональным» решениям.

Преимущество рынков предсказаний на третьем этапе эксперимента

Для подтверждения важности увеличения объема информации, доступной на рынках предсказаний, мы провели отдельную серию из шестнадцати экспериментов на рынках, в которых с самого начала представили участникам реальные комментарии и мнения, собранные нами в ходе традиционного маркетингового исследования, побуждая респондентов к активному обсуждению своих решений.

Это действительно помогло повысить точность прогнозов. В ходе шестнадцати тестов с использованием вышеупомянутых микровыборок из пятнадцати человек мы получили тринадцать верных прогнозов, что соответствует общей точности прогнозирования на уровне 80%. Это позволяет рынкам обогнать традиционные маркетинговые исследования по обсуждаемому показателю.

Понятно, что это лишь один простой эксперимент, однако он несомненно подтверждает предположение о том, что информация помогает толпе принимать более рациональные решения. Для тех, кто уверен в своем мнении, это может послужить стимулом к использованию предоставленных сведений и их тщательному обдумыванию для получения выгоды. Например, при тестировании кружек респонденты смогли угадать, что лучше всего из них продается зеленая. При этом были отмечены такие комментарии: «Садоводство — популярное времяпрепровождение, и кружка с садовыми инструментами может стать хорошим подарком».

Общий анализ

Результаты

Варианты ответа

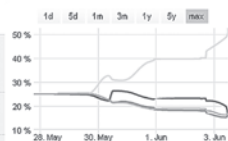
Это задание на прогнозирование в рамках опроса GMI Prediction Survey 2014. Все свои сообщения и комментарии по всем приведенным вопросам направляйте, пожалуйста, по ссылке: GMI Prediction Survey 2014 C.

На изображениях представлены обсуждаемые товары, а ниже приведены комментарии участников к ним.

Синяя кружка	Оранжевая кружка	Зеленая кружка	Красная кружка
Синий цвет популярный	Потому что она оформлена в кухонной тематике	Прелестная	Потому что она ярко красная, привлекает взгляд
Синий кажется более утонченным	Она ярче	Привлекут рисунки садовой тематики	Яркая и эффектная
Отдых на природе очень популярен	Появился цвет	Садоводство нравится многим	Чем-то понравится большинству
Притягивает взгляд	Кухонные принадлежности наводят на мысли о чае	Тема садоводства пользуется популярностью у многих	Такую купят мужчины

Зеленая кружка	52.54 %
Оранжевая кружка	16.17 %
Синяя кружка	16.02 %
Красная кружка	15.26 %

Изменение стоимости (Операции)



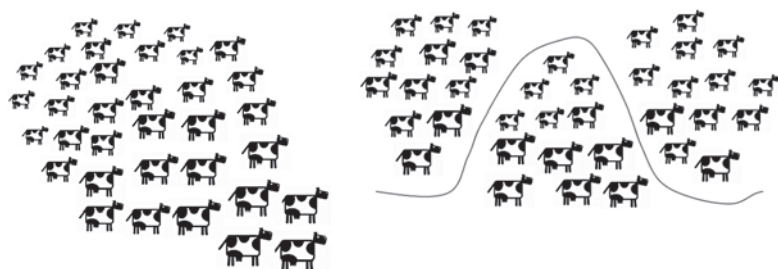
Лучшие участники (Рейтинг)

	US\$
1. Claire Breeds	2.81
2. Denise Butler	1.79

Хотя добавление информации, несомненно, улучшило ситуацию, мы по-прежнему не были удовлетворены показателем точности прогнозов на уровне 80%. По нашим предположениям, стадный эффект все так же оказывал довольно значимое влияние. Мы хотели выяснить, есть ли способ добиться дальнейшего повышения уровня точности прогнозов.

Разделим стадо!

В итоге мы пришли к очень простой идее, которую считаем очень эффективным решением для дальнейшего повышения надежности прогнозов,— концепции разделения стада. Что произойдет, если вместо работы с единым рынком предсказаний мы разделим «стадо» на несколько аналогичных рынков, работающих независимо друг от друга, а затем соберем с них данные? Эта методика практически с самого начала успешно применялась в качественных исследованиях.



Для проверки этой идеи мы создали три отдельных рынка предсказаний, на каждом из которых участники составляли шестнадцать одинаковых прогнозов независимо друг от друга. В конце мы собирали данные об их коллективном мнении. Восемь из шестнадцати прогнозов оказались идентичными на всех трех рынках, в семи случаях мнения разделились, а еще в одном все три рынка дали разные ответы. Однако собрав все данные, мы оценили каждый прогноз по степени уверенности участников каждого рынка и обнаружили, что пятнадцать из шестнадцати прогнозов оказались верными, что равно 90%-ному уровню точности.



И вновь необходимо подчеркнуть, что это был лишь небольшой единичный эксперимент, и его следует рассматривать как частный случай, не доказывающий потенциальной ценности применения метода разделения «стада».

Кроме того, следует отметить для всех скептически настроенных специалистов, что метод разделения «стада» обладает встроенным механизмом защиты от ошибочных прогнозов. Если вы проводите опрос на трех рынках и все они прогнозируют одно и то же, существует 88%-я вероятность надежности прогноза. Если же возникают противоречия, вы можете либо агрегировать полученные данные, либо пойти дальше и провести опрос на большем количестве подобных рынков, так как создаются они очень легко и быстро.

Научные труды, посвященные рынкам предсказаний

Данная работа посвящена в основном нашему исследованию, однако рынки предсказаний в течение вот уже многих лет остаются увлекательной темой для основательных научных штудий. Поэтому мы посчитали необходимым в завершение нашей статьи осветить несколько ключевых тезисов из научных работ, опубликованных по некоторым из поднятых нами вопросов.

Объем инвестиций. Служит ли сам факт выделения большого объема инвестиций на рынке предсказаний показателем уверенности прогнозиста в своем ответе, подтверждением наличия у него более высокого уровня знаний, чем у того, кто вкладывает меньшие суммы, и становится ли при этом его прогноз более точным?

Профессор Бернардо Хуберман, научный сотрудник компании HP, проанализировал этот вопрос в 2003 г. Удивительно, но он не обнаружил никаких убедительных доказательств того, что более объемные инвестиции означают наличие у участника эксперимента более высокого уровня знаний или способности к прогнозированию. Однако потом у него возникла другая идея. Некоторым людям нравится рисковать, ставя на кон более высокие суммы денег, чем это делают другие. Сначала он проанализировал поведение респондентов на множестве различных рынков, чтобы получить сведения о склонности каждого из них к риску. Затем он проанализировал полученные данные еще раз, модифицируя индивидуальные ставки в соответствии со склонностью к риску, определенной на первом этапе работы. Сделав это, Хуберман обнаружил существенное улучшение точности прогнозов по сравнению с несовершенными ры-

ночными методами. После исключения выбросов удалось добиться снижения ошибочности прогнозов приблизительно на 20—30%.

Гендерные проблемы. Довольно объемные данные Prediki по результатам торгов демонстрируют, что на любом отдельно взятом рынке предсказаний женщины в среднем инвестируют существенно меньше мужчин. Поэтому специалисты этой компании разработали внутренние алгоритмы оптимизации прогнозов для улучшения точности прогнозов, анализируя модель инвестирования каждого из участников в прошлом. Такие оптимизированные прогнозы можно использовать в качестве дополнительной информации, которая зачастую оказывается полезной при более детальном анализе.

Реальные и игровые деньги. Лучше ли справляются со своей задачей рынки предсказаний, если игроки вынуждены подкреплять свое мнение реальными деньгами? Или же достаточно использовать игровые деньги?

Это важный вопрос, так как проблема нелегальности подобных ставок в некоторых странах — например, в США — не позволит в таком случае использовать рынки предсказаний. Профессор Джастин Уолферс из Школы имени Уортона изучил этот вопрос в 2003 г. В его исследовании читателей газеты USA Today просили принять участие в биржевой игре и спрогнозировать результаты матчей команд Американской футбольной лиги. К удивлению специалистов, ощутимых различий в качестве прогнозов выявлено не было. В некоторых случаях с задачей лучше справились рынки с игровыми деньгами, в других — рынки с реальными деньгами. Однако и те и другие показали гораздо более высокие результаты по сравнению с отдельными экспертами и опросами. Для специалистов по маркетинговым исследованиям это хорошая новость. Метод рынков предсказаний можно использовать исключительно на основе игровых денег, избежав таким образом всех возможных юридических трудностей и географических ограничений.

Важность вознаграждения. Получается, при использовании рынков предсказаний денежное вознаграждение респондентов уже не является необходимым условием? Ведь было бы так здорово, если бы мы могли прибегать к коллективной мудрости совершен-

но бесплатно! Однако на практике мы регулярно видим снижение активности на рынках прогнозов, вызванное тем, что клиент оказывается не в состоянии предоставить респондентам достаточное денежное вознаграждение или привлекательные призы. Для получения качественных результатов рынок предсказаний должен привлекать к участию хорошо осведомленных людей, они должны быть достаточно мотивированы и, по Канеману, использовать при ответе систему № 2. В 2006 г. профессор Стефан Люкнер из Университета Карлсруэ протестировал три возможных варианта: награждение призами разной ценности участников исследования, занявших первое, второе и третье места, награждение призами за определенные достижения и фиксированное денежное вознаграждение, как в традиционных опросах. Неудивительно, что вариант с призами для трех лучших участников позволил получить лучшие результаты по сравнению с фиксированным вознаграждением, ведь респондентам приходилось прикладывать больше усилий для победы. Неожиданным оказалось то, что вариант с присуждением призов за отдельные достижения показал худшие результаты, чем фиксированное вознаграждение. По мнению специалистов из Prediki, причина таких показателей — небольшое, но гарантированное вознаграждение в отсутствие каких-либо действий со стороны участников группы, в которой призы присуждались за достижения. Респонденты просто боялись потерять то, что уже получили.

Число участников торгов. Сколько игроков необходимо для получения точных результатов? Специалистам в области маркетинговых исследований хорошо знакомы такие понятия, как возможности стимулирования, статистическая значимость, демографические показатели, отбор участников и так далее. Но как вся эта математика работает в случае рынков прогнозов? В 2007 г. Джед Кристиансен из Букингемского университета решил это выяснить. Он поступил рационально, выбрав будущее спортивное мероприятие, крайне мало освещавшееся в прессе и не обладавшее какой-либо значимостью для жизни страны — соревнования по гребле. Участникам исследования надо было предсказать, кто станет победителем. Сложная задача, учитывая тот факт, что никто из экспертов не высказывал своего мнения по этому поводу в прессе, как это обычно

происходит в футболе. Однако Кристиансен отобрал респондентов из числа членов команд по гребле и их немногочисленных болельщиков на интернет-сайте сообщества любителей гребли; другими словами, он нашел экспертов. Полученные результаты оказались просто поразительными. Волшебным числом оказалось «16». Рынки с числом участников от шестнадцати человек можно назвать хорошо откалиброванными. Если игроков меньше, результаты оказываются недостаточно точными.

Эксперты — хорошие или плохие прогнозисты? А что означает само слово «эксперт», есть ли среди них люди, которые просто производят шум, но не повышают — а то и снижают! — точность прогнозов? Одно из исследований компании Prediki, проведенное в 2009 г. совместно с профессором Андреасом Милдом из Венского экономического университета, показало, что этот фактор может оказаться крайне важным. Исследователи использовали данные, накопленные за семь лет работы различных рынков политических, спортивных и корпоративных предсказаний для выявления среди их участников «супертрейдеров», «неудачников» и «среднячков». Было установлено, что хорошие результаты на одном рынке указывают на наличие хороших результатов и на других рынках. Затем исследователи пошли дальше и проанализировали ставки на рынках, нейтрализовав прогнозы «неудачников» и придав больше веса прогнозам «супертрейдеров». В результате точность предсказаний повысилась на 10–20%. Недавно рабочая группа продолжила это исследование. Была проведена оценка прошлых результатов трейдеров по отдельным областям, так как тот факт, что кто-то хорошо справляется с прогнозами в автомобильной отрасли, не означает, что он добьется подобных успехов в прогнозировании объемов продаж зубной пасты или результатов выборов.

Подведение итогов

Принимая во внимание, что в данной работе представлен большой объем полезной информации, мы хотели бы подвести итоги, перечислив основные выводы, сделанные нами в процессе проведения всех вышеописанных экспериментов.

1. Толпа может сделать очень мудрый прогноз, однако только в том случае, если она обладает определенными коллективными знаниями по данной теме. В противном случае она может оказаться абсолютно несведущей.
2. Прогнозы толпы о стоимости товаров могут несколько отставать от реальных цен, а женщины определенно лучше осведомлены о ценах на бытовые товары по сравнению с мужчинами.
3. В отсутствие необходимых коллективных знаний мы больше полагаемся на мнение окружающих, из-за чего некоторые прогнозы становятся зависимыми от воздействия эффекта внушения и стадного инстинкта.
4. Наши эмоции могут серьезно помешать нам сделать объективный прогноз.
5. Крайне важно то, как именно вы задаете прогностические вопросы. Они должны быть сформулированы объективно, как «проблемы», которые необходимо «решить» данной аудиторией.
6. С одной стороны, прогностический взгляд на ситуацию может быть более объективным, но, с другой стороны, ему не хватает детализации, поэтому прогностические вопросы необходимо задавать вместе с личными.
7. Методы рынков предсказаний позволяют более эффективно дифференцировать возможные варианты ответа, чем стандартное ранжирование элементов, однако максимизация этой эффективности при помощи геймификации на практике не приводит к дальнейшему улучшению качества сделанных прогнозов.
8. Можно повысить точность прогнозов путем взвешивания ответов, полученных из разных источников, на основе их надежности.
9. Определенно, существуют хорошие и плохие прогнозисты, однако их сложно выявить простыми тестами. (Возможно, все дело в том, что мы просто пока не нашли такой тест.)
10. Наши способности к прогнозированию определяются нашими знаниями и опытом в соответствующей области. Поэтому

можно повысить точность прогнозов, учитывая при оценке ответов компетентность (или надежность) прогнозистов.

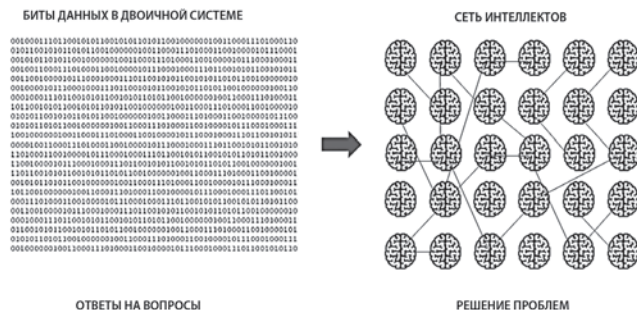
11. Уверенность человека в своем ответе напрямую соотносится с качеством его прогнозов так же, как и величина сделанной на него ставки. Поэтому эти два показателя можно считать косвенными индикаторами наличия необходимых знаний. Однако при этом необходимо классифицировать респондентов, чтобы учитывать базовый уровень их уверенности.
12. Метод рынков предсказаний определенно позволяет получить более надежные предсказания на микровыборках по сравнению с традиционными опросами при условии, что участникам предоставляется необходимая информация, а их работа квалифицированно модерируется.
13. Метод разделения «стада», вероятно, представляет собой неплохой способ смягчения воздействия стадного эффекта.
14. Для предоставления надежных прогнозов требуются микровыборки, отобранные по стандартам маркетинговых исследований, и мы уверены, что рынки предсказаний потенциально представляют собой идеальный вариант для улучшения качественных проектов, а также исследований сообществ в формате MROC.

Выводы

К каким основным выводам мы пришли, проанализировав все представленные в этой работе данные?

Для достижения максимальной эффективности методов прогнозирования в рамках обычных опросов мы рекомендуем отказаться от представления об участниках как об отдельных носителях информации и вместо этого воспринимать их как единую сеть интеллектуалов, которую можно использовать для решения проблем. Поставьте перед ними вопрос, который вы пытаетесь решить, и предоставьте им всю имеющуюся у вас информацию для его решения.

НОВЫЙ ВЗГЛЯД НА РЕСПОНДЕНТОВ



Информация, несомненно, является ключевым условием эффективного прогнозирования. Чем больше информации вы предоставляете респондентам и чем больше вы стимулируете их делиться имеющимися у них сведениями с другими прогнозистами, тем более точными получаются в итоге их прогнозы.

Мы также порекомендовали бы выбрать более «стереоскопический» подход к проведению исследований в целом. Прогнозирование невозможно использовать отдельно от вопросов о личном мнении респондентов по выбранной теме. Поэтому мы советуем внедрять больше элементов прогнозирования в исследования параллельно с увеличением количества личных вопросов. Как мы видели, эти два способа взглянуть на ситуацию могут предоставить нам разные точки зрения. С одной мы сможем увидеть лес, с другой — деревья.



Как все эти методы можно использовать для модернизации наших представлений о проведении исследований?

В завершение хотелось бы представить краткое описание способов применения знаний, полученных нами в ходе написания этой работы, для модернизации наших представлений о проведении исследований.

Традиционные методы:

- респондент воспринимается как отдельный носитель данных;
- вы собираете общую картину ситуации на рынке из отдельных кусочков, словно пазла;
- большинство вопросов направлены в прошлое для получения количественной оценки поведения людей;
- респондентам задают ряд вопросов;
- никакого сотрудничества или обмена мнениями.

Метод прогнозов:

- использование интеллектуальных возможностей респондентов для решения проблем;
- коллективное мнение респондентов представляет собой мудрость толпы;
- переход к вопросам, направленным в будущее;
- опрос предлагается в форме игры по предсказанию будущих тенденций на рынке;
- открытый обмен мыслями и мнениями.

Контрольные вопросы традиционных исследований:

- Какие бренды Вам известны?
- Вы приобретали товары этого бренда?
- Оцените этот бренд.
- С какими из этих утверждений о бренде Вы согласны, а с какими не согласны?
- Вы будете покупать продукцию этого бренда в будущем?

Контрольные вопросы прогностических исследований

- Как Вы считаете, у какого из этих брендов больше возможностей для роста?
- Не могли бы Вы объяснить причины такого ответа?
- В чем сильные стороны этого бренда?

- Как меняется модель Вашего поведения?
- Какие тенденции Вы замечаете?
- Вот мысли других респондентов по данному вопросу...

Благодарность

Выражаем особую благодарность Уолкеру Смиту из Futures Company за помощь в выборе правильного направления работы, а также за полезные советы в процессе ее написания.

Выражаем благодарность Рэю Пойнтеру из Vision Critical, именно его философские размышления о том, что влияет на нашу способность к точному прогнозированию, и что ей мешает, подвигли нас отправиться в это увлекательное путешествие.

Мы также благодарны участникам конференции 2014 MRS BIG Conference, которые помогли нам, подсказав некоторые из идей, рассмотренные в этой работе.

ЛИТЕРАТУРА

1. Chen K., Fine L. and Huberman B. Predicting the Future // *Information Systems Frontiers*. 2003. 5:1. P. 47–61.
2. Christiansen J.D. Prediction Markets: Practical experiments in small markets and behaviours observed // *Journal of Prediction Markets*. 2007.1.P. 17–41.
3. Galton F. 1907a. Letters to the editor. *Nature* 75. 1952.
4. Galton F. 1907b. Vox populi. *Nature* 75. 1949. P. 450–451.
5. Gigerenzer G. Why heuristics work. *Perspectives on Psychological Science*. 2008. 3. 1. P. 20–29.
6. Gordon K. Group judgements in the field of lifted weights. // *Journal of Experimental Psychology* 7. 1924. P. 389–400.
7. Goldstein D.G., McAfee R.P., Suri S. The wisdom of smaller, smarter crowds // 15th ACM conference on Economics and computation. 2014.
8. Herzog S.M. and Hertwig R. The wisdom of many in one mind: Improving individual judgements with dialectical bootstrapping // *Psychological Science*. 2009.20. P. 231–237.
9. Klugman S.F. Group judgement for familiar and unfamiliar materials // *Journal of Genetic Psychology* 32. 1945. P. 103–110.
10. Knight H. A comparison of the reliability of group and individual judgements. M.S. thesis. Columbia University. unpublished. 1921.
11. Levitt S.D., Miles T.J., Rosenfield A.M. Is texas hold'em a game of chance? a legal and economic analysis // *The Georgetown Law Journal* 101. 2012. P. 581–636.
12. Lorenz J., Rauhut H., Schweitzer F., Helbing D. How social influence can undermine the wisdom of crowd effect // *Proceedings of the National Academy of Sciences* 108. 2011. 22. P. 9020–9025.
13. Lorge I., Fox D., Davitz J., Brenner M. A survey of studies contrasting the quality of group performance and individual performance, 1920–1957 // *Psychological Bulletin* 55. 1958. 6. P. 337–372.

14. Lovie A.D., Lovie P. The flat maximum effect and linear scoring models for prediction // *Journal of Forecasting* 5. 3. 1986. P. 159–168.
15. Luckner S. How Do Incentive Schemes Affect Prediction Accuracy? Dagstuhl Seminar Proceedings 06461. 2007.
16. Mannes A.E., Soll J.B., Larrick R.P. The wisdom of small crowds. Manuscript in Preparation. 2013.
17. McMurray J.C. Aggregating information by voting: The wisdom of the experts versus the wisdom of the masses // *The Review of Economic Studies* 80. 1. 2013. P. 277–312.
18. Muchnik L., Aral S., Taylor S.J. Social influence bias: A randomized experiment // *Science* 341. 6146. 2013. P. 647–651.
19. Puleston J., Sleep D. The game experiments: Researching how game techniques can be used to improve the quality of feedback from online research. 2011. Esomar Congress Amsterdam.
20. Servan-Schreiber E., Wolfers J., Pennock D.M., Galebach B. Prediction Markets: Does Money Matter? *Electronic Markets*, Volume 14. 3. 2004.
21. Simmons J.P., Nelson L.D., Galak J., Frederick S. Intuitive biases in choice vs. estimation: Implications for the wisdom of crowds // *Journal of Consumer Research* 38. 1. 2011. P. 1–15.
22. Surowiecki J. *The Wisdom of Crowds*. 2005.
23. Tetlock P.E. *Expert Political Judgment: How Good Is It? How Can We Know?* 2006.
24. Treynor J.L. Market efficiency and the bean jar experiment // *Financial Analysts Journal* 43. 3. 1987. P. 50–53.
25. Waitz M., Mild A. Improving forecasting accuracy in Corporate Prediction Markets // *Journal of Prediction Markets* 3 3. 2009. P. 49–62.

Раздел 6

Методология онлайн-исследований: качественные исследования

Качественные онлайн-методы

*Рожественская Елена
(Национальный исследовательский университет
«Высшая школа экономики»)*

Немецкий футуролог и исследователь трендов будущего Маттиас Ксоркс пишет о технологии, представляя технологический прогресс не в линейной форме, но в коэволюции с человеческой культурой, ревизией дигитального в широком смысле и переходом от инноваций к синновациям [53]. Связывая цифровую революцию с энергетической, американский экономист Джереми Рифкин [42] говорит о поворотном моменте в истории, который влияет на повседневную жизнь, практики неслыханным образом и продолжать их менять далее. Распространение цифровых режимов коммуникации через интернет и спутниковую связь заинтересованных сторон по всему миру в считанные секунды позволяет преодолеть временные и географические барьеры в буквальном смысле.

Вопрос о физическом соприсутствии двух партнеров по коммуникации не существенно тех пор, пока оба в равной степени имеют доступ к одним и тем же интернет-технологиям. Как это влияет на научно-исследовательский процесс в смысле предмета исследования и методологии исследования?

Увеличение технологической и экономической доступности интернета широким социальным слоям открывает дорогу альтернативным способам сбора данных в социально-исследовательской практике. Интернет широко используется в рамках количественных методов сбора данных, но также есть многочисленные примеры успешной реализации качественных онлайн-исследований.

Надежных данных о распространении онлайн-исследований, однако, не хватает. Качественные методы, используемые онлайн, мало отличаются по своим принципам, как правило, от традиционных офлайн исследовательских подходов. В самом общем виде можно идентифицировать три качественных метода онлайн: опрос, наблюдение и анализ документов. Новые условия, в которых проходит процесс исследования, требуют не только использования инновационных технологических разработок для того, чтобы войти в контакт с полем, но и адаптации или приспособления заслуживших признание в офлайн- методологических приемов взаимодействия между исследователями и респондентами. Если качественное онлайн-исследование и не предоставляет приемлемого метода для любой исследовательской ситуации, по крайней мере, оно предлагает, как мы покажем далее, богатые возможности для постановки широкого круга научных вопросов.

Качественный онлайн-опрос

Качественные онлайн-опросы имеют относительно долгую традицию на рынке и в социологических исследованиях. Сегодня исследователи обладают широким спектром различных качественных неструктурированных и полуструктурированных методик опроса для решения различных задач (например, лейтмотивное или нарративное интервью, см. [17]). Качественные формы исследования на базе интернета могут быть классифицированы по различным параметрам: единичное против группового интервью [38; 40], синхронное, в форме почти одновременного диалога [11;4] против асинхронного (в форме последовательных сессий вопросов и ответов [2;3]), метод опроса против метода дневников [22] и многие другие. Различные технические возможности реализации синхронных и асинхронных форм опроса в интернете перечислены ниже.

Асинхронный опрос	Синхронный опрос
E-mail	Веб-чат
Рассылка по списку	Рассылка мгновенных сообщений
Новостные группы	Видеоконференция
Дискуссионный форум	
Открытый веб опрос	

Одну из самых популярных форм опроса на основе интернета представляет собой асинхронное единичное интервью, в котором респонденты получают исследовательские вопросы, с просьбой отослать ответы по электронной почте по отдельности или в тематических блоках. Если в традиционном интервью вопросы задают по очереди, то текстуальный опрос по электронной почте соблазняет задавать несколько вопросов сразу, чтобы ускорить процесс интервью и не мучить респондента слишком долго многими отдельными вопросами. Хотя подобное действительно доказано для стандартизированных, письменных опросов при включении соответствующих альтернатив ответов, это не обязательно рекомендуется в качественных онлайн-подходах. Открытые вопросы, которые используются в качественных исследованиях, требуют значительно больше когнитивных усилий, чем при работе со стандартизированной анкетой, поскольку ответы не просто пассивно распознаются и выбираются, а должны вначале активно конструироваться самим респондентом. Чтобы не демотивировать респондента длинным списком вопросов, слишком обширным ввиду возможной усталости или разочарования, которые могут привести к отказу от взаимодействия, нужно обратить внимание на баланс вопросов и опросных эпизодов. Электронная почта является одной из форм общения, которое построено на довольно коротких автономных единицах текста. Поэтому лишь несколько тематически связанных вопросов могут быть сгруппированы в одном послании электронной почты (см. [4]). Помимо простых интервью в опросе на основе электронной почты также могут быть использованы и специальные методы измерения. Baker [3] использовала электронную почту только в качестве

доставки, благодаря чему респонденты знакомились с содержанием кейса и включались в обсуждение. Однако, как показали Hessleretal [22], реализуем и сбор данных по электронной почте в течение более длительного периода. Вышеуказанные авторы попросили подростков в рамках исследования о рискованном поведении в течение десяти недель вести дневник и посылать им эти страницы по почте.

Преимущества и недостатки онлайн-опросов (см. [2;4])

Преимущества	Недостатки
Снижение затрат	Низкая спонтанность
Экономия времени	Высокая доля задержки ответов
Большая открытость	Отсутствие невербальных сигналов
Деиерархизация коммуникации	Необходимость технической инфраструктуры
Доступ к особым социальным группам	

В сравнении с офлайн-интервью онлайн-интервью влечет за собой два типа смещений, связанных с двумя основными измерениями человеческого опыта. Во времени взаимодействие между интервьюером и респондентом чаще асинхронно, с паузами различной длины от всплесков связи или «эпизодов», а в пространстве взаимодействие осуществляется дистанционно, посредством электронной связи. Асинхронность электронного интервью имеет несколько последствий — задержки во взаимодействии между исследователем и респондентом могут варьироваться от нескольких секунд (практически в режиме реального времени) до нескольких часов или дней. Существенно, что собеседник не обязан ответить в кратчайшие сроки, что создает преимущество электронного интервью, поскольку участники коммуникации не принуждены определять в переговорах взаимно удобное время для беседы. Длительные паузы между сообщениями в электронном интервью дают собеседникам время, чтобы выстроить ответ на конкретный вопрос, найти информацию, хотя исследователь в таких случаях не знает, какие ресурсы собеседник привлечет, отвечая на вопрос, на что обращает внимание Ли [30,

p.118]. Полагают, что асинхронность может снизить нервность опрошенных, давая люфт для размышлений, но, с другой стороны, электронное интервью также теряет в спонтанности, безусловно, способной обогащать данные. Кроме того, информант может оспорить точность транскрипции при контрасте с произнесенным офлайн; в онлайн-интервью процесс коммуникации тщательно задокументирован, но и желаемое впечатление продумано. Тем не менее спонтанность не является недостижимой, как сообщают Бэмптон и Каутон [4], а также Кристел [13]; в их опыте некоторые быстро поступившие ответы часто содержат виды орфографических и грамматических ошибок, характерные для электронной почты. Можно предположить, что они не были отредактированы. Возможное снижение спонтанности может быть преимуществом для сенситивных исследовательских тем, к которым подход респондентов нарративно заблокирован. В электронном интервью у коммуникантов есть время, чтобы отреагировать на развитие диалога. Одной из практических проблем, которые возникают в момент ожидания ответа от информанта, является неясность происхождения этой паузы (занятость, отложенный ответ, обдумывание, коммуникативный сбой, техническая проблема, отказ). Неопределенность может быть отчасти снята отправленным напоминанием, но исследователи идут на это неохотно [4], боясь, что этот жест будет расценен как необоснованное давление на информанта. Проблема в том, что физическая удаленность делает ситуацию очень трудно считываемой. Однако как и для любой социально щекотливой ситуации, может быть разработан репертуар подходящих тактик и набор фраз, соответствующих контексту исследования. Альтернативой является гиперссылка всех поступивших сообщений электронной почты, простой список протоколов, которые возобновляются как напоминание о возможно пропущенном сообщении ввиду ненадежности сетевого общения, а также как приглашение к его завершению, если становится очевидным исчерпанность темы. Одним из способов снижения риска срыва интервью является ознакомление собеседника со всеми вопросами исследования сразу, что немыслимо в интервью офлайн. Но в результате предоставления нескольких вопросов интервью может показаться для информанта сложным, он

может задержаться с ответом или даже оказаться не в состоянии реагировать вообще. Есть и техническое возражение Кристел [13]: в отношении электронной почты распространено ожидание, что цепочка текстообмена должна быть полностью видна на экране без необходимости его прокрутки. При нарушении данного принципа теряется ощущение неформальности длящегося разговора. Вместо того чтобы отправить большое число вопросов сразу, другие вопросы и новые темы возникают в последовательных эпизодах интервью. Тем не менее должен быть баланс между нагрузкой на респондента в рамках одного сеанса связи и числом сеансов, также утомляющим информанта.

Следующий практически нерешенный вопрос заключается в определении того, когда интервью близится к концу. В офлайн-интервью интервьюер, как правило, чувствует, когда время исчерпано. Соответственно, он перестраивается, отбирая наиболее важные вопросы и пренебрегая второстепенными. Однако существуют некоторые возможные сигналы необходимости скорого завершения интервью и онлайн. Например, ответы становятся все более затянутыми, качество ответов скудеет — все это признаки ослабевающего интереса. С точки зрения западного исследовательского менталитета [31], решение лежит в установлении и заключении предварительных правил «на берегу». Это может касаться продолжительности интервью, конфиденциальности и режима напоминаний. В отличие от электронного интервью, офлайн респонденты часто работают дольше, чем предварительно пообещали [44]. Но в зазеркалье электронного интервью сюжет времени не очевиден. Скорее всего, интервьюер останется в неведении, насколько его собеседник ангажирован к участию. Итак, прагматичным решением было бы в случае полуструктурированных интервью послать информанту примерный список вопросов, подлежащих раскрытию.

Опосредованность коммуникации интервьюера и информанта порождает и другие проблемы. Например, сигналы невербального поведения «могут означать, что собеседнику становится непросто или его беспокоит расспрашивание» [9, р. 319]. Говоря о потерях данных, Брайман замечает: «Качественные исследователи часто заинтересованы не просто в том, что говорят информанты, но и в том,

как они это говорят» [9, р. 319]. В этом смысле распознавание информации в электронном интервью ограничено [31]. Итак, при всех усилиях, дистанцированность интервьюера и интервьюируемого в электронном интервью снижает богатство сообщения, порождая недоразумения и сбой в коммуникации. Плюсы: информант может быть защищен от переживаний, смущения, социального контроля при условии физического отсутствия интервьюера, форма интервью менее интервенирующая, степень анонимности выше, чем в интервью офлайн. Степень анонимности — безусловно амбивалентный сюжет, угрожающий в ситуации онлайн потерей идентификации защиты от идентификации собеседника. Кто на самом деле дает интервью, — многократно муссируемый, но не закрытый вопрос. Он отчасти решается на уровне «привратников»- контролеров, гейтскайперов поля. В этом убеждены западные коллеги, приводя комичный для нас пример сотрудника тюрьмы или священника, если осужденный преступник дает онлайн-интервью. Прагматичнее кажется замечание относительно текстуальности, выдающей информанта «с головой» — сленговые выражения, использование нетрадиционных фраз и наличие опечаток и грамматических ошибок [13]. Кроме того, электронное интервью может защитить оба лица в онлайн коммуникации, предлагая степень анонимности через принятие ника или дегендеризации и снижая предвзятость в отношении пола, возраста, этничности, социального статуса.

Итак, интернет как средство опроса имеет ряд преимуществ и недостатков. Учитывая (в основном) ограничения ресурсов, одним из основных аргументов в пользу использования интернета является экономия времени и материальных затрат. Благодаря асинхронной коммуникации по электронной почте уже нет нужды в одновременном соприсутствии респондентов и интервьюеров. Это становится преимуществом во времена, когда в связи с плотной структурой дня труднодоступных людей взять у них интервью возможно только в отдельные, ограниченные временные интервалы. Наконец, метод опроса по электронной почте обеспечивает такое финансовое преимущество, как снятие потенциальных дополнительных расходов на обеспечение места для интервью, транспорт и транскрибирование (см. [2]). Кроме того, интернет дает одновре-

менный доступ к конкретным группам людей, разделенных в пространстве (например, людям из других стран), чего прежде было бы трудно или невозможно достичь. Далее, отсутствие прямой личностной интеракции с интервьюером и иногда связанного с этим давления в пользу необходимости сразу давать ответ, что особенно касается нервных, социально внушаемых респондентов, позволяет им быть более расслабленными. Таким образом, они начинают отвечать более свободно и открыто (см. [2]). Тем не менее онлайн-опросы по электронной почте также связаны с некоторыми неудобствами. Проблематичным может быть асинхронный характер опроса из-за возможно долгого периода реагирования, так как может не лучшим образом сказаться на спонтанности ответа (см. [4]). В случае задержки в интервью по электронной почте респонденты имеют возможность дольше подумать о своих ответах, могут привлечь дополнительную информацию и сформулировать ответы лучше, чем это было бы возможно при синхронных формах опроса. Это может (но все же не должно) ухудшить качество ответов, потому что, как показали Vamptson и Cowton [4], на основе анализа входящей электронной почты, многие ответы не представляются сильнее продуманными или лучше сформулированными по сравнению со спонтанными ответами. Тем не менее, могут возникнуть трудности, когда респондент замолкает надолго. Часто невозможно определить, то ли у респондента просто не было времени ответить и ответ еще поступит, то ли он не смог ответить на вопрос по каким бы то ни было причинам. Так, при интервью по электронной почте бывает трудно понять, не возникло ли у респондента желание завершить интервью, поскольку соответствующие невербальные сигналы, которые считываются в личной беседе, в данном случае отсутствуют. Указанием на эту ситуацию могут быть все более краткие ответы и увеличение интервалов между репликами [4]. Поэтому коммуникативное содержание онлайн-интервью значительно беднее именно благодаря отсутствию невербальных сигналов, таких как тон голоса, выражение лица или телесные позы, что в изобилии присутствует в сопоставимых интервью *face-to-face*. Можно попытаться скомпенсировать этот недостаток благодаря использованию эмодиконов, коротких текстовых или графических репрезентаций

эмоционального содержания, но это вряд ли эквивалентно полноте демонстрируемых человеком эмоций ([31]).

С традиционными методами обследования более схожи синхронные формы опроса, которые проводятся посредством веб-чата или через рассылку мгновенных сообщений. В отличие от асинхронной электронной почты, они имеют много преимуществ, особенно в виду отсутствия задержки ответов. Так как, например, в личной беседе реакции следуют непосредственно в ответ на заданный вопрос, то от респондента при стимулировании на повествование ожидаются спонтанные и распространенные ответы. Однако несмотря на более низкую по сравнению с электронной почтой задержку ответов, для проведения синхронного онлайн-интервью требуется значительно больше времени (почти в два раза), чем для интервью face-to-face; в то же время ответы в среднем получаются менее длинные [38], а также [14]. В синхронных опросах сильное влияние на ответы оказывают скорость чтения и печати респондентов. Еще большей задержке ответов может способствовать отсутствие невербального подкрепления (например, через зрительный контакт) ([31]). Особенностью коммуникации в чате является его текстуальная структура. Если личные беседы характеризуются линейной последовательностью вопроса и ответа, эта структура в онлайн-чатах часто нарушается.

В связи с временной задержкой ответ или следующий вопрос появляются не всегда в той последовательности, которая отвечает содержанию диалога, а со смещением. Это становится проблемой, если ответ можно отнести как к текущему вопросу, так и к предыдущему, что может привести к ненужной многозначности в рамках последовательного содержательного анализа. К подобным трудностям способно привести и использование сарказма и метафор. Сарказм может быть легко проинтерпретирован как оскорбление из-за отсутствия невербальных стимулов [14], что приведет к ошибочным интерпретациям.

Особое значение интернет в качестве средства качественного исследования приобретает еще и в связи с тем, что предлагает ранее практически не гарантированную возможность анонимных сценариев коммуникации. Если в асинхронном онлайн-интервью это играло до сих пор скорее подчиненную роль, то большая аноним-

ность и приватность принципиально осуществимы в синхронном онлайн-опросе. Анонимизированные онлайн-взаимодействия позволяют исключить невербальные сигналы, которые часто бессознательно влияют в ходе личных бесед, и помогают оттеснить на задний план культурно обусловленные иерархии, например, по полу, возрасту, или физическим характеристикам.

Это может потенциально привести к большей эгалитарности среди участников дискуссии и более открытому разговору, сфокусированному на сюжете обсуждения. Особое значение имеет анонимность качественных онлайн-опросов прежде всего для обсуждения сенситивных, потенциально смущающих проблематик, о которых сложно говорить в открытом личном диалоге (ср. [2]), например, в сфере сексуальности [14; 11; 32]. Так, например, Чейни и Дью [11] рекрутировали респондентов для исследования сексуальной зависимости у гомосексуальных мужчин в соответствующие чатах, где они ожидали контакта с другими анонимными участниками. После того, как к респондентам обратились, им сообщили об исследовательских намерениях и пригласили к однократному интервью через мессенджер (рассылка мгновенных сообщений). Как показало это исследование, качественный онлайн-опрос вполне реализует доступ к представителям особых социальных групп, которые вряд ли достижимы или с большим трудом доступны с помощью традиционных методов обследования. Например, ислам запрещает женщинам общаться с неженатыми мужчинами. В связи с этим запретом в последние годы в арабском мире сформировались весьма активные онлайн-сообщества, в которых женщины и мужчины могут контактировать друг с другом. В таких чатах авторы [2] смогли опросить арабских женщин, к которым они как мужчины в противном случае вряд ли имели бы доступ. Наконец, нужно принять во внимание и немаловажный момент необходимости технической инфраструктуры для реализации онлайн-интервью. Она должна иметь место не только с обеих сторон, интервьюера и респондента (использование электронной почты — обычное дело, но реализация синхронных интервью уже технически сложна); также полученные данные должны быть должным образом обработаны, а необходимые для этого медиакомпетенции присутствуют не во всех социальных группах.

В частности представители пожилых или менее технически образованных групп людей часто испытывают трудности в обучении и сценарии применения нового программного продукта, которые сопровождаются (часто необоснованно) страхами что-то испортить из-за неправильного употребления. В результате либо приходят к полному отказу от использования программного продукта по сопровождению онлайн-опроса, либо к преувеличенным предосторожностям, которые проявляются в очень нерешительном обращении с инструментом коммуникации, что может повлиять на онлайн-диалог. Таким образом, структурные, а также эмоциональные барьеры усложняют или делают невозможным в определенных обстоятельствах рекрутирование и интервьюирование весьма специфических выборок; таким образом, в указанных случаях исследовательские онлайн-средства не могут быть предложены.

Несмотря на потенциальные подводные камни, качественные методы синхронного онлайн-опроса уже прижились в различных содержательных областях исследовательской практики. Например, весьма широкая тенденция к использованию онлайн-опросов наблюдается в контексте исследования потребительского поведения. Так, авторы [39] сделали отчет об использовании техники лэддеринг¹ в интернете, полу-стандартизированного метода обследования для выявления цепочек конечных мнений. При этом они исходили из свойств продукта, которые замерялись в качественном интервью, и делали заключение через компоненты потребностей о ценностях и целях потребителей. В рамках сравнительного исследования по имеющимся в продаже MP3-плеерам исследователи [39] смогли

¹ Лэддеринг (метод лестницы) — это полуформализованный качественный метод в маркетинговых исследованиях, с помощью которого стремятся понять, почему люди покупают и используют те или иные продукты и услуги. С его помощью выясняется, какой продукт особенно важен для пользователей продукции и обладает эмоциональными преимуществами для покупателей или пользователей, таким образом устанавливая причинно-следственные связи между свойствами продукта или услуги и выгодами от использования и ценностями потребителей. Таким образом, индивиды покупают в оглядку на свои эмоции и убеждения, затем рационализируя свой выбор функциональными преимуществами. Авторы Томас Дж. Рейнольдс и Джонатан Гутман разработали и внедрили лэддеринг в 1988 году на основе теории means-end-chain Гутмана.

продемонстрировать более высокую эффективность качественных интервью в чате с использованием техники лэддеринг по сравнению с анкетированным опросом.

Групповая дискуссия

Групповые дискуссии (фокус-группы) образуют центральный элемент исследований рынка и общественного мнения, который все более разворачивается к интернету. Подобно индивидуальным интервью, они могут принимать синхронную и асинхронную форму. В синхронной фокус-группе, организованной в рамках чата или телеконференции, участники тут же обмениваются информацией, в то время как асинхронные опросы, в которых участники реагируют с временным смещением, размещают на форумах, а также проводятся посредством электронной почты или группового списка рассылки (см. [40]).

Большое значение как руководитель фокус-группы в случае синхронной дискуссии в чате приобретает модератор, главная задача которого заключается в инициации и ведении группового обсуждения. В присутственных дискуссиях модератор уже своим физическим статусом, мимикой и жестами существенно контролирует группу. В виртуальной среде всего этого не хватает. Таким образом, модератор должен полагаться исключительно на текст, письменное слово, чтобы влиять на групповые процессы. Это представляет собой сложную задачу даже для опытных модераторов. В синхронной онлайн фокус-группе модератор должен также принимать во внимание множество различных потоков информации и не потеряться в разветвлениях дискуссионных потоков. С ростом группы это может стать сложной задачей. На практике часто оказывается полезным реализовывать модерацию не в одиночку, а при поддержке ассистента или привлекая заранее подготовленные фрагменты текста, которые нужно только скопировать и вставить [32; 49]. Коммерческие системы веб-чатов для проектирования групповых дискуссий могут помочь модераторам рядом полезных функций в их работе. В так называемом «ник-листе» на дисплее представлены все участвующие в дискуссии лица. Возможных участников, скры-

тых наблюдателей могут видеть только модераторы. В зависимости от функционала программного обеспечения управление процессом дискуссии позволяет ограничить доступ к чату конкретных лиц, за отдельными участниками признать ограниченное право говорить, а кого-то лишить на время голоса. Также при модерации некоторых веб-чатов возможно загрузить сформулированные лейтмотивы заранее и затем в ходе обсуждения выбирать отдельные вопросы из гайда. Это экономит усилия по вводу вопросов. Словесные вставки могут касаться и гиперссылок, которые, если они не кликабельны для всех участников, позволяют открыть новое веб-окно. Тем самым возможна презентация внешнего стимульного материала.

В идеале все обсуждения участников протоколируются в базе данных и совмещаются с временем внесения и никнеймом участника, чтобы затем манипулировать дифференцированными оценками. Современные конференционные системы на базе интернета позволяют импорт аудио- и видеоинформации сигналов (с помощью микрофонов, компьютерных динамиков и веб-камеры), а также Power Point презентаций. Дискуссионный чат хотя и присутствует обычно в этих системах, но часто имеет второстепенное значение. Современные системы также часто интегрируют интерактивные интерфейсы пользователей, и не просто пассивно отражают их содержание, но также могут генерировать новое содержание. Так называемые белые доски, интерактивные стены для записей позволяют респондентам через интернет, как на листе бумаги, написать на экране или оставить графические сообщения. Это может быть использовано для объединения нескольких участников вместе в целях разработки когнитивных карт или креативных идей относительно каких-либо продуктов.

Преимущества и недостатки онлайн-фокус-групп практически те же, что и для персональных отдельных интервью. Они экономят время и экономически более эффективны, более удобны для участников и ведут к большей открытости ответов [7 цит. по 43]. Онлайн-фокус-группы дают возможность доступного обсуждения на региональном уровне или среди территориально рассеянных участников [40]. В этом исследовании респонденты-студенты, которые проживали по всей территории Соединенных Штатов, были

включены в список рассылки, который служил дискуссионной платформой на тему участия в системе дистанционного обучения. Получив приглашающее обращение исследователя, студенты одного из курсов отвечали на него, а также на последовавшие содержательные реакции других участников обсуждения. Так, с течением времени может развиваться плодотворная асинхронная дискуссионная культура. Из-за анонимности онлайн фокус-групп потенциальные коммуникативные проблемы исключены благодаря эффектам модерации (социально-демографические аспекты самого модератора). Кроме того, они также позволяют застенчивым студентам, принять участие в обсуждении [32].

Преимущества и недостатки онлайн-фокус-групп (см. [32, цит. по 43])

Преимущества	Недостатки
Быстро, недорого	Необходимы базовые медиа компетенции
Удобно для исследователей и участников	Трудно идентифицировать ложь
Легче обсуждать чувствительные вопросы	Трудно установить отношения
Снижен эффект экспериментатора	Отсутствие невербальных сигналов
Нет необходимости в транскрипции данных	

В контексте сравнения коммуникативного и дискуссионного поведения в обычной групповой дискуссии лицом к лицу и в онлайн групповой дискуссии удалось пронаблюдать множество методических различий [18]. Так, в чате происходят «параллельные дискуссии», то есть участники отклоняются от обсуждения общей темы и переходят на коммуникацию друг с другом. Вклад в дискуссию в виде развивающих тему высказываний в чате часто осуществляется отдельными лицами. Напротив, участники дискуссии лицом к лицу ориентированы скорее на групповое обсуждение в целом. Возможно, из-за анонимной ситуации опроса и естественной, привычной среды, в которой находятся участники в процессе обсуждения (как правило, дома на родном компьютере), в онлайн-чате регистрируется большое количество открытых мнений (см. также [35]).

Что касается числа производимых слов, то, как было обнаружено [18], участники офлайн-дискуссии производят их в три раза больше, чем онлайн. С другой стороны, участники онлайн-чата чаще берут слово, то есть частота сообщений в чате выше при том, что в целом здесь производится меньше текста. Тем не менее некоторые авторы считают, что в онлайн-фокус-группах из-за отсутствия невербальных стимулов, когда участники только читают с экрана компьютера, вряд ли можно инициировать реальную групповую динамику между участниками. Таким образом, традиционные фокус-группы продуктивнее онлайн-версий. Поэтому автор [43] выразительно подчеркивает, что онлайн-фокус-группы не могут быть заменой для офлайн-фокус-групп, но вполне их дополняют.

Качественное наблюдение

Особой популярностью онлайн-методология пользуется в сфере наблюдения. Благодаря разнообразным доступным комбинациям данных наблюдение представляет собой один из самых распространенных качественных методов сбора данных в интернете. В традиционном исследовательском дизайне наблюдение, несмотря на очень продуктивные результаты, не часто используется из-за организационных и экономических причин. Наблюдение требует много времени и задействованного персонала, чтобы собрать валидные данные. Интернет, однако, часто дает прямой доступ к интересующей группе и позволяет за счет автоматического протоколирования log-файлов (например, в форме диалогов на дискуссионных форумах) в значительной степени беспрепятственно получить данные. Таким образом, онлайн-методы набирают особый вес в контексте виртуальной этнографии. Интернет для многих людей уже не просто средство коммуникации, но в растущей степени социальное пространство коммуникации, в котором протекает часть повседневной жизни. Соответственно, онлайн-наблюдения открывают доступ к взаимодействиям индивидов и групп, которые трудно обнаружить традиционными методами.

Наблюдения могут быть дифференцированы относительно уровня или меры участия исследователя — в форме участвующего наблюдения, когда «наблюдатели активно участвуют в происходящем» [8,

р. 267], и не участвующего наблюдения, когда наблюдатель ставит себя вне происходящих событий и регистрирует события поля вне активного вмешательства. Многочисленные источники данных для неучаствующего наблюдения можно обнаружить в форме случайных конфигураций данных на размещенных в интернете форумах, новостных группах и онлайн-чатах.

Основной объект изучения представляет собой, как правило, определенные аспекты коммуникативной структуры наблюдаемого сообщества, группы или отдельных участников [10; 35; 16]. Примером может послужить исследование [35], в котором были проанализированы паттерны коммуникации в тайских чат-форумах. Контент-анализ чатов выявил значительное взаимовлияние гендера и культуры. Женщины, которые, как правило, занимают в тайском обществе подчиненное положение, через интернет-общение существенно расширяют свои социальные возможности. Они часто вовлекаются в обсуждение и вызывают больше реакции на свое участие в дискуссиях, чем офлайн как мире «доминирующих» мужчин. Другие ключевые вопросы исследований возникают из основных тем, на которые ориентированы анализируемые форумы.

Так, имеются многочисленные примеры качественных онлайн-наблюдений в различных областях. Например, в области образования — в контексте семинаров по онлайн-обучению, которые часто сопровождают дискуссионные форумы, чтобы поддержать онлайн внедрение учебного процесса [20; 21]. Или в медицинском контексте — в виде онлайн-групп самопомощи [27; 29]. Или в контексте анализа изображений, который осуществляется в онлайн-дискуссиях об отдельных компаниях и их продуктах [52]. Исследователи изучали в экспериментальном дизайне, какое поведение преподавателя курса на онлайн-форуме поощряет критический дискурс учебного материала [20]. И на основе контент-анализа форумных высказываний они пришли к выводу, что требовательный руководитель стимулирует учащихся к основательно аргументируемым ответам. Процедура опроса потребителей по поводу предпочитаемых продуктов упрощается, если, как в случае навигационной системы компании «Tom Tom» (2006), производитель обзаводятся целевой группой интернет-пользователей. На сайте этой фирмы есть дискус-

сионный форум для пользователей, на котором клиенты обсуждают перспективы продукции компании и, таким образом, обеспечивают богатую информацию для будущих инноваций и дальнейшего развития.

В качественном онлайн-исследовании акцент сделан на «субъективных смыслах и взаимодействиях» [19, р. 158], при этом востребован более специализированный способ доступа в форме включенного наблюдения. Оно занимает доминирующее положение в изучении киберкультур [6], предполагающем описание индивидуальных или коллективных субъективностей с целью понимания их вклада в локальную культуру. В классической этнографии, не учитывающей специфики компьютерно опосредованной коммуникации, исследователь лично погружается в изучаемое сообщество, знакомится с участниками и их повседневной деятельностью, чтобы разобраться в их опыте. Специфика интернета дает возможность модификации привычного дизайна этнографического исследования. Задокументированная и большей частью текстуализированная коммуникация внутри, например, виртуального сообщества позволяет обратить внимание не только на интерпретацию текстов. Обсуждая онлайн-исследования в культурологии, Дж. Стерн [48] предлагает сместить фокус с интерпретации текстов к анализу контекста происхождения текста, события или практики. Тогда мы иначе формулируем исследовательские вопросы, нам важно не столько то, что означает данное событие для участников, но как и благодаря каким условиям локальных практик эти значения становятся возможными [25, р. 262]. Развивая эти идеи, последователи киберэтнографии даже формулируют интерактивную методiku на основе эпистемологии *doing*, делая акцент на самовоспроизводстве и процессуальности онлайн-идентичности [41]. Эта методология предполагает, что субъекты / объекты самовоспроизводят себя через текстуальные продукты — отчеты, письма, работу с изображениями, создание аватаров, цифровых видео- и аудио- материалов, и по большому счету участвуют в практиках повседневной жизни через эти интерфейсы.

Приведем пример. В этнографическом наблюдении скандинавских исследователей была решена задача о том, как женщины обходятся с жизненной драмой рака груди и часто связанной с этой

болезнью социальной изоляцией [23]. Они не только наблюдали за происходящим в списке рассылки для самопомощи, но и сами принимали активное участие, писали посты, соучаствовали в радостях и горестях пострадавших и таким образом смогли получить богатый материал, раскрывающий субъективное значение и важность этих форумов для пострадавших, что значительно шире простого офлайн-наблюдения. Хотя включенное наблюдение доказало свою уместность во многих исследовательских областях, вновь и вновь здесь порождаются научно-этические дилеммы. С одной стороны, общие руководящие принципы или стандарты для онлайн сбора данных (ADM 2001) рекомендуют априори объяснение целей исследования и согласия сторон в нем участвовать. С другой стороны, исходя из исследовательско-методологических соображений, это требование, особенно при наблюдении, как правило, контрпродуктивно, так как последнее направлено на замер «естественного» поведения, а естественность ввиду привлечения внимания наблюдаемых к целям исследования как раз оказывается под угрозой.

То, в какой мере посты в публично доступных форумах отражают физические следы в смысле, употребленном Борц и Деринг, и могут быть собраны и проанализированы (в соответствии с обычными этическими стандартами) [8, р. 325] или, если речь о частной коммуникации, представляют собой предмет авторского права, — еще довольно спорно [17]. Исходя в первую очередь из прагматических соображений, большинство исследователей придерживается первой позиции, описанной выше, — работать без эксплицитного предвещения участников о целях предстоящего исследования. Часто метод включенного наблюдения используется, чтобы исследовать онлайн-сообщества, какие существуют, например, в так называемых многопользовательских онлайн-играх, виртуальных мирах, в которых могут онлайн участвовать игроки со всего мира (см, например, [28; 50]. Участвующее наблюдение дает некоторые преимущества в изучении таких виртуальных социальных пространств. Во-первых, такие исследования всегда обнаруживают субъективный, интерпретативный аспект, поскольку поведение таких групп едва ли наблюдаемо вне той среды, в которой оно разворачивается. Только через активное участие в (игровом) действии исследователю открывают-

ся — как часть рассматриваемого явления — предмет исследования и возможность дальнейшего его изучения. Кроме того, активное участие облегчает понимание игры, неявных правил поведения, и может повысить доверие игроков, показав, что исследователь принадлежит сообществу. Так, например, исследователи [55] проанализировали правила социального взаимодействия в многопользовательской онлайн-стрелялке «Counter-Strike». В качестве источников данных для этого исследования был доступен: (а) обширные лог-файлы из нескольких игр на различных серверах, с помощью которых ученые могли извлечь разговоры участников во время игры и закодировать их; (б) собственные наблюдения, которые они смогли собрать в качестве участников игры. Также были составлены графические репрезентации игроков (Логос) и зафиксированы невербальные интеракции. Вопреки общественному мнению, согласно которому такие жестокие игры часто имеют отрицательный эффект, Райт и др. пришли к выводу, что истинный смысл игры для участников менее всего заключается в акте виртуального насилия или просмотра кровавых изображений, но в большей степени в социальных взаимодействиях друг с другом. Таким образом, игроки обучаются способам социального поведения, которые выстроены в основе установленных кодов поведения в реальном мире, и могут поэкспериментировать в виртуальном пространстве с нарушением правил и их следствиями — но уже без страха перед последствиями офлайн.

Анализ документов

Анализ документов является очень расплывчатым термином в контексте качественного онлайн-исследования, в том числе и потому, что метод наблюдения главным образом исходит из документов (как записанных разговоров). В дальнейшем анализ документов фигурирует как исследовательский подход, который базируется на веб-документах, которые не включают коммуникативное содержание в смысле прямого диалога. Тем не менее это определение не лишено проблем. Так, феномен блоггинга находится в серой зоне между простой публикацией статьи и репликой на другие записи. Документы для анализа представлены в интернете как правило на сайтах всех

видов. В этой связи можно различить два научно-исследовательских подхода: вебсайты в качестве источника данных [26] и веб-сайты как предмет исследования [51; 15].

Примером первого подхода является следующее исследование [26]. Для того чтобы изучить субъективные жизненные реалии аутистов, были проанализированы самоописания повседневной жизни и опыта людей с этой болезнью, которые они опубликовали самостоятельно в интернете на общедоступных сайтах. В этом случае обнаруженные сайты представляют собой только средство публикации для действительно интересного контента — самоописаний аутистов. С другой стороны, могут быть предметом исследования и сами сайты. Так, исследователи [51] проанализировали почти 70 сайтов, которые предлагают лечение алкоголиков, с точки зрения эмпирического обоснования предлагаемых методов терапии. Результатом исследования стал вывод, что большая часть предлагаемых способов терапии не основывается на эмпирически проверенных знаниях и имеет сомнительное качество. Особый случай простых веб-сайтов — так называемые веб-блоги. Это личные дневники или хронологически выстроенные комментарии к индивидуально значимым темам, которые регулярно публикуются в интернете отдельными лицами или группами. Блоги обеспечивали существенный подъем в интернет-сфере до недавнего времени, правда, эта тенденция сейчас уже перекрыта популярностью социальных сетей. Тем не менее с ростом и стабилизацией популярности и распространения блогов в интернете, на них обратили внимание ученые [24; 37]. Исследователям Huffaker и Calvert [24] удалось показать в своем анализе самопрезентаций в веб-блогах, что, несмотря на анонимность, которую интернет предлагает как социальную норму, подростки в своих блогах часто разглашают личную информацию, например, их настоящие имена, возраст или место жительства. Вопреки распространенному мнению, подростки представляют себя на сайтах так же, как и в коммуникативных ситуациях офлайн.

Самораскрытие играет важную роль в интернете, поскольку другие пользователи смогут узнать ровно столько, сколько контрагент захочет открыть. Таким образом, самораскрытие — это не только самовыражение, но и форма взаимодействия онлайн и офлайн. В ки-

берпространстве стремление к самораскрытию сильнее желания остаться анонимным. Возможно, это связано с тем, что отсутствие в киберпространстве характерного для других форм коммуникации социального контекста требует больше информации для распознавания других [45]. Что касается степени самораскрытия онлайн, то открытость очень высока во время общения с пользователями, состоящими в тесной взаимосвязи. Последние склонны быть более открытыми и честными в режиме взаимодействия по электронной почте, чем в контакте лицом к лицу или по телефону. Более того, Спраулл и Кислер [47] пришли к выводу, что в онлайн-фокус-группах респонденты выражают свои мнения более честно и свободно, чем в фокус-группах лицом к лицу.

Заключение

Быстрое в последние годы распространение интернета принесло многочисленных изменения в форматы теоретических и эмпирических исследований. Сферы применения онлайн-исследований являются, соответственно, многообразными. Анализ изображений, тесты продукции, опросы потребителей, анализ паттернов коммуникации в онлайн-обучении, форумов самопомощи, а также описание социальных групп в виртуальных сообществах — лишь несколько примеров текущих тем качественных онлайн-проектов.

Как показывают многочисленные исследования, потенциал интернета как ресурса для сбора данных не ограничивается только количественным уровнем. Исследователи, работающие в русле качественного подхода, предлагают разнообразные методологические решения, чтобы использовать сеть не только как предмет, но и как способ исследования. При этом различные, присутствующие в нашем распоряжении качественные методы (опрос, наблюдение и анализ документов) лишь незначительно отличаются от своих традиционных вариантов. В частности, качественные формы опросов в интернете предлагают с учетом коммуникативной синхронности некоторую вариативность. Для онлайн-опросов не обязательно требуется одновременное соприсутствие обоих коммуникативных партнеров. Благодаря электронной почте или дискуссионным форумам

можно также получить, учитывая специфику медиа, такие же богатые объемные данные, как и в синхронных формах. Тем не менее, качественные онлайн-методы по их информативному содержанию не обязательно сопоставимы с традиционными исследованиями офлайн. Особенно онлайн-опросам зачастую не хватает спонтанности и эмоциональной невербалики в полученной информации. Наблюдения, которые могут быть сделаны в присутствии респондента, иногда порождают неожиданные идеи на основе невербальных реакций респондентов, невозможны в онлайн-интервью. Тем не менее это преимущество офлайн-интервью может и не сыграть своей роли, поскольку выразительность мимики и непосредственность реакций в целом далеко не всегда идентифицируются текстуально в транскрипте интервью.

Поэтому в онлайн-интервью, возможно, даже больше, чем в интервью лицом к лицу, от модератора или интервьюера требуются особые навыки и компетенции. Он должен основываться исключительно на напечатанном слове, заключая на его основе об аффективном и психологическом опыте респондентов, и соответственно управлять потоком разговора.

Памятуя о всех возможных ограничениях, онлайн-исследования предлагают в дополнение к организационным и экономическим преимуществам возможность достижения специфических социальных групп, которые трудно набрать для опроса обычными способами (например, потребителей наркотиков), или для исследования сенситивных тем, которые сложно открыто обсуждать о личной беседе (например, противозаконное поведение или сексуальные практики). Поэтому в целом нельзя делать утверждения за или против использования интернета в процессе социального исследования. Вопрос преимуществ и недостатков соответствующей методологии в каждой исследовательской ситуации приходится решать заново

ЛИТЕРАТУРА

1. Arbeitskreis DeutscherMarkt- und Sozialforschungsinstitute e. V. (ADM) (2001): Richtlinien fuer Online-Befragungen. www.admev.de/fileadmin/user_upload/PDFS/R08_D_07_08.pdf.
2. Al-Saggaf Yeslam, Williamson Kirsty (2004): Online Communities in Saudi Arabia: Evaluating the Impact on Culture through Online Semi-Structured Interviews. In:

- Forum Qualitative Sozialforschung, 5 (3), Art. 24, [43 Absätze]. www.qualitative-research.net/fqs-texte/3-04/04-3-24-e.htm.
3. Baker Andrea (2000): Two By Two in Cyberspace: Getting Together and Connection Online. In: *CyberPsychology & Behavior*. 3. S. 237–242.
4. Bampton Roberta, Cowton, Christopher J. (2002): The E-Interview. In: *Forum Qualitative Sozialforschung*, 3 (2), [27 Absätze]. <http://www.qualitative-research.net/index.php/fqs/article/view/848>.
5. Baur Timo, Kolo, Castulus (2001): Feldforschung in Multiplayer-Onlinespielen. Vortrag bei der German Online Research Conference in Göttingen. www.psych.unigoettingen.de/congress/ger-2001/contrib/baur-timo.
6. Bell D. (2001) *An introduction to cybercultures*. London: Routledge.
7. Bloor Michael, Frankland Jane, Thomas, Michelle, Robson, Kate (2001): *Focus Groups in Social Research*. London.
8. Bortz Juergen, Doering, Nicola (2002): *Forschungsmethoden und Evaluation*. Berlin.
9. Bryman A. (2001). *Social research methods*. Oxford: Oxford University Press
10. Cassell Justine, Tversky Dona (2005): The Language of Online Intercultural Community Formation. In: *Journal of Computer-Mediated Communication*. 10. <http://jcmc.indiana.edu/vol10/issue2/cassell.html>.
11. Chaney Michael P., Dew Brian J. (2003): Online Experience of Sexually Compulsive Men Who Have Sex with Men. In: *Sexual Addiction & Compulsivity*. 10. S. 259–274.
12. Cox Glenda, Carr Tony, Hall Martin (2004): Evaluating the Use of Synchronous Communication in two blended Courses. In: *Journal of Computer Assisted Learning*. 20. S. 183–193.
13. Crystal D. (2001). *Language and the internet*. Cambridge: Cambridge University Press.
14. Davis Mark, Bolding Graham, Hart G., Sherr L., Elford Jonathan (2004): Reflecting on the Experience of Interviewing Online: Perspectives from the Internet and HIV Study in London. In: *AIDS Care*. 16. S. 944–952.
15. Doering Nicola (2002): Personal Home Pages on the Web: A Review of Research. In: *Journal of Computer-Mediated Communication*. 7. <http://jcmc.indiana.edu/vol7/issue3/doering.html>.
16. Durham Mercedes (2003): Language Choice on a Swiss Mailing List. In: *Journal of Computer-Mediated Communication*, 9. <http://jcmc.indiana.edu/vol9/issue1/durham.html>.
17. Dzyk Waldemar (2001): *Ethische Dimensionen der Online Forschung*. In: *Koelner Psychologische Studien*. 6. S. 1–30.
18. Erdogan Guelten (2001): *Die Gruppendiskussion als qualitative Datenerhebung im Internet. Ein Online-Offline-Vergleich*. kommunikation@gesellschaft. 2. www.unifr.ch/fb03/KG/B5_2001_Erdogan.pdf.
19. Flick Uwe (2000): *Qualitative Forschung*. Reinbek.
20. Gerber Sue, Scott Logan, Clements Douglas H., Sarama Julie (2005): Instructor Influence on Reasoned Argument in Discussion Boards. In: *Educational Technology Research & Development*. 53. S. 25–39.
21. Heckman, Robert, Annabi Hala (2005): A Content Analytic Comparison of Learning Processes in Online and Face-to-Face Case Study Discussions. In: *Journal of Computer-Mediated Communication*, 10. <http://jcmc.indiana.edu/vol10/issue2/heckman.html>. Zugriff: 02.11.10.
22. Hessler Richard M., Downing Jane, Beltz Cathleen, Pelliccio Angela, Powell Mark, Vale Whitley (2003): Qualitative Research on Adolescent Risk Using E-Mail: A Methodological Assessment. In: *Qualitative Sociology*. 26. S. 111–124.

23. Hoybye Mette T., Johansen, Christoffer Tjornhoj-Thomsen. Tine (2005): Online Interaction. Effects of Storytelling in an Internet Breast Cancer Support Group. In: *Psycho-Oncology*. 14. S. 211–220.
24. Huffaker David A., Calvert Sandra L. (2005): Gender Identity, and Language Use in Teenage Blogs. *Journal of Computer-Mediated Communication*. 10. <http://jcmc.indiana.edu/vol10/issue2/huffaker.html>.
25. Jones S. (1999) *Studying the Net: Intricacies and issues* // Steve Jones (Ed.), *Doing Internet research*. Thousand Oaks, Ca: Sage. P.1–27.
26. Jones Robert S., Zahl Andrew, Huws Jaci C. (2001): First-Hand Accounts of Emotional Experiences in Autism: A Qualitative Analysis. In: *Disability & Society*. 16. S. 393–401.
27. Keski-Rahkonen, Anna Tozzi, Federika (2005): The Process of Recovery in Eating Disorder Suffers' Own Words: An Internet-Based Survey. In: *International Journal of Eating Disorder*. 37. S. 80–86.
28. Kolo Castulus, Baur Timo (2004): Living a Virtual Life: Social Dynamics of Online Gaming. In: *Game Studies — The International Journal of Computer Game Research*. 4. www.gamestudies.org/0401/kolo/.
29. Lasker Judith N., Sogolow Ellen D., Sharim Rebecca R. (2005): The Role of an Online Community for People With a Rare Disease: Content Analysis of Messages Posted on a Primary Biliary Cirrhosis Mailinglist. In: *Journal of Medical Internet Research*. 7. www.jmir.org/2005/1/e10/.
30. Lee R.M. (2000). *Unobtrusive methods in social research*. Buckingham: Open University Press
31. Mann Chris, Stewart Fiona (2000): *Internet Communication and Qualitative Research: A Handbook for Researching Online*. London.
32. McClelland Mark (2002): Virtual Ethnography: Using the Internet to Study Gay Culture in Japan. In: *Sexualities*, 5, S. 387–406.
33. Naderer Gabriele, Wendpap Marion (2000): Online-Gruppendiskussionen. Möglichkeiten und Grenzen. Vortrag am BVM-Kongress in Basel. <http://www.mediensprache.net/archiv/pubs/2948.htm>.
34. Opinionpanel (2010).Online discussion group with whiteboard. <http://www.opinionpanel.co.uk/services/online-qualitative/online-focus-groups/> Letzter
35. Panyametheekul Siriporn, Herring, Susan C. (2003): Gender and Turn Allocation in a Thai Chat Room. In: *Journal of Computer-Mediated Communication*, 9. http://jcmc.indiana.edu/vol9/issue1/panya_herring.html.
36. Prickarz Herbert / Urbahn Julia (2002): Qualitative Datenerhebung mit Online-Fokusgruppen: Ein Bericht aus der Praxis. In: *Planung & Analyse*, 1, S. 63–70.
37. Reichmayr Ingrid F. (2005): Weblogs von Jugendlichen als Bühnen des Identitätsmanagements. Eine explorative Untersuchung. kommunikation@gesellschaft, 6. www.soz.uni-frankfurt.de/KG/B8_2005_Reichmayr.pdf.
38. Reid Donna J., Reid Freiser J.M. (2005): Online Focus Groups: An In-Depth Comparison of Computer-mediated and Conventional Focus Group Discussion. In: *International Journal of Market Research*. 47. S. 131–162.
39. Reppel Alexander, Gruber Thorsten, Szmigin Isabelle, Voss Ruediger (2006): Online Laddering — Development of Innovative Laddering Data Collection Methods. Vortrag bei der General Online Research Conference in Bielefeld.
40. Rezabek, Roger (2000): Online Focus Groups: Electronic Discussions for Research. In: *Forum Qualitative Sozialforschung*. 1 (1). [67 Absätze]. <http://www.qualitative-research.net/index.php/fqs/article/view/1128>.

41. Rybas N., Gajjala R. Developing Cyberethnographic Research Methods for Understanding Digitally Mediated Identities. In: Forum Qualitative Sozialforschung, Volume 8. No. 3. Art. 35 September 2007.
42. Rifkin Jeremy (2004): Wirtschaft und Gesellschaft in einer digital vernetzten Welt. Vortrag auf den Muenchner Medientagen «Merging Media» in Muenchen.
43. Salfinger Brigitte (2005): Gender Specific Analysis of Virtual Communication in Panel Online Focus Groups. Recent Developments and Applications in Social Research Methodology. Proceedings of the RC33 Sixth International Conference on Social Science Methodology. Amsterdam.
44. Saunders M.N.K.; Lewis P. & Thornhill, A. (2000). *Research methods for business students* (2nd edition). London: Financial Times Management.
45. Short J., Williams E., Christie B. (1976). *The Social Psychology of Telecommunications*. Wiley. London.
46. Sifry David (2007): The state of the liveeb. www.sifry.com/alerts/archives/000493.html.
47. Sproull L., Kiesler S. (1986) Reducing social context cues: electronic mail in organizational communication. *Management Science* 32 (11). 1492–1512.
48. Sterne J. (1999) Thinking the Internet: Cultural studies versus the millennium // Steve Jones (Ed.). *Internet research: Critical issues and methods for examining the Net*. Thousand Oaks. CA: Sage. P.257–288.
49. Strickland Ora L., Moloney Margarete F., Dietrich Alexa S., Myerburg Stuart, Cotsonis, George A., Johnson Robert V. (2003): Measurement Issues to Data Collection on the World Wide Web. In: *Advances in Nursing Science*. 26. S. 246–256.
50. Suler John (1996): One of Us — Participant Observation Research at the Palace. *The Psychology of Cyberspace*. www.rider.edu/~suler/psyber/partobs.html.
51. Goldsmith, Lori Toll, Heather J. (2003): What Do Internet-Based Alcohol Treatment Websites Offer? In: *Cyber Psychology & Behavior*. 6. S. 581–584.
52. TomTom (2006): Tom Tom Go. www.meintomtomgo.de/forum/.
53. Xorx M. (2008) *Technolution. Wie unsere Zukunft sich entwickelt* Frankfurt. New York: Campus Verlag.
54. Wood Richard T.A., Griffith Mark D., Eatough Virginia (2004): Online Data Collection from Video Game Players: Methodological Issues. In: *Cyber Psychology & Behavior*. 7. S. 511–518.
55. Wright. Talmadge, Boria, Eric Breidenbach, Paul (2002): Creative Player Actions in FPS Online Video Games — Playing Counter-Strike. In: *Game Studies — The International Journal of Computer Game Research*. 2. www.gamestudies.org/0202/wright/.

Проблема способа проведения качественных интервью: скайп, онлайн, лицом к лицу

*Колозариди Полина Владимировна
(Национальный исследовательский университет
«Высшая школа экономики»)*

Введение

Онлайн-методы всё чаще используются на разных этапах организации исследований. Иногда речь идёт об онлайн-рекрутинге через социальные сети, нередко респондент и интервьюер вступают в переписку, чтобы договориться о контакте, вопросы и другие материалы высылаются по электронной почте участнику (или потенциальному участнику). Сами интервью также могут проходить с помощью разных онлайн-инструментов, каждый из которых обладает своими особенностями.

Условия качественного интервью предполагают, что исследователь и респондент общаются в удобной для каждого среде. Интервьюер, организуя этот процесс и одновременно ведя беседу, получает неформализованные ответы на свои вопросы. Их интерпретация происходит не как в анкете, в соответствии с продуманными вариантами ответов. Наоборот, интерпретируется сказанное, влияние исследователя не исключается, скорее, напротив: изучается диалог, то, что сказали оба участника.

Примат процесса над содержанием, конечно, имеет свои пределы, но заставляет нас внимательно относиться к условиям. Поэтому вопросы методологии качественных интервью — это во многом вопросы коммуникации. Насколько ситуация комфортна для участников? Подходит ли она для того, чтобы собеседники могли

реализовать свои цели: интервьюер мог задать необходимые вопросы, а респондент — правдиво на них ответить? Поэтому вопрос опосредования, т.е. использования различных средств коммуникации в качественных интервью неслучайный и со временем всё более актуальный. С появлением телефонов появилась возможность брать интервью по телефону, затем возникла электронная почта, онлайн-чаты и наконец, скайп и подобные ему программы для видеосвязи. Всё это используется исследователями, но вопрос в том, какие особенности появляются при использовании разных способов и технических средств проведения интервью? В частности, что происходит при применении онлайн-методов? И что, строго говоря, является онлайн-методом, а что — нет?

Для того чтобы прояснить эти вопросы, мы предлагаем классифицировать различные виды использования технологий при опосредованном проведении интервью. В этом аналитическом обзоре мы показываем, какое влияние они оказывают на результативность и в целом, какие аспекты влияния технологий рассматривают исследователи. В центре нашего внимания — индивидуальные качественные интервью.

Обзор формируется на основании изучения публикаций в ключевых журналах, связанных с обсуждением качественных методов в международной и русскоязычной среде. Выделяется несколько основных аспектов: качество исследования, проблемы конфиденциальности и приватности, этические вопросы, распределение власти и приёмы, которые используются для организации интервью (имеется в виду и организация исследования, и сам процесс: уточнение, поощрение респондента, поведение при обнаружении сбоев). Есть и специфические проблемы онлайн-методов — идентификация собеседника, анонимность и выбор оптимальных технологий.

Прагматика использования цифровых технологий разнообразна, в частности, это позволяет значительно сократить расходы на организацию исследований. Особенно важно это бывает в международных исследованиях, а также при больших расстояниях между населёнными пунктами, как, например, в России. Среди изучаемых примеров и методологий присутствуют не только непосредственно онлайн-интервью, но также телефонные. Это связано с тем, что

в современной ситуации использование интернета всё чаще¹ происходит с помощью мобильных технологий. На современных смартфонах приложения для звонков и сообщений могут использовать разные способы связи. Имеется в виду, что звонить можно (зачастую это выгодно) не только через стандартные GSM линии, но также по Скайпу, с использованием WhatsApp, Фейсбука и других подобных приложений.

Поэтому исследования онлайн-методов в качественных интервью — тема, которая меняется в результате развития технических средств коммуникации и приобретает новые значения.

История и терминология использования цифровых технологий в качественных интервью

Термин «онлайн-интервью» относится не ко всем видам опосредования технологиями. Его используют, чтобы обозначить интервью с помощью чатов, мгновенных сообщений или электронной почты. Более ранний и сейчас реже используемый синоним — электронное интервью (electronic interview), или е-интервью (e-interview). Для видеointервью с использованием приложений чаще используется название «скайп-интервью» (Skype interview), или «видеointервью» (video interview).

Онлайн-интервью — наиболее изученный вариант опосредованной цифровыми технологиями коммуникации и при этом общее название. Использование сервисов обмена сообщениями было одной из ключевых функций сетей даже до появления интернета как Всемирной сети с единым протоколом [1]. Поэтому применение этого инструмента для интервью было закономерным. Для многих респондентов это достаточно привычный и комфортный способ общения. Технически используются разные сервисы — от распространённых уже много лет назад ICQ, QIP или Miranda до современных мессенджеров (WhatsApp, Telegram, Viber и так далее). В условиях,

¹ Более 50 миллионов из 80 миллионов пользователей, по данным TNS Россия. (данные TNS Россия, 24.04.2015 http://www.tns-global.ru/press/news/344111/?sphrase_id=68633)

когда рекрутинг респондентов, а нередко и исследования полного цикла проходят с использованием социальных сетей, используются системы обмена сообщениями между пользователями. Некоторые из них представляют собой отдельные приложения (Telegram, Facebook Messenger), некоторые встроены в платформу социальной сети (ВКонтакте, Одноклассники). Также может использоваться электронная почта.

Применение видеоприложений также распространено, наиболее популярная программа — Скайп (Skype), также есть Google Hangout, Facebook Video, иногда исследователи применяют специальные приложения, которые работают онлайн или устанавливаются на компьютер. Применение специальных приложений может быть связано с необходимостью конфиденциальности, необходимостью установить качественную связь, существуют также специальные приложения для проведения онлайн-фокус-групп.

Анонимность и идентичность

В связи с использованием онлайн-инструментов возникает несколько вопросов, и в первую очередь это идентификация собеседников. Если нет камеры, невозможно проверить, действительно ли вы общаетесь с человеком, с которым планировали говорить². Но есть и другой нюанс, связанный с тем, что общение в интернете в целом даёт дополнительную свободу в конструировании своего поведения и представления себя другим. Положительной стороной является то, что это «может защитить оба лица в е-коммуникации, предлагая степень анонимности через принятие ника или дегендеризации и снижая предвзятость в отношении пола, возраста, этничности, социального статуса» [2] —, пишет Е. Рождественская, подчёркивая, что при этом проблема того, кто на самом деле даёт (и берёт, добавим от себя) интервью — эта проблема не снимается.

Возможно ли совместить визуальный контакт и общение с помощью онлайн-методов? Такая попытка предпринята в исследовании

² Строго говоря, камера тоже не даёт окончательной ясности, хотя и значительно затрудняет фальсификацию.

У. Давидович и Х. Ур. В их исследовании участвовали 199 респондентов возрастом от 14 до 24 лет [3]. Общение шло в онлайн-чате, но при этом были включены камеры. Респонденты отмечают ощущение анонимности в таком общении, но следует учитывать, что молодые респонденты склонны общаться в таком жанре регулярно. Авторы исследования подчёркивают, что сегодня эффективность таких исследований, в первую очередь в общении с молодыми респондентами, может быть велика. При этом те же молодые пользователи сомневаются в конфиденциальности передаваемой информации, опасаются, что данные могут быть перехвачены. Поскольку безопасность передаваемых через стандартные системы данных под вопросом, некоторые исследователи (коммерческие и академические) используют специально разработанные приложения. Например, такой опыт описан в статье о видеоинтервью Б. Чейни и соавторов [4], они также подчёркивают возможность записи, которая позволяет возвращаться и делать своего рода расшифровки не только текста, но также обращать внимание на выражение лица интервьюируемых. С другой стороны, это же считается недостатком, что подчёркивает П. Ханна [5]. По его мнению, исключительно «текстовый» характер интервью по телефону — это достоинство метода, так как другие варианты, включая Скайп и интервью лицом к лицу, скорее добавляют лишний³ контекст.

Ситуация с интервью в сообщениях постепенно меняется, поскольку интернет становится менее анонимным пространством, и используя, например, рекрутинг в социальных сетях, мы с большей вероятностью и общаться с собеседником будем в социальной сети. Тогда, по крайней мере, вопрос об идентичности как соответствию своему имени становится менее актуальным. Термин «виртуальный» ещё применяется для описания исследований, связанных с интернетом, но это происходит всё реже. Второй аспект, конструирование личности в онлайн-общении, напротив, по-прежнему важен. Теоретические и методологические вопросы, связанные с этой

³ Какой контекст является лишним, а какой полезным — это отдельный и нерешённый по сей день вопрос, который едва ли можно решить только анализируя онлайн-исследования. Скорее они заостряют его актуальность и добавляют новые особенности проблематизации.

темой — насколько общение в интернете, анонимное и неанонимное, предполагает свободное конструирование личности, что влияет на этот процесс. Также следует понять, как это влияет на качество интервью. В целом эта проблема относится к вопросу о том, что именно описывается респондентом в интервью: то, что он/она хочет показать или же «действительное» положение вещей. Оценивая интервью с конструктивистских позиций, мы не можем считать образ, «сконструированный» в онлайн-интервью менее достоверным, чем образ, «сконструированный» в интервью лицом к лицу. Более того, с помощью онлайн-данных можно дополнить ту информацию, которую мы получаем в офлайн-интервью, таким образом эти методы можно использовать вместе, хотя это ставит этические и методологические проблемы [6].

Качество данных

Напрямую сравнивать качественные интервью достаточно сложно. Если респонденты — разные люди, различия можно приписать их личным особенностям. Если это один и тот же человек, возможно, и он или она ведёт себя по-разному, это может быть связано не только с технологией, но и с любым другим неизвестным фактором. При небольшой выборке результаты таких сравнений всегда оказываются под сомнением. И всё же, при всех оговорках, исследования, нацеленные на сравнение методов, проводятся. Например, в работе С. Фогль были проведены 112 полуформализованных интервью с детьми разных возрастов, с каждым из которых говорили как по телефону, так и лицом к лицу. В результате выяснилось, что респонденты с готовностью соглашались на телефонные интервью, им проще говорить на чувствительные и сложные темы, но при этом они меньше вовлечены [7].

Конструктивистский подход предполагает, что критерий соответствия интервью нормам — то, что оба собеседника в процессе соконструируют одну ситуацию [8]. Например, что респондент говорит в жанре интервью, а не монолога. Или что исследователь задаёт развёрнутые вопросы, а не формализованные варианты, на которые обычно отвечают «да» или «нет».

Исходя из этого, один из ключевых критериев качественного интервью — его соответствие формату. Если речь идёт о биографическом интервью, оно не должно становиться экспертным (то есть человек использует свой опыт, а не знания о теме вообще). Существует разница между полуформализованным и неформальным, экспертным и биографическим интервью, и эти нюансы должны быть учтены, особенно, если речь идёт, например, о крупном исследовании, где работают несколько интервьюеров. При необходимости сверять результаты, полученные от разных людей. Соблюдение правил, следовательно, становится критерием научности как воспроизводимости результатов. Интервьюеру не следует полностью контролировать беседу, поэтому его задача — одновременно создавать условия для беседы и при этом задавать вопросы в соответствии с опросником. Таким образом, у него одновременно две роли: быть в своём роде «режиссёром», задающим общий сценарий и при этом — исполнителем роли задающего вопросы. В случае с онлайн-интервью это означает ответственность, в том числе и за условия интервью, обеспечение комфортной среды. Зачастую необходимо понимать, насколько респондент быстро печатает (важно для интервью через текстовые сообщения) и насколько владеет разными сервисами, можно ли предложить ему или ей альтернативные варианты. Например, в использовании видеоприложений следует уточнить, возможны ли другие варианты, кроме интервью по Скайпу (самого популярного и известного из таких приложений). Возможно, респондент предпочтёт другой сервис, например Google Hangouts или Facebook Video. С другой стороны, в некоторых странах сервисы Google не работают, и тогда нужно изучить, что доступно в конкретной стране. Такие технические нормы оказываются аналогами выбора удобного пространства в случае интервью лицом к лицу.

Отсутствие прямого контакта делает более явной проблемой коммуникативные сбои, при этом они оказываются более сложными для объяснения. Так, если в интервью с помощью мгновенных сообщений возникает пауза, неясно, каково её происхождение: это может быть «занятость, отложенный ответ, обдумывание, коммуникативный сбой, техническая проблема, отказ» [2]. В интервью лицом к лицу можно интуитивно догадаться и отреагировать в соответ-

ствии с выражением лица или общей ситуацией, но в чате собеседникам сложнее понять друг друга. Также возможно общение через системы «виртуальных миров», в таком случае через поведение персонажа можно понять что-то, что исчезает из взаимодействия, построенного исключительно на тексте [11].

В феминистском исследовании А. Триер-Бьеник также подчёркивается предположительно большая искренность респондентов. Автор приписывает это «виртуализации», тому самому свойству конструировать идентичность в процессе ответа на вопросы [12]. Другая особенность телефонных интервью (особенно в случае мобильных или радиотелефонов) подчёркивается в работе С. Хольт, она приводит примеры того, как респонденты ходили по дому с телефонной трубкой и могли оставаться в одиночестве, даже если их домочадцы заходят в дом [13]. Это имеет и обратную сторону — интервьюер не может видеть, кто находится рядом с их собеседником или собеседницей, поэтому гарантировать, что это диалог один на один, невозможно. То же самое относится и ко всем видам онлайн-интервью, даже если включена видеосвязь: «слепое пятно» оказывается очень большим. Если речь идёт о политически опасных, сензитивных темах, использование онлайн-инструментов может оказаться неэтичным и потенциально опасным.

Этика, конфиденциальность, приватность

Качественные интервью предполагают, что беседа включает развёрнутые ответы на вопросы, переходящие в «самодостаточный рассказ» [14] респондента. При этом считается, что интерпретация происходит в процессе диалога, и в таком случае важно, чтобы интервьюируемый понимал задачи исследования. Это обуславливает необходимость соблюдения дополнительных этических норм.

Этика интервью обычно включает в себя соблюдение баланса между директивностью и максимально возможным уровнем свободы высказывания для респондента, а также соблюдением дистанции, что особенно важно для сензитивных тем. Если респондент — человек из другой среды или социального слоя, необходимо нивелировать различия, сохраняя при этом дистанцию. Интервьюер

также не должен восприниматься враждебно. Общая рекомендация при этом — соблюдать правила приличия, но они могут отличаться у людей из разных сред. Это же важно для интервью с помощью он-лайн-инструментов, когда сетевой этикет в каждом случае должен соблюдаться. Например, для интервью с помощью сервисов обмена мгновенными сообщениями это предполагает, что речь приближена к устной, она не изобилует сложными конструкциями, уместными в письме.

Наоборот, в интервью по электронной почте такие конструкции будут уместны, так как способ интервьюирования позволяет располагать большим временем для ответа. В исследовании Н. Джеймс и Х. Бушер описывается, как психологи и педагоги давали интервью по электронной почте [15]. Вопросы включали как сами текстовые формулировки, так и дополнительные материалы в виде изображений. Кроме того, респондентов просили оценить опыт использования электронной почты как инструмента для интервью. Выяснилось, что рефлексивность ответов достаточно велика, возможность подумать над ответами и посмотреть визуальные материалы оценивается респондентами положительно. Сама возможность увидеть свой ответ в виде текста предполагает большее вовлечение респондента. Он или она видит свои ответы не как малоструктурированные реплики в ответ на непредсказуемые вопросы, а как цельный нарратив. Это, в свою очередь, позволяет снизить директивность исследователя, потому что респондент может отвечать на вопросы в удобном для него темпе и последовательности, обращаться к чему-то вне его непосредственных ответов «здесь и сейчас». Конечно, это подходит не для всех интервью, тем более, что «золотой стандарт» предполагает скорее формат беседы, не выходящей за рамки совместного присутствия. Но для ряда интервью, в частности, экспертных, такой вариант может оказаться более подходящим, если респондент готов потратить время и, возможно, дополнительные усилия на написание письма. Исходя из этого, можно предположить, что электронная почта может быть подходящим вариантом для экспертных интервью с образованными профессионалами, но они меньше подходят для нарративных интервью или личных тем, где нужно участие интервьюера, его или её поддержка и внимание.

Наиболее строгие этические рамки, как представляется, предлагаются в «феминистских» интервью [16]. Этот методологический подход настаивает на устранении директивности и нивелировании различия в ролях между респондентом и интервьюером. В идеальном варианте респонденту следует рассказывать и о целях исследования, и о личной мотивации интервьюера. Это особенно важно для сензитивных и личных тем. Такая беседа необходима и возможна не во всех случаях, однако в случае беседы в онлайн-среде подобные требования могут выполняться не так, как в личной беседе. Например, интервьюер может заранее рассказать о себе и исследовании, а в процессе разговора обращаться к этому знанию.

Синхронность и присутствие

Выделяются следующие свойства интервью: они могут быть согласованы по времени и пространству или рассогласованы. В онлайн-интервью с помощью скайпа и мгновенных сообщений собеседники находятся в одном времени, но разных пространствах, при переписке по электронной почте общение асинхронное. Наконец, в традиционном интервью лицом к лицу происходит синхронная беседа в одном пространстве. Само по себе это не говорит нам о качестве и других особенностях метода. Например, в обзоре Х. Дикин и К. Вэйкфилд различие в местонахождении трактуется скорее как достоинство, так как оно «увеличивает» вторжение в частное пространство респондента [17]. Собеседник едва ли будет захвачен врасплох, но какие-то детали его частной жизни могут оказаться в центре внимания интервьюера без желания последнего.

Само по себе исчезновение физического сопresутствия не только затрудняет понимание пауз и возможных сбоев, как мы писали выше, но может поставить под угрозу сам процесс интервью. Из-за проблемы со связью или по собственному желанию одного из участников соединение может прерваться без объяснения причин.

Интервьюеру в онлайн-интервью сложнее использовать стимулы, которые в обычных интервью передаются невербально — понимающе смотреть, кивать, издавать одобряющие или заинтересованные звуки. В общении в чате это необходимо делать, понимая систему

общения в этой среде (трактуя и используя смайлики, стикеры или специальные слова, если респондент их использует). В интервью через электронную почту это вовсе невозможно. В интервью через Скайп или по телефону использование невербальных способов общения может быть ограничено. Отдельно эта тема затрагивается в статье А. Ирвина, П. Дрю и Р. Сайнсбери [18]. В ней представлены результаты конверсационного анализа: на небольшом материале (11 интервью) измеряется количество и частота поощряющих реплик и звуков со стороны интервьюера. В телефонных разговорах требуется больше таких реплик, что приписывается как раз отсутствию визуального контакта. Продолжая проблематизировать пространство коммуникации, мы можем заметить и появление «третьего». Как показывает исследование автора, в общении участвуют не только интервьюер и респондент, но и «третий», которым оказывается техническое средство коммуникации, будь это компьютер, планшет или телефон [19]. Интервью оказывается одновременно общением человека с человеком и человека с компьютером. Это может вести к рассредоточению внимания обоих участников, но может и наоборот, снижать директивность.

Заключение

Метод онлайн-интервью — это двойственное понятие. С одной стороны, оно означает использование техники в интервью в качестве инструмента. С другой — непосредственное общение с цифровым интерфейсом вместо собеседника. Используя первую трактовку, следует проблематизировать также и другие используемые для интервью технологии, например, диктофонную запись. Если сужать онлайн-методы до методов, где общение происходит в онлайн-интерфейсе, интервью с помощью видеоприложений следовало бы рассматривать отдельно. В них способы коммуникации не меняются, тогда как общение через мгновенный обмен сообщениями без онлайн-инструментов невозможно.

Если мы используем относительно узкую трактовку, то понимаем, что в случае каждого интервью необходимо выбирать между разными вариантами, важными в случае каждого конкретного ис-

следования. Возможные преимущества, которые нужно принимать во внимание:

- конфиденциальность/удобство (при выборе общедоступного или закрытого способа технического опосредования);
- анонимность/искренность (в случае сензитивных или политически сложных тем);
- спонтанность/структурированность (для интервью по электронной почте).

Использование разных онлайн-методов подходит для разных типов качественных интервью. Как показывает наш обзор, можно выделить несколько типов, наиболее подходящие для следующих типов:

- интервью по электронной почте — наиболее подходящий вариант для экспертных интервью и тех исследований, где самостоятельный нарратив респондента является ценностью;
- интервью с помощью систем мгновенных сообщений — подходит для общения с молодой аудиторией и особенно в случае анонимности — общения на сензитивные темы;
- интервью посредством видеоприложений во многом заменяют интервью лицом к лицу, но интервью без видео, как и телефонные, сохраняют часть анонимности и могут использоваться для сензитивных тем;
- при этом следует учитывать, что не зная локального контекста, в котором находится собеседник, исследователь рискует конфиденциальностью данных.

Развитие методологии и технологий происходят одновременно, и на примере обзора использования качественных интервью с применением онлайн-методов, мы видим, что это развитие оказывается связанным. Вероятно, в следующие годы мы увидим развитие и решение некоторых из обозначенных проблем. Сегодня возникает несколько вопросов, которые появление онлайн-инструментов ставит перед самой методологией интервью:

- Обогащает или ухудшает качество интервью дополнительный контекст? Следовательно, должны ли исследователи стремиться сделать интервью максимально текстуализованными?

- Что представляет собой искренность или достоверность в интервью, например, может ли она трансформироваться в условиях онлайн-коммуникации?

Эти вопросы, которые для онлайн-интервью оказываются скорее методологическими предпосылками к рефлексии и выработке ясных тактик, обнаруживают сложные аспекты самого метода интервью и предполагают возможные варианты его трансформации.

ЛИТЕРАТУРА

1. Berners-Lee T. WWW: Past, present, and future // *Computer*. 1996. T.29. №10. P. 69-77.
2. Рождественская Е. Ю. Визуализация идентичности в Е-интервью (доклад) VI Всероссийская Конференция, посвященная памяти первого декана факультета Социологии Крыштановского А.О «Современная социология - современной России». М., 2012)
3. Davidovich U., Uhr H. . (2006) .Qualitative research online: Self-reported pros and cons of being chat-interviewed online with web cameras // General Online Research Conference (GOR). 2006
4. Chaney B. H. Barry A. E., Chaney J.D., Stellefson M.L., Webb, M.C. Using screen video capture software to aide and inform cognitive interviewing // *Quality & Quantity* January 2012, URL: <http://link.springer.com/article/10.1007%2Fs11135-012-9669-4>
5. Hanna P. Using internet technologies (such as Skype) as a research medium: a research note // *Qualitative Research*. April 2012. Vol. 12. 2: P. 239-242.
6. Curasi C. F. A critical exploration of face-to-face interviewing vs. computer-mediated interviewing // *International Journal of Market Research*. 2001. T. 43. №. 4. P. 361.
7. Vogl S. Telephone Versus Face-to-Face Interviews: Mode Effect on Semistructured Interviews with Children. // *Sociological Methodology* August. 2013. 43. P. 133-177
8. Jacobsson K., Åkerström M. Interviewees with an agenda: learning from a 'failed' interview // *Qualitative Research*. 2013. T. 13. №. 6. P. 717-734.
9. Веселкова Н.В. Методические принципы полуструктурированного интервью // *Социология*: 4М. 1995. № 5-6.
10. Vähäsantanen K., Saarinen J. The power dance in the research interview: manifesting power and powerlessness // *Qualitative Research*. 2012. C. 1468794112451036.
11. Salmons J. (ed.). Cases in online interview research. Sage. 2011.
12. Trier-Bieniek A. Framing the telephone interview as a participant-centred tool for qualitative research: a methodological discussion // *Qualitative Research* December 2012 12. P 630-644.
13. Holt A. Using the telephone for narrative interviewing: a research note // *Qualitative Research*. February 2010. 10. P. 113-121, doi: 10.1177/1468794109348686
14. Квале С. Исследовательское интервью. М.: Смысл. 2003. С. 301.
15. James N., & Busher, H. Credibility, authenticity and voice: Dilemmas in online interviewing // *Qualitative Research*. August 2006. Vol. 6. no. 3. P. 403-420
16. Клименкова Т. Феминистские стратегии интервьюирования и анализа данных. По страницам статьи М. Девуат // *Возможности использования качественной методологии в гендерных исследованиях: Материалы семинаров* / Под ред. М. Малышева. 1997.

17. Deakin H., Wakefield K. SKYPE interviewing: reflections of two PhD researchers // Qualitative Research. 2013. P. 1468794113488126.
18. Irvine A., Drew P., Sainsbury, R. 'Am I not answering your questions properly?' Clarification, adequacy and responsiveness in semi-structured telephone and face-to-face interviews // Qualitative Research. February 2013. 13. P. 87-106
19. Колозарики П. В. Видеоинтервью с помощью онлайн-приложений: методический опыт // Социология: Методология, методы, математическое моделирование. 2014. № 38. С. 96-127.

Раздел 7

Практикум. Опросные методы и онлайн-панели

Онлайн-панель как инструмент телевизионных измерений

*Искра Иван Сергеевич (GfK)
Кетов Сергей Валерьевич (GfK)*

Измерения аудитории телевизионных передач с помощью зрительских панелей являются основным способом исследования телевизионной аудитории во многих странах мира. Основным, но не единственным. Методы инструментального измерения телеаудитории, часто называемые пиплометрией¹, создавались прежде всего как понятный и прозрачный инструмент для продажи рекламных возможностей телеканалов (рекламная «валюта») и лишь отчасти помогают в решении задач управления контентом телеканала. Пиплометрические измерения позволяют ответить на вопрос «сколько» (сколько людей посмотрели телепрограмму), но не отвечают на вопрос «почему» (почему программа получила высокий или низкий рейтинг, как зрители восприняли телепрограмму, что им понравилось или не понравилось). Поэтому наряду с измерениями объема аудитории крупнейшие телевизионные компании регулярно проводят исследования отношения зрителей к телевизионному контенту. Этот тип телевизионных исследований получил название «оценочные опросы» (appreciation survey). Исследования зрительских оценок телевизионного контента позволяют заглянуть вглубь, понять причины успеха или неуспеха телепрограммы у зрителей.

С самого начала своей истории и до недавнего времени «appreciation survey» проводились в основном методом дневнико-

¹ Мы будем пользоваться этим не вполне точным, но распространенным обозначением практики измерения телеаудитории с помощью пиплметров — приборов, позволяющих в интерактивном режиме автоматически собирать информацию об аудитории телепередач, времени, в течение которого осуществлялся просмотр и т.п.

вых исследований или регулярных опросов телезрителей. Однако все эти методы обладали рядом существенных недостатков. Основными проблемами было то, что зрители, как правило, оценивали телепрограммы через большой промежуток времени после просмотра, и результаты таких исследований становились доступны аналитикам телеканалов с большим опозданием.

С появлением пиплометрических панелей в некоторых странах Европы были сделаны попытки получения зрительских оценок телепрограмм через специальный интерфейс пиплметра. После завершения телепрограммы на экране пиплметра может появиться просьба дать оценку телепрограмме, и респондент может ввести свою оценку (например, от 1 до 9) с помощью пульта управления пиплметром. Такие оценки отражают впечатление зрителей сразу после просмотра телепрограммы, и могут собираться и обрабатываться очень оперативно. Но это создает дополнительную нагрузку на участников панели, и ставит под угрозу качество основного измерения — рейтинговой информации. Вдобавок этот способ измерения зрительских оценок не может обеспечить требуемой глубины информации и поэтому имеет ограниченную полезность.

Быстрое распространение интернета и, как следствие, появление и развитие онлайн-методов исследований с начала 2000-х годов создали принципиально новые возможности для исследования зрительских оценок телевизионного контента. Онлайн-исследования позволяют быстро и экономично собирать большой объем информации о мнениях и оценках телезрителей, а онлайн-инструменты обладают необходимой гибкостью и способностью адаптироваться к опыту и привычкам телезрителей.

В начале 2000-х годов компания GfK, являющаяся одним из мировых лидеров в области телевизионных исследований, разработала специальную методологию изучения зрительских оценок на базе современных онлайн-инструментов, получившую название «GfK Appreciation Panel System».

«GfK Appreciation Panel System» (APS) — это онлайн-панель телезрителей, которые оценивают на ежедневной основе телепрограммы, просмотренные накануне, и дают телепрограммам развернутую количественную и качественную оценку.

Первый проект по методологии «GfK Appreciation Panel System» был начат в 2002 году в Нидерландах для крупнейшего общественного вещателя NPO [1]. Почти сразу вслед за этим — в 2004 году — онлайн-панель зрительских оценок была создана в Великобритании по заказу BBC [2].

Разработанная GfK методология «Appreciation Panel System» получила широкое распространение в Европе. Онлайн-панели зрительских оценок GfK были созданы для крупнейших общественных и коммерческих телеканалов во многих европейских странах, таких как Нидерланды, Великобритания, Ирландия, Германия, Испания, Бельгия, Словакия. С 2013 года исследование по методологии «GfK Appreciation Panel System» — проект «ТВ-Компас» — проводится в России по инициативе крупнейшего независимого телевизионного холдинга «СТС Медиа».

Таблица 1. Панели зрительских оценок GfK

Страна	Вещатель	Объем панели	Аудитория	Год создания
Нидерланды	NPO	n = 8.000	ТВ, радио, интернет	2002
Великобритания	BBC	n = 22.000	ТВ, радио, интернет	2004
Великобритания	ITV	n = 8.000	ТВ	2006
Германия	ZDF	n = 7.300	ТВ	2006
Ирландия	RTE	n = 3.000	ТВ, радио	2007
Испания	TV3	n = 1.500	ТВ	2008
Бельгия	VRT	n = 4.200	ТВ, радио	2012
Словакия	RTV	n = 1.000	ТВ	2013
Россия	СТС Медиа	n = 6.000	ТВ	2013

Предпосылки панели зрительских оценок в России

Долгое время основным препятствием для развития онлайн- исследований телевизионной аудитории в России оставался относительно невысокий уровень проникновения интернета в стране и, что

важнее, значительная неравномерность проникновения интернета в разных географических и возрастных сегментах.

По данным GfK, к началу 2015 года уровень проникновения интернета среди населения России в возрасте от 16 лет и старше составил 68%. Таким образом, аудитория интернета в России превысила 80 миллионов человек. Но проникновение интернета среди разных групп населения России было и остается очень неравномерным. Если среди молодежи и людей среднего возраста интернетом пользуются более 85%, то среди людей старшего возраста (55+) — только 25%. Наиболее высокий уровень проникновения интернета — в Москве и в крупных городах (70–80%), но в малых городах и селах этот показатель значительно ниже.

С точки зрения телевизионных исследований, существующая неравномерность проникновения интернета особенно критична для крупных федеральных телеканалов, значительную часть аудитории которых составляют люди старшего возраста и для которых важен охват всей территории страны, включая малые города и села. Но уже сейчас онлайн-инструменты вполне применимы для полноценного исследования контента телеканалов, ядро аудитории которых составляют зрители молодого и среднего возраста и стратегия вещания которых ориентирована преимущественно на крупные города. Поэтому вполне объяснимо, что инициатором создания онлайн-панели зрительских оценок в России выступил телевизионный холдинг «СТС Медиа». «СТС Медиа» управляет четырьмя телевизионными каналами в России — СТС, Домашний, Перец, СТС Love — ориентированными прежде всего на молодежную аудиторию и телезрителей среднего возраста.

Целевую аудиторию российской панели зрительских оценок «ТВ-Компас» составляют жители городов с населением от 100 тыс. и более в возрасте от 15 до 50 лет². В этой целевой группе проникновение интернета уже сейчас составляет 90%, то есть границы

² Панель «ТВ-Компас» охватывает только Европейскую часть России и Урал. Такая география панели позволяет охватить 83% населения городов 100+, избежав при этом сложностей и дополнительных затрат, связанных с различиями в сетках программ телеканалов по часовым поясам.

онлайн-аудитории максимально приближены к границам телевизионной аудитории в целом.

Рекрутирование панели зрительских оценок

Российская панель зрительских оценок «ТВ-Компас» была создана в августе-октябре 2013 года. Первоначально панель охватывала 4500 респондентов, а через год — с сентября 2014 года — объем панели был увеличен до 6000 респондентов.

Большая часть респондентов панели были рекрутированы онлайн. Для этого использовались онлайн-панели GfK и партнеров. Онлайн-рекрутирование в целом является быстрым и эффективным способом формирования выборки онлайн-панели. Но необходимо было учитывать, что достижимость разных сегментов внутри целевой группы с помощью онлайн-методов не одинакова.

По нашему опыту, наибольшую проблему для онлайн-рекрутирования представляли две группы респондентов:

- (а) тинэйджеры, которые плохо соглашались на участие в долгосрочных панельных исследованиях;
- (б) люди старшего возраста (в данном случае 40–50 лет), жители периферийных городов, люди с низким уровнем образования — т.е. те группы населения, уровень интернет-активности которых ниже и которые поэтому плохо представлены в онлайн-панелях.

Поэтому для формирования российской панели зрительских оценок использовалось также офлайн-рекрутирование респондентов, прежде всего в наиболее труднодостижимых сегментах. Офлайн-рекрутирование позволяет компенсировать дефицит некоторых групп респондентов в онлайн-панелях. Также офлайн-рекрутирование создает возможности для более тесного и доверительного контакта с потенциальными респондентами и таким образом помогает преодолеть повышенный уровень отказов от участия в панельном проекте в указанных выше группах. В целом, офлайн-рекрутирование хотя и является более дорогим способом формирования панели, позволяет улучшить структуру панели и включить в нее группы, которые невозможно достичь только онлайн-методами.

При создании российской панели зрительских оценок офлайн-методами были рекрутированы 17% респондентов (в том числе офлайн были рекрутированы около 75% тинейджеров и более 30% респондентов в группах людей старшего возраста, жителей малых городов, людей с низким уровнем образованием). Офлайн-рекрутирование используется также в процессе поддержания панели с таким расчетом, чтобы доля респондентов, рекрутированных офлайн, не опускалась ниже 15%.

Дизайн и структура панели

Онлайн-панель зрительских оценок имеет ряд особенностей по сравнению с традиционными панельными исследованиями. Основу панели составляет заполнение респондентами ежедневной анкеты, но при этом от панелистов не требуется заполнять анкету каждый день. Такое требование было бы очень обременительно (а иногда и просто невыполнимо) для большинства респондентов и привело бы к существенному снижению уровня согласия на участие в проекте.

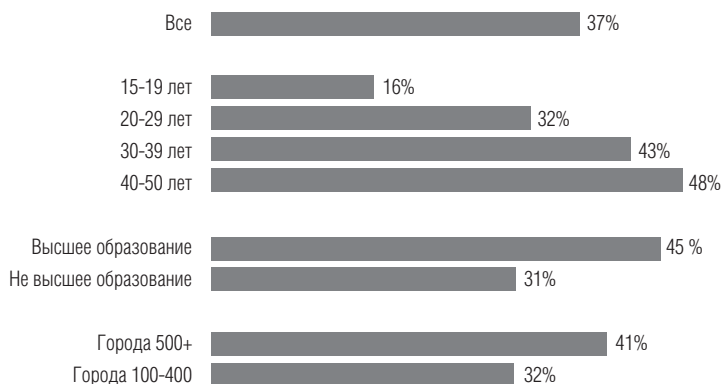
По условиям участия в проекте «ТВ-Компас» панелисты должны заполнять ежедневную анкету не менее 3-х дней в неделю. Опыт показал, что это вполне комфортный режим участия в проекте для большинства респондентов. Заполнение 13-ти ежедневных анкет в течение месяца является минимальным порогом для начисления ежемесячного вознаграждения панелистам. При этом если участник панели заполнил менее 13-ти ежедневных анкет за месяц, это не является основанием для исключения из проекта.

В среднем в день ежедневную анкету проекта «ТВ-Компас» заполняют 37% панелистов (в будние дни отклик немного выше, в выходные — немного ниже), и ежедневная выборка панели составляет около 2000 респондентов. Таким образом, панель существует как бы на двух уровнях — Gross-панель ($N=6000$), т.е. все респонденты, участвующие в проекте, и Net-панель ($N=2000$), т.е. панелисты, заполнившие ежедневную анкету в конкретный день.

Gross-панель формируется в соответствии с заданной матрицей контрольных параметров (по полу, возрасту, образованию респон-

дентов, а также с учетом региона и размера города). Но поскольку активность разных групп по заполнению ежедневной анкеты не одинакова (см. Рисунок 1), то в ежедневной выборке (Net-панель) возникают смещения, которые корректируются ежедневным взвешиванием данных (в соответствии с заданной матрицей контрольных параметров панели).

Рисунок 1. Средний процент отклика на ежедневную анкету в социально-демографических группах



«ТВ-Компас», июнь 2015

Первоначально Gross-панель формировалась в соответствии с матрицей контрольных параметров, пропорциональной структуре целевой группы исследования (расчеты были сделаны по данным официальной статистики, а также Всероссийского Омнибуса «ГФК-Русь»). По мере накопления опыта и данных об уровне ежедневной активности различных групп респондентов был рассчитан диспропорциональный план структуры Gross-панели.

Так, среди тинейджеров наблюдается самый низкий уровень активности по заполнению ежедневной анкеты: в среднем в день ежедневную анкету заполняют около 16% тинейджеров, зарегистрированных в Gross-панели. Это примерно в 2 раза меньше, чем в среднем по выборке. Соответственно, для того чтобы тинейджеры были хорошо представлены в ежедневной выборке, их доля в Gross-

панели должна быть в 2 раза выше, чем при пропорциональном распределении.

С другой стороны, респонденты старших возрастных групп демонстрируют более высокий уровень ежедневной активности (больше 40%), а значит, их доля в Gross-панели может быть несколько ниже, чем при пропорциональном распределении.

Диспропорциональная структура Gross-панели, учитывающая разный уровень активности респондентов, обеспечивает более равномерное формирование ежедневной выборки (Net-панели) и, таким образом, улучшает результаты ежедневного взвешивания.

Ежедневная анкета APS

Ежедневная анкета, которую заполняют респонденты, участвующие в российской панели зрительских оценок «ТВ-Компас», состоит из следующих основных блоков:

1. интерактивная сетка телепрограмм за вчерашний день, в которой респонденты отмечают то, что они посмотрели накануне³;
2. базовые вопросы с оценкой всех телепрограмм, которые респондент посмотрел вчера;
3. углубленные вопросы (в том числе открытые вопросы), задаваемые опционально про некоторые телепрограммы (не более чем про три телепрограммы).

Заполнение ежедневной анкеты занимает у респондентов от 1 до 3 минут (в зависимости от объема телесмотрения за вчерашний день).

Данные ежедневных анкет, заполненных участниками панели, оперативно обрабатываются утром следующего дня и предоставляются в виде баз данных и стандартных отчетов. Таким образом, «GfK Appreciation Panel System» позволяет получать зрительские оценки для широкого списка телепрограмм уже на второй день после эфира.

³ В ежедневной анкете проекта «ТВ-Компас» показывается сетка программ 8-ми федеральных телеканалов: Первый канал, Россия 1, НТВ, ТНТ, СТС, РЕН ТВ, Домашний, Перец.

Рисунок 2. Фрагмент интерактивной сетки телепрограмм в ежедневной анкете

						
18:00	18:00	18:00	18:00	18:00	18:00	18:00
18:00 Вечерние новости с субтитрами	18:15 Прямой эфир	19:00 Сегодня	18:00 Интерны	18:00 Она написала убийство	18:00 Война без правил	18:00 Уральские пельмени
18:45 Давай поженемся!	19:35 Местное время, Вести	19:40 Новая жизнь сыщика Гурова	18:30 Интерны	18:55 Одна за всех	19:00 Информационная программа 112	18:30 Уральские пельмени
19:50 "Пусть говорят" с Андреем Малаховым	20:00 Вести	21:30 Шеф	19:30 Интерны	19:00 Не родись красивой	19:30 Новости	19:00 Воронины
21:00 Время	20:50 Спокойной ночи, малыши!	23:30 Сегодня	20:00 Сашатаня	19:50 Не родись красивой	20:00 Бумер	19:30 Воронины
21:35 Дом с лилиями	21:00 Своя чужая	01:45 Слето в СССР	20:30 Физрук	20:45 Доктор Хаус	22:00 Водить по-русски	20:00 Кухня
22:35 Как избежать наказания за убийство	21:55 Своя чужая	02:40 Холм одного дерева-9	21:00 Конеди Клуб	21:35 Доктор Хаус	23:00 Новости	20:30 Кухня
01:10 Обязанная кость	22:55 Чужое гнездо	04:55 Всё будет хорошо!	22:30 Закон каменных джунглей	22:30 Рублёвка на выезде	23:25 Борджиа	21:00 Кухня
02:50 Драконий мечтунг Эволюция	23:50 Чужое гнездо	00:00 Дом-2. После заката. Спецвыпуск	23:00 Дом-2. Город любви	23:30 Одна за всех	01:30 Водить по-русски	21:30 Сонная лошадь
	00:50 Табачный капитан	01:00 Великолепная афера	00:00 Дом-2. После заката. Спецвыпуск	00:30 Отцовский инстинкт	02:00 Борджиа	23:30 Уральские пельмени
	02:45 Престиж, мама	03:15 Пригород	01:00 Великолепная афера	02:25 Быть с ним	04:00 "Территория заблуждений" с Игорем Проклепенко	00:00 Даёшь молодёжь!
						00:30 Большая разница

Как и в любых опросах, качество заполнения ежедневной анкеты во многом зависит от того, насколько полно и точно респондент вспомнит и отметит в анкете телепрограммы, которые он посмотрел вчера. Методология APS включает ряд методических инструментов, обеспечивающих качество заполнения ежедневных анкет:

- *панельный принцип исследования*: респонденты заранее знают, что необходимо будет вспомнить при заполнении ежедневной анкеты и «откладывают» это в своей памяти;
- *короткая ежедневная анкета*: не создает мотивации искусственно «укорачивать» анкету, отмечая не все просмотренные вчера телепрограммы;
- *интерактивная сетка телепрограмм*: служит удобной и наглядной «подсказкой», помогающей респондентам вспомнить все, что они посмотрели вчера.

Респонденты заполняют ежедневную анкету на следующий день после просмотра телевизора (анкета доступна для заполнения с раннего утра и до 2 часов ночи по местному времени). При этом по факту большинство респондентов заполняют анкету утром следующего дня или в дневные часы, т.е. меньше чем через сутки после просмотра, и только небольшая часть респондентов откладывает заполнение анкеты на вечер (см. Таблицу 2).

Таблица 2. Типичный график заполнения ежедневной анкеты респондентами

Часть дня	Доля заполненных анкет
Утром, до 12 часов	53%
Днем, с 12 до 18 часов	24%
Вечером, после 18 часов	23%

«ТВ-Компас», апрель 2015

Качество данных панели зрительских оценок

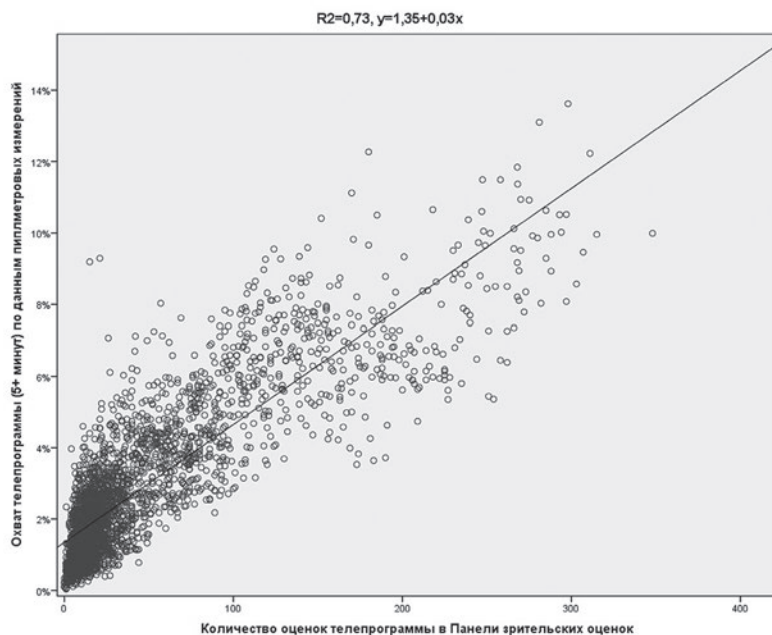
Один из ключевых вопросов, возникающих при оценке качества данных «appreciation survey», получаемых с помощью опросов (все равно — онлайн или «бумажных»), — это их сопоставимость с результатами измерений телевизионной аудитории с помощью пиплметров. Известно, что результаты опросов всегда в большей или меньшей степени отличаются от данных пассивной регистрации телесмотрения с помощью пиплметров. Для ответа на этот вопрос было проведено сравнение охватов телепрограмм в панели зрительских оценок «ТВ-Компас» с данными пиплметровой панели.

Для анализа были взяты данные о просмотре телепрограмм четырех каналов — Первый канал, ТНТ, СТС и Перец — за два месяца (март и апрель 2015 года). Из сравнения были исключены телепрограммы, которые не удалось точно сопоставить в сетках программ двух источников данных. Всего для сравнительного анализа были взяты 2827 эфирных событий (выходов телепрограмм). Поскольку по условиям заполнения ежедневной анкеты «ТВ-Компас» респонденты должны отмечать только те телепрограммы, которые они смотрели не менее 5 минут, данные панели зрительских оценок сравнивались с 5-минутными охватами телепрограмм в пиплметровой панели⁴.

⁴ Авторы статьи благодарят специалистов департамента исследований «СТС Медиа» за помощь в проведении этого анализа.

Проведенный анализ подтвердил высокий уровень корреляции ($R^2=0,73$) между количеством ответов по телепрограммам в панели зрительских оценок и охватом телепрограмм по данным пиплметровой панели (см. Рисунок 3).

Рисунок 3. Сравнение охватов телепрограмм в панели зрительских оценок с данными пиплметрии



Наряду с общим сравнением в ходе анализа проверялась версия о том, что при заполнении анкеты на следующий день после просмотра проблема «вспоминания» все же существует, и короткие телепрограммы запоминаются и отмечаются в анкете несколько хуже, чем длинные. Для проверки этой гипотезы было проведено сравнение охватов телепрограмм разной длительности. Все телепрограммы были разбиты на три группы — длительностью до 45 минут, от 45 до 60 минут и больше 60 минут. Сравнение данных из двух ис-

точников подтвердило, что чем длиннее телепрограмма, тем лучше панелисты вспоминают факт ее просмотра на следующий день:

- для программ длительностью до 45 минут — $R^2=0.67$;
- для программ длительностью от 45 до 60 минут — $R^2=0.73$;
- для программ длительностью больше 60 минут — $R^2=0.74$.

Эффективность припоминания зависит также от того, сколько времени прошло между фактическим просмотром телепрограммы и заполнением анкеты. Онлайн-методы помогают решать эту проблему — большинство респондентов заполняют анкету утром или днем на следующий день, т.е. меньше, чем через сутки после просмотра. Но все же разница в полноте и точности ответов респондентов, обусловленная временем выхода телепрограмм, остается.

В ходе анализа телепрограммы были разбиты на две группы по времени выхода в эфир — дневные (время начала с 12:00 до 17:45) и вечерние (время начала с 17:45 до 23:45). Анализ показал, что для дневных телепрограмм связь между результатами опроса и данными инструментальных измерений слабее, чем для вечерних телепрограмм. С одной стороны, это может быть обусловлено большей временной удаленностью дневных телепрограмм относительно момента заполнения анкеты — в результате программы припоминаются хуже. С другой стороны, в дневное время выше доля коротких телепрограмм, что также повлияло на результаты сравнения.

Таким образом, данные онлайн-панели зрительских оценок в целом хорошо сходятся с результатами измерений с помощью пиплметров. Онлайн-панель, с одной стороны, обеспечивает оперативный сбор информации, максимально приближенный по времени к просмотру телепрограмм. С другой стороны, удобный интерфейс анкеты помогает респондентам легко найти и вспомнить все телепрограммы, которые они посмотрели накануне. Все это позволяет минимизировать проблемы «вспоминания» более коротких телепрограмм и временного лага между просмотром и заполнением анкеты.

Базовые инструменты диагностики телепрограмм

Основная задача исследований, проводимых по методологии «GfK Appreciation Panel System», заключается в диагностике вос-

приятия телепрограмм зрителями на ежедневной основе. Для этого респондентам задается стандартный набор базовых вопросов про все телепрограммы, которые они посмотрели вчера. В российской панели «ТВ-Компас» для этого используется следующий набор базовых вопросов:

1. Appreciation Index — оценка телепрограммы по 10-балльной шкале.
2. Был ли просмотр телепрограммы запланированным или случайным?
3. Смотрел ли респондент телепрограмму полностью или частично?
4. Смотрел ли респондент телепрограмму дома или вне не дома?
5. Смотрел ли респондент телепрограмму один или совместно с другими людьми?

В разных странах базовые метрики могут немного отличаться, но «Appreciation Index» (AI) неизменно используется в качестве ключевого элемента системы диагностики. Респондента просят оценить программу, которую он посмотрел вчера, по 10-балльной шкале. Полученные по конкретной телепрограмме оценки усредняются и умножаются на 10. В результате получается индекс в диапазоне от 0 до 100, который и называют «Appreciation Index». Приведение к 100-балльной шкале делает оценку более удобной для восприятия, так как часто разница в оценках телепрограмм может быть не велика.

Таблица 3: Средний 'Appreciation Index' (AI) по телеканалам за месяц

Телеканал	Appreciation Index
Первый канал	85.6
ТНТ	85.6
НТВ	84.1
Россия 1	86.7
СТС	85.7
РЕН ТВ	84.4
Перец	88.5
Домашний	85.9

«ТВ-Компас», март-апрель 2015

«Appreciation Index», как правило, рассчитывается для телепрограмм, набравших не менее 25 оценок в конкретный день. Даже с учетом этого ограничения исследование позволяет получать ежемесячно огромный массив данных для анализа. В среднем в месяц по 8-ми телеканалам, которые охватывает исследование «ТВ-Компас», «Appreciation Index» рассчитывается более чем для 1300 выходов телепрограмм (получивших 25 и более оценок зрителей). Однако надо учитывать, что количество программ для анализа сильно варьируется от канала к каналу и зависит от его популярности (см. таблицу 1). Например, желание анализировать даже наиболее популярные программы на телеканале «Домашний» потребовало бы существенного увеличения панели.

Таблица 4: Количество эфирных событий в среднем за месяц, набравших 25+ оценок

Телеканал	Количество выходов телепрограмм, получивших 25+ оценок
Первый канал	371
ТНТ	268
НТВ	212
Россия 1	187
СТС	163
РЕН ТВ	109
Перец	42
Домашний	4
ВСЕГО	1355

«ТВ-Компас», март-апрель 2015

Важным преимуществом панельного исследования является быстрое накопление бенчмарков по AI. В результате появляется возможность сравнивать программы как в целом, так и в разрезе телеканалов, жанров, тайм-слотов и с учетом сезонности.

Дополнительные возможности APS

Помимо базовой диагностики «GfK Appreciation Panel System» предоставляет широкие возможности для более углубленного анализа реакций аудитории и проведения разного рода дополнительных исследований.

Одним из основных преимуществ методологии APS является гибкость — возможность оперативно настраивать содержание ежедневной анкеты, включая в нее при необходимости дополнительные вопросы для зрителей (или, наоборот, не зрителей) конкретной телепрограммы в определенные дни. Специальное программное обеспечение «Questionnaire Management Tool» (QMT), разработанное GfK специально для панелей зрительских оценок, позволяет вносить оперативные изменения в ежедневную анкету, а также хранит всю историю о содержании ежедневной анкеты исследования за любую дату.

Важным элементом углубленной диагностики являются открытые вопросы, в которых участники панели могут давать развернутую качественную оценку просмотренным телепрограммам. В российской панели «ТВ-Компас» открытые вопросы, как правило, задаются только про новые телепрограммы (в течение первой недели после премьеры), но при необходимости могут быть добавлены в ежедневную анкету и для телепрограмм, давно выходящих в эфир. Опыт показывает, что по новым телевизионным проектам на открытые вопросы отвечают более 90% респондентов, посмотревших телепрограмму. При этом, как правило, 50–60% ответов содержат развернутую качественную оценку просмотренных телепрограмм — чем именно понравился или не понравился зрителям новый телевизионный продукт.

Ответы на открытые вопросы в панели зрительских оценок дают богатый материал для качественного анализа телепрограмм и в некоторых отношениях могут даже соперничать с фокус-группами (сохраняя при этом все преимущества онлайн-исследования — реакция на телепрограмму в условиях реального телепросмотра, высокая оперативность сбора данных, возможность их количественной обработки). Например, за первую неделю показа на телеканале СТС сериала «Принц Сибири» (в мае 2015) было получено более 350

зрительских оценок, в том числе более 200 конкретных и содержательных ответов на открытый вопрос о том, чем понравились или не понравились зрителям первые четыре серии нового проекта.

Кроме углубленных вопросов, добавляемых в ежедневную анкету, онлайн-панель телезрителей является эффективной основой для проведения дополнительных исследований любой сложности. Дополнительные исследования на панели зрительских оценок проводятся в тех случаях, когда необходимо опросить респондента по сложной и длинной анкете, которая никак не вписывается в формат ежедневной анкеты APS.

По опыту российского проекта «ТВ-Компас», а также других аналогичных проектов GfK в Европе, предметом дополнительных исследований на базе панели зрительских оценок чаще всего становятся:

- углубленная оценка вышедших в эфир телепрограмм;
- тестирование пилотов новых телевизионных продуктов;
- тестирование промо-роликов и рекламных макетов;
- анализ эффективности эфирного и внеэфирного промо;
- имиджевые исследования телеканала;
- исследование популярности звезд телеканала;
- исследования кросс-платформенного потребления ТВ-контента.

Ощутимыми преимуществами дополнительных исследований, проводимых на базе панели зрительских оценок, является возможность формировать целевые выборки в привязке к фактическому зрительскому поведению респондентов, а также максимальная оперативность и низкая стоимость таких исследований по сравнению с обычными *ad hoc*-проектами.

Двухлетний опыт российской онлайн-панели зрительских оценок «ТВ-Компас» показал, что современные онлайн-инструменты могут с успехом применяться для изучения аудитории телеканалов. Высокий уровень проникновения интернета в городах 100+ делает онлайн-методы вполне пригодными для измерения аудитории федеральных телеканалов, ориентированных на аудиторию молодого и среднего возраста. Результаты таких онлайн-исследований по некоторым ключевым параметрам хорошо сопоставимы с данными традиционных пиплометрических измерений телевизионной аудитории.

При этом онлайн-панель зрительских оценок обладает целым рядом достоинств и возможностей, которых лишены традиционные опросные методы:

- **оперативность** — возможность получать зрительские оценки уже на второй день после эфира телепрограммы;
- **широкий список телепрограмм**, получающих зрительскую оценку на достаточно большой ежедневной выборке;
- **гибкость** — возможность углубленного изучения зрительских реакций и мнений, проведения дополнительных исследований.

Все это делает онлайн-панель зрительских оценок удобным и полезным исследовательским инструментом, способным обеспечить телеканалам постоянную обратную связь со своей аудиторией. Онлайн-панель по методологии «GfK Appreciation Panel System» позволяет оперативно и глубоко изучать мнение зрителей как о телепрограммах, уже идущих в эфире, так и готовящихся к выходу на экран, получать ответы на вопросы «почему?» (зрители приняли телепрограмму так, а не иначе) и «что можно улучшить?» (в телевизионных продуктах) и, таким образом, является эффективным источником информации для управления программной политикой телеканала и понимания зрительского поведения.

ЛИТЕРАТУРА

1. M. Hammersma, M. Appel. Internet Panel as a Tool for TV Programme Appreciation Research. ESOMAR WAM. Cannes. France. 2002.
2. N. North, S. Smith. Finger on the Pulse. How the BBC is Revolutionising Audience Relationship. ESOMAR WM3. Shanghai. China. 2006.

Поиск реальной целевой аудитории своего бренда

Ткачев Олег Викторович
(Wellhead)

Целевая аудитория. Это понятие одно из немногих, по которому в среде маркетологов и окружающих их представителей других специальностей (креативщиков, продавцов, топ-менеджеров, акционеров) существует взаимопонимание и осознание его важности. В отличие от многих других маркетинговых терминов в данном случае понимание значения этого термина у различных специалистов во многом совпадает.

И все было бы хорошо, если бы мы время от времени не встречали на улицах старичка, пьющего «Клинское», хрупкую девушку в брутальном джипе или бородача, покупающего «Даниссимо». В этот момент система «целевая аудитория» в ее классическом понимании дает сбой и вызывает недоумение: кто же на самом деле наш потребитель?

Исторически на формирование описания ЦА посредством социально-демографических характеристик большое внимание оказало медиа-планирование. Это направление деятельности является в системе маркетинга самой затратной (для крупнейших брендов), а значит, оно «тянет за собой» свою систему координат, включая удобную для нее систему описания потребителя. Для медиа-планирования социально-демографические характеристики удобны своей простотой и однозначной системой идентификации. Также в опросы часто добавляют параметры уровня доходов потребителей, однако, это вызывает определенные трудности. Например, некоторое время назад российский потребитель старательно избегал ответов на эти вопросы — накладывали свой отпечаток внутрисемейные трансферы (от детей пожилым родителям, фактическая оплата жил-

площади и питания для молодежи, живущей с родителями), потом к этой проблеме добавились кредиты. В итоге девушка-секретарь может позволить себе последнюю версию смартфона, поскольку живет с родителями и свободных денег у нее больше, чем у начальника отдела в той же компании с двумя детьми, неработающей женой, ипотекой и кредитом за машину.

Отказ от понятия «целевая аудитория» повергает менеджеров в ужас. Даже внутри одной целевой аудитории у людей различные вкусы, пристрастия, эстетические предпочтения. Если же мы начинаем работать «для всех», то ориентиры исчезают окончательно. Есть ли альтернативная классификация (не социально-демографическая) потребителей? Да, на рынке существует несколько таких подходов (RULS от КОМКОН, ряд методик компании TNS на базе M`Index), но у них две проблемы.

1. Первая: пытаясь охватить множество критериев в разных плоскостях, система размывает характер и специфику респондента. Человек вполне может быть новатором в выборе пива и консерватором при покупке сосисок. Еще больше различий обнаруживается при переходе к изучению дорогих товаров длительного пользования. В итоге мы получаем «слегка» различающиеся группы. И зачастую на их различия большее влияние оказывает вопрос, ходят ли они в фитнес-центр, нежели потребляемые марки. Когда же мы начинаем смотреть на потребителей каких-то торговых марок, в большинстве случаев различия становятся минимальными или же очевидными и не требующими сложных классификаций (например, покупки, связанные с высоким уровнем дохода и наличием свободных денежных средств на премиальные марки).
2. Вторая: точки контакта с каждой группой потребителей размываются и усложняются. Можно попробовать показать ТВ-ролик «новаторам», а наружную рекламу повесить для консерваторов, но насколько это будет эффективно? Сомневаюсь, что достаточно для такого дорогого контакта.

Не скажу, что у нас есть единственно верное волшебное средство, но у нас есть работающая система, подтвердившая свою эффективность на примере наших клиентов.

Идея состоит из двух частей.

Первое. Мы работаем только с реальными регулярными покупателями. Для FMCG продуктов человек, покупающий товар с полки и сам его выбирающий, и является реальным потребителем. Мы исключаем «потенциальных покупателей», т.к. если это особо не оговорено брифом и идеей проекта, завоевать такую аудиторию крайне сложно. Это работа либо специальных нишевых проектов, которые преодолевают барьеры, либо лидеров, расширяющих весь рынок.

Мы ограниченно исследуем «группы влияния», т.к. либо это жестко установленные правила («муж другой колбасы не ест»), которые переводят человека в категорию консерватора, либо эпизодические покупки «по просьбе».

Второе. Опираясь на задачи проекта, специфику продукта и его потребления, мы формируем массив критериев (часто с помощью качественных исследований) и по ним опрашиваем респондентов, отобранных на первом этапе. Далее полученный массив подвергается факторному и кластерному анализу, что приводит к выявлению типов потребителей.

Такая классификация уникальна для изучаемой товарной категории и, возможно, региона (в зависимости от параметров проекта).

Её ценность в том, что она более характерно и четко выявляет особенности каждой группы потребителей.

Как это связано с онлайн-исследованиями?

Есть один, но очень важный нюанс: онлайн-панель позволяет без увеличения стоимости опроса охватывать несколько населенных пунктов или регионов. В нашем случае это имеет решающее значение: фокусироваться на одном городе не имеет смысла, а традиционное поле в нескольких регионах увеличивает стоимость исследования.

Из истории наших проектов

Антикризисное исследование

Конец 2014 — начало 2015 годов. Падение курса рубля, санкции и экономическая стагнация вызывают смятение и неуверенность у наших клиентов.

Непонятны настроения потребителей.

В различных публикациях фиксируются отдельные факты падения спроса в разных товарных категориях, однако общая картина не ясна.

Решение.

Чтобы помочь клиентам разобраться с происходящими в среде покупателей изменениями, мы решили провести двухэтапное исследование, направленное на изучение изменений в потребительском поведении.

Первый этап представлял собой серию личных интервью, необходимых для того, чтобы понять, как изменилось поведение людей в кризис, на какие группы и типы они поделились.

Уже первый этап позволил сделать вывод, что кризис влияет на потребителей неравномерно. Также мы смогли описать как пострадавшие, так и незатронутые кризисом целевые группы и дать первые рекомендации владельцам брендов.

Второй этап — масштабный количественный опрос, проведенный совместно с компанией OMI (700 человек), позволил описать количественные изменения по основным товарным группам, изучить степень трансформации потребительских привычек, выявить типаж потребителей и дать их подробное описание. И, разумеется, подготовить рекомендации для брендов различных ценовых и продуктовых категорий.



Основной вывод: люди очень по-разному реагируют на кризис. На их реакцию влияет не только реальное финансовое состояние, но и эмоциональные оценки и ожидания¹.

Профилирование интернет-площадки под нужды потребителей

Агентство Wellhead и компания Молоток в течение нескольких лет совместно работали над репозиционированием имиджа сайта Molotok.ru в сознании потребителей.

Работа велась в поэтапном режиме с последовательным изучением мнения целевой аудитории и корректировкой наиболее важных составляющих имиджа в рекламных кампаниях.

В определенный момент возникло понимание, что формат работы портала может быть удобен не для всех типов потребителей. Было принято решение провести детальное изучение моделей потребительского поведения, поставлена задача выявить типы, численность и размеры расходов разных групп интернет-покупателей.

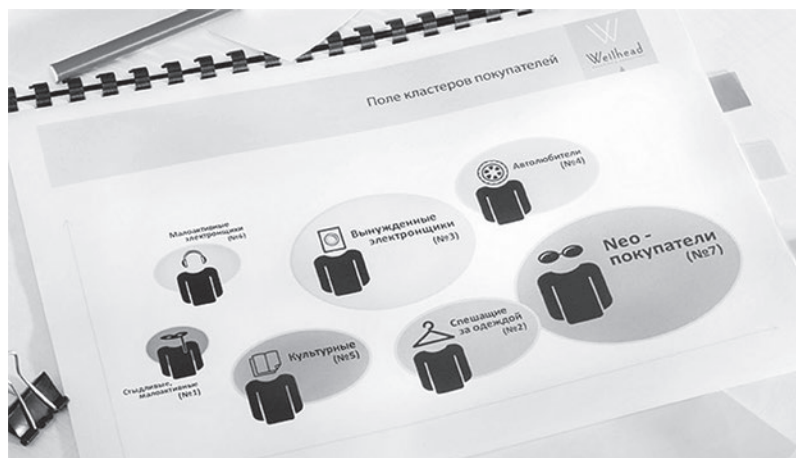
Реализация и результат проекта.

На первом этапе, в ходе качественных исследований, был сформирован максимально широкий перечень потребительских привычек, характерных для онлайн-покупок.

Данные результаты легли в основу количественных исследований, которые благодаря обработке данных с использованием различных статистических методов (частотный, регрессионный и кластерный анализ) позволили сформировать как общую картину рынка, так и выявить характерные типы потребителей. Последних удалось не только описать, но и оценить их суммарные расходы в общих объемах интернет-торговли.

Понимание структуры потребителей на рынке интернет-торговли позволило заказчику одновременно более фокусировано позиционировать проект Molotok.ru и одновременно запустить ряд новых проектов (slando.ru, brands&brands.ru, TIU.ru), нацеленных на другие аудитории, увеличив таким образом общую долю компании на рынке интернет-коммерции.

¹ С отчетами можно ознакомиться на сайте агентства Wellhead.



Увеличение объемов продаж за счет выхода на новых потребителей

В начале 2013 г. компания Ромкор обратилось в агентство Wellhead с целью усилить свои позиции на рынке Уральского региона. У производителя был широкий ассортимент продукции, известная марка, хорошие позиции на родном Челябинском рынке, а также на рынках Свердловской и Тюменской областей.

На предварительном этапе возникла идея расширения портфеля брендов.

Для оценки перспективности этой идеи был проведен масштабный количественный опрос.

Последующая обработка данных методом кластерного и факторного анализа позволили выявить 5 основных типов потребителей на уральском рынке и предложить позиционирование каждого бренда клиента на свою целевую аудиторию.

Исследование показало, что каждая группа потребителей — это особый сплав финансовых возможностей, предпочтений по ассортименту, свежести продукции, ментальности и многих других параметров. Комплексный учет интересов каждой группы, а также анализ возможностей и выгод заказчика позволили определить перспективные аудитории для двух брендов: существующего (Ромкор) и нового (Домашняя марка).



Использование методов онлайн-исследований для прогнозирования поведения потребителей в кризис

*Микаелян Инга Ваагновна,
Хитров Сергей Сергеевич
(ЗАО «Росбизнесконсалтинг»)*

В настоящее время маркетинговые онлайн-исследования не только являются важнейшим поставщиком актуальной информации относительно текущей ситуации на потребительском рынке, но также помогают лучше понимать психо-эмоциональные настроения россиян. Научившись правильно интерпретировать такие данные, их можно использовать для прогнозирования потребительского поведения в будущем, что особенно актуально в условиях кризиса, когда точные экономические прогнозы позволяют компаниям лучше планировать свою деятельность, избегать перерасхода бюджетов и в конечном итоге опережать своих конкурентов.

С помощью опросных методов можно достаточно оперативно «поймать» зарождающиеся сдвиги в потребительском поведении, являющиеся основой возникновения новых тенденций на рынке, которые определяют будущее его развития. В условиях колоссального роста использования интернета в России уже сегодня онлайн-опросы можно применять с целью изучения не только онлайн-, но и офлайн-рынков. Это делает онлайн-опросы все более и более актуальными.

Впрочем, для применения методов онлайн-опросов в качестве достоверного способа получения адекватной социологической информации, а также обоснованных выводов по результатам исследований необходимо строгое следование правилам корректного построения выборки. Важнейшим условием при этом является соблюдение условий репрезентативности выборки — в ней должны быть представ-

лены респонденты разного пола, возраста и географии проживания, причем эти пропорции должны максимально соответствовать общей генеральной совокупности всего населения рассматриваемого региона, территории или страны.

Кроме того, важнейшим условием возможности использования методов онлайн-опросов для поиска новых тенденций в потребительском поведении являются определенные правила построения социологических анкет. Прежде всего, они должны содержать блок одинаковых вопросов, которые задаются респондентам с определенной периодичностью: раз в месяц, раз в квартал или раз в год.

Поскольку условия проведения опроса, блоки вопросов и структура выборки для разных во времени социологических исследований полностью совпадают, то появляется возможность осуществлять поиск и прогнозирование трендов в потребительском поведении с помощью применения специальных алгоритмов. Заметим, что полученную информацию можно использовать для понимания будущего развития всего потребительского рынка.

РБК.research имеет большой опыт применения подобных методик. Наши социологические онлайн-опросы обычно проводятся среди жителей российских городов с численностью населения более 100 тыс. чел., где розничный рынок наиболее развит. Структура выборки подбирается таким образом, чтобы она соответствовала всей генеральной совокупности населения в городах 100 тыс.+ по полу, возрасту и географии проживания респондентов.

Мы живем в эпоху экономической турбулентности, под которой следует понимать особую сложную траекторию движения в условиях неопределенности и рисков, в которых как отдельным компаниям, так и национальным экономикам в целом приходится выживать и обеспечивать поступательное движение вперед. Действительно, уровень взаимного переплетения связей между странами на сегодняшний день крайне высок и не ограничивается географическим соседством. Благодаря интернету, социальным сетям, мобильным телефонам сегодня мы существуем в границах единого информационного пространства. В данных условиях многократно возрастают риски: экономические проблемы, непредсказуемые изменения рыночной, политической или социальной конъюнктуры одной страны

в условиях существующих крепких взаимосвязей тотчас отражаются на экономике других стран. Финансовый кризис 2008 года четко обрисовал всю механику подобных процессов. Ипотечный кризис в США стал причиной банкротства банков и привел к мировому экономическому кризису, иногда называемому «великой рецессией». Для российской экономики он привел к сокращению ВВП на 7,8%, розничный товарооборот страны упал на 5,1%.

В сегодняшних реалиях мировая экономика накопила множество неразрешенных экономических, политических и социальных вопросов, которые в любой момент могут стать спусковым крючком для нового глобального кризиса, не говоря уже о небольших потрясениях. В одной из своих статей Джон А. Каслионе, эксперт по глобальным стратегиям управления, отметил, что «глобальная коммуникация, в «мирное время» приносящая пользу всем участникам экономической деятельности, превращается в среду распространения вируса негативных настроений, как только где-нибудь происходит неладное. И этот вирус в силу эффекта домино мгновенно разносится по всему миру» [1]. В ближайшие десятилетия в этом «вихревом потоке» придется выживать и даже расти многим компаниям. Это признают не только эксперты, но и политики. Так, Владимир Путин, выступая на инвестиционном форуме «Россия — 2012», отметил, что турбулентность, нестабильность глобальной экономики приобрела длительный характер¹.

В сложной экономической реальности, когда непродолжительные периоды экономического благополучия и роста чередуются с кризисами и стагнацией, требуются новые подходы к руководству компаниями. На сегодняшний день богатейший арсенал экономико-математических методов и моделей не в силах помочь в предсказании начала кризисов, а тем более их последствий [2]. Нарушение инерционности развития экономик, неизбежно происходящее в периоды рецессии, создают серую зону неопределенности, не позволяющую экспертам в области экономики в полной мере оценить глубину кризиса, а также возможный урон для отдельных отраслей экономики. В этой связи системой раннего оповещения о возмож-

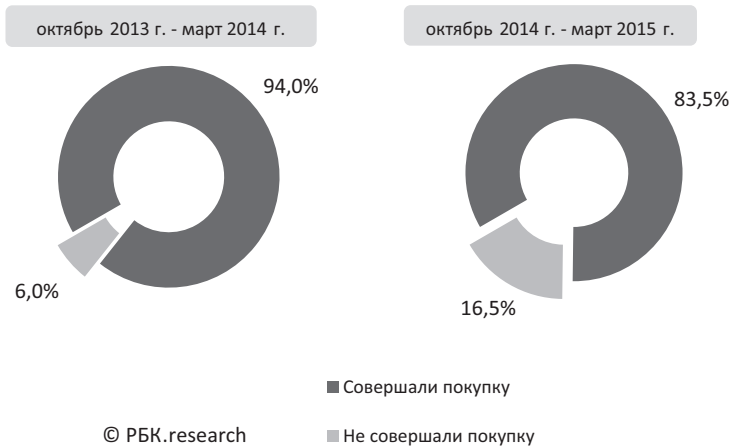
¹ <https://news.mail.ru/politics/7986731/>

ности и даже начале рецессии могут стать потребительские онлайн-опросы. Они позволяют в кратчайшие сроки произвести измерение и отследить настроения потребителей, оценив текущий уровень их потребительской уверенности. Ведь известно, что преобладание депрессивных настроений среди населения страны, а именно высокий уровень беспокойства и сомнений, неизбежно влияет на потребительские расходы, определяя как их структуру, так и склонность к сбережению. А падение спроса на «конечный продукт» сказывается на всей цепочке производства, отражаясь на индикаторах роста экономики страны в целом.

Наряду с отслеживанием настроений потребительские опросы могут носить и прогностический характер, позволяя оценить уровень падения продаж как отдельных отраслей, так и сегментов внутри них. РБК.research, аналитическое подразделение компании РБК, проводя широкий спектр исследований, накопил достаточный опыт, в том числе и в проведении подобных потребительских опросов. Так, социологические онлайн-исследования, проведенные в марте 2014 года и марте 2015 года, показали падение доли россиянок, совершавших покупки женской одежды. Интересно отметить, что эти данные были получены и опубликованы раньше, чем официальная статистика Росстата за I квартал 2015 года. Таким образом, игроки рынка смогли оперативно получить всю необходимую информацию о динамике спроса, что в условиях сложной и быстроизменяющейся экономической ситуации приобретает высокую ценность.

Наряду с изменением доли покупателей анализировались динамика среднего чека, структуры и частоты покупок. Также были выявлены потребительские настроения относительно будущих покупок женской одежды, что позволило уже в начале 2015 года оценить, на сколько процентов упадут продажи в этом сегменте, и получить достоверный и обоснованный прогноз развития рынка женской одежды на весь 2015 год. В условиях невозможности (или существенно возрастающей сложности) применения традиционных эконометрических методов прогнозирования в кризис изложенный выше подход выглядит куда более объективным, чем простые экспертные оценки развития того или иного показателя.

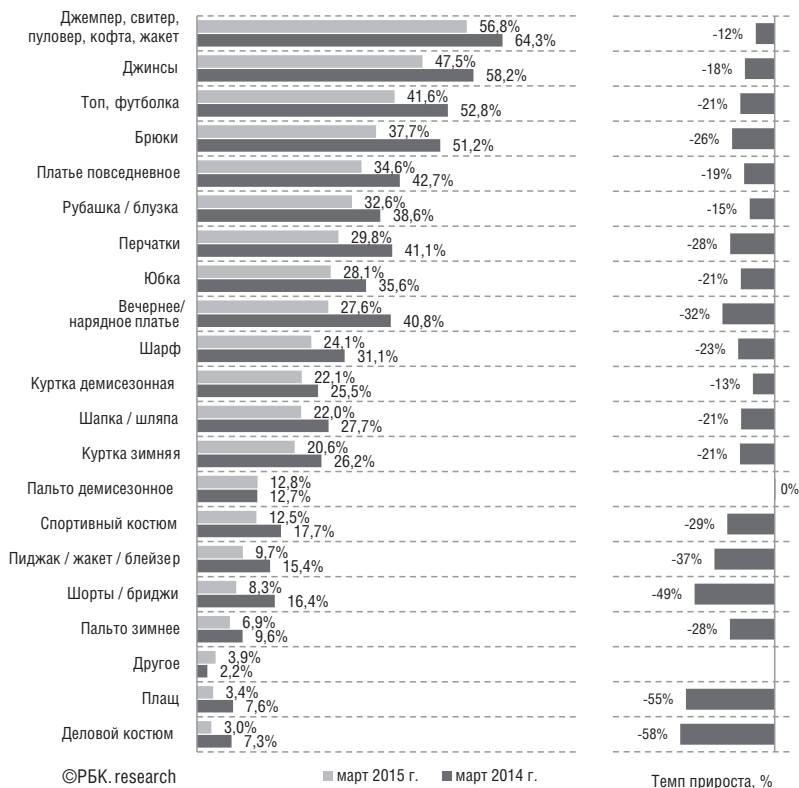
Рисунок 1. Распределение респондентов по факту покупки женской одежды в октябре 2013 г. — марте 2014 г., октябре 2014 г. — марте 2015 г., % от опрошенных женщин



Источник: исследование РБК.research «Российский рынок женской одежды 2015», дата актуализации — март 2015 года

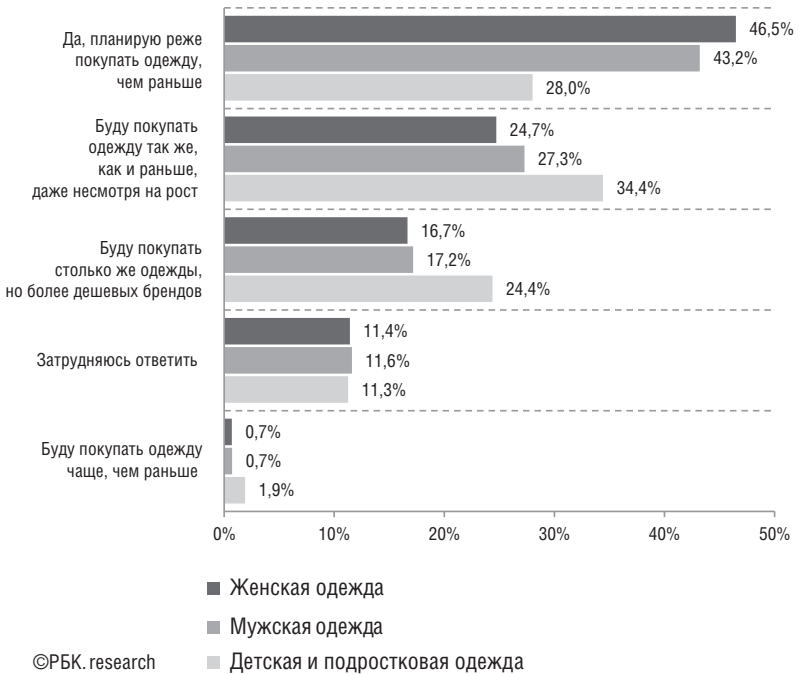
Потребительские онлайн-опросы помогают лучше понимать покупателей, выявлять их потребности и прогнозировать будущее поведение. Например, с помощью опроса можно выяснить, от чего потребители скорее откажутся в кризис. Применительно к рынку одежды нам удалось определить, что многие покупатели готовы отказаться от покупки новой одежды себе, но не своим детям. Таким образом, мы получили важный вывод о том, что рынок детской и подростковой одежды оказывается более стрессоустойчивым к кризису, поскольку родители зачастую не готовы экономить на своих детях так же сильно, как на себе. Подобная аналитика оказалась особенно полезна тем компаниям, которые работают в разных сегментах рынка одежды, включая как детскую и подростковую, так и мужскую или женскую одежду. Выводы наших исследований помогли им более эффективно перераспределить внутренние ресурсы и выбрать более правильные акценты для своего развития в кризис.

Рисунок 2. Динамика количества покупателей некоторых предметов женской одежды, март 2014 г.— март 2015 г.,
% от опрошенных женщин, покупавших женскую одежду



Источник: исследование РБК. research «Российский рынок женской одежды 2015», дата актуализации — март 2015 года. Сумма не равна 100%, так как респондент мог указать несколько вариантов ответа

Рисунок 3. «Планируете ли Вы экономить на покупках женской, мужской, детской и подростковой одежды в 2015 году?», % от опрошенных



Источник: исследование РБК.research «Розничные сети по продаже одежды 2015», дата актуализации — март 2015 года

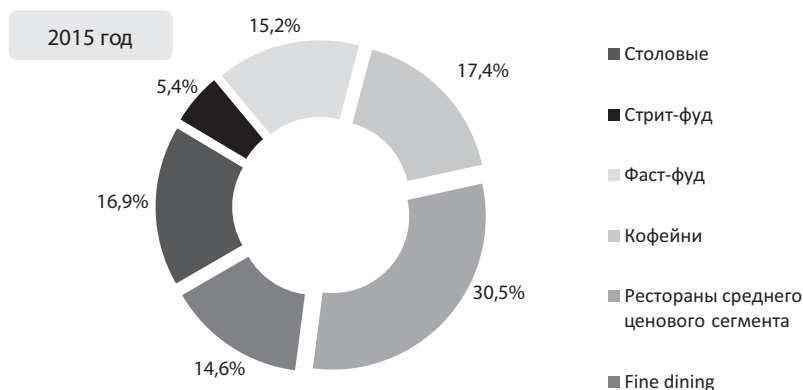
Отметим, что, к сожалению, исследователям не всегда доступна официальная статистика развития некоторых потребительских рынков и их отдельных сегментов. Например, Росстат собирает данные по рынку общественного питания в целом, не выделяя его отдельных сегментов — например, сегмента быстрого питания, демократичных ресторанов, кофеен, элитных ресторанов и других концепций. Однако на всем рынке общественного питания работает огромное количество компаний, которые специализируются

на разных сегментах общего рынка и им крайне необходима статистика по тому сегменту, в котором они представлены.

Онлайн-опросы позволяют получать такую статистику. Так, социологическое онлайн-исследование, проведенное РБК.research в мае 2015 года, позволило оценить доли отдельных сегментов в общем объеме российского рынка общественного питания. Для получения этих данных респондентам задавались вопросы о частоте посещения и сумме затрат на посещения ресторанов и кафе различных форматов.

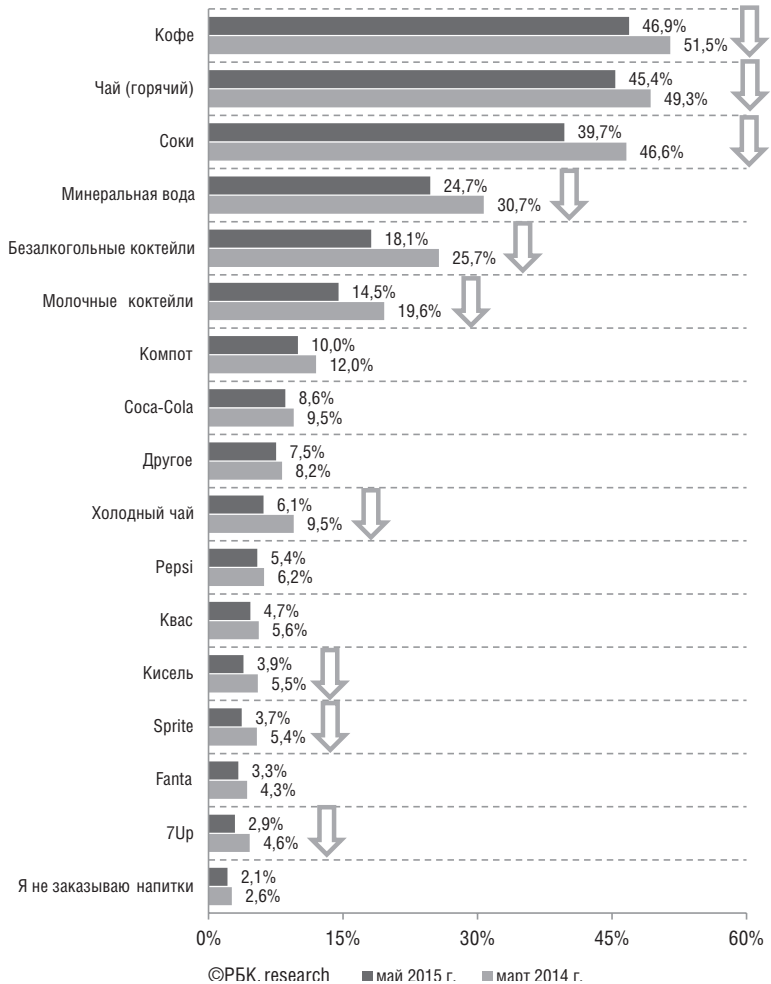
В результате были собраны данные о структуре российского ресторанного рынка, уникальность которых обусловлена отсутствием какой-либо иной статистики. Причем хочется особо отметить, что рассчитанная структура является более близкой к истинному положению дел на рынке по сравнению, например, с экспертными субъективными оценками представителей ресторанной индустрии, которые часто оказываются не в состоянии оценить картину сразу в нескольких сегментах рынка общественного питания.

Рисунок 4. Структура российского рынка общественного питания, 2015 г., %



Источник: исследование РБК.research «Российский сетевой рынок общественного питания 2015», дата актуализации — май 2015 года

Рисунок 5. «При посещении ресторанов с обслуживанием какие безалкогольные напитки Вы чаще всего заказываете?», 2014–2015 гг., % от опрошенных посетителей ресторанов с обслуживанием



Источник: исследование РБК.research «Российский рынок ресторанов среднего ценового сегмента 2015», дата актуализации — май 2015 года

Наряду с этим социологическое исследование помогло найти ответы на вопросы о том, как российские потребители адаптировались к кризису. Так, в 2015 году многие россияне при посещении ресторанов и кафе начали экономить на безалкогольных напитках. Данная статистика позволила нам сделать вывод о том, что россияне в условиях экономического кризиса и падения собственных доходов стали оптимизировать свои затраты — например, заказывая один коктейль на двоих и т.п. Интересные результаты были получены и при анализе заказов алкогольных напитков, а также всевозможных кулинарных блюд.

На основе полученных данных ведущим российским рестораторам был предоставлен ряд рекомендаций относительно изменения своих меню под новые нужды российских потребителей.

Таким образом, онлайн-опросы:

1. Могут использоваться в качестве барометра текущего восприятия людей относительно состояния экономики и их собственных потребительских настроений.
2. Позволяют очень быстро и качественно «ловить» изменения в потребительских настроениях и ожиданиях.
3. Дают возможность оценивать глубину падения самых разных рынков и их отдельных сегментов в сложные экономические периоды еще до появления официальных статистических данных.
4. Помогают выявлять наименее стрессоустойчивые сегменты потребительского рынка и поиск возможностей для адаптации ассортимента игроков рынка с целью максимального удовлетворения изменившихся предпочтений потребителей.
5. Позволяют выявлять зарождающиеся сдвиги в потребительском поведении, что дает возможность успешно прогнозировать развитие спроса в будущем.

ЛИТЕРАТУРА

1. Philip Kotler, John A. Caslione. Chaotics: The Business of Managing and Marketing in the Age of Turbulence. // AMACOM. 2009. P.150–200.
2. Писарева О. М. Методы прогнозирования развития социально-экономических систем. М. 2007. С. 20–2

Раздел 8

Практикум. Безопросные методы: Big Data, аналитика социальных сетей

Большие вопросы — большие ответы

Выручит ли индустрию вещания обуздание «умных данных» об аудитории?

Том Вейс (Genius Digital)

Франк Тюрп (GfK)

Ник Вееш (GfK)

Большие данные на данный момент всё ещё пребывают в младенческом возрасте. Различные компании предлагают всё больше и больше сервисов, предполагающих подписку или регистрацию, работающих на разных устройствах и собирающие горы данных об их пользователях и аудиториях. Но по большей части потенциал этих данных остаётся неиспользованным. Однако же благодаря тем открытиям, что несут нам Большие Данные, индустрия вещания может кардинально измениться. Вещательные компании, создатели контента и рекламодатели могут использовать Большие Данные для составления прогнозов, которые в корне изменяют бизнес-модели и направление денежных потоков. Согласно Джонатану Сайксу (Johnathan Sykes), «главное — понять, что поведенческие данные могут помочь в любых областях бизнеса, от бизнес-стратегии до менеджмента операций, от рекламы до привлечения клиентов и контроля оттока клиентов».

В этом докладе ключевые менеджеры и управленцы из 14 медиagrupp в формате интервью рассказывают о преимуществах использования Больших Данных в индустрии вещания. Совокупное число подписчиков этих компаний — более 70 миллионов, а ежедневный охват — около миллиарда человек.

Астро (Astro), Ченнел Фор (Channel 4), Диджисел (Digicel), Джиниус Диджитал (Genuis Digital), ГФК (GfK), Либерти Глобал Интернешнл (Liberty Global International, LGI), Мейджин

Тиви (Magine TV), Максдоум (Maxdome), Оранж Франс (Orange France), OCH (OSN), Скай АйКью (Sky IQ), ЮЭфЭй (UFA), Верайзон (Verizon) и Виакон Интернешнл Медиа Нетворкс (Viacom International Media Networks, VIMN)

Несмотря на то, что современные медиа и индустрия развлечений находятся на передовой этих изменений, путь к использованию Больших Данных ещё неблизкий. Согласно Нико Вешче (Niko Waesche), главе департамента Медиа и индустрии развлечений в ГФК, «все бизнес-модели, строящиеся вокруг данных, до сих пор находятся в той фазе мышления, когда мы отталкиваемся от «своих» данных, а не пытаемся выстроить целостную экосистему данных».

Филипп О'Феррал (Philip O'Ferral), старший вице-президент мероприятия Виакон Интернешнл Медиа Нетворкс (Viacom International Media Networks, VMN) утверждает, что «нам нужно понять: данные так же важны, как и превосходный контент, и эти два мира необходимо свести вместе».

Данная статья откроет путь к определению и разработке последовательной и успешной стратегии в подходе к Большим Данным всем тем, кто изучает их потенциал и возможные приложения.

Бизнес Больших Данных

Большие данные всё чаще находят признание как инструмент для стратегического планирования и управления операциями компаний. Именно это и станет важным различием между вещательными сервисами в условиях жёсткой конкуренции, когда каждое решение имеет критическое значение.

Чтобы использовать Большие Данные, необходимо дать ответ на несколько вопросов. Какие рабочие структуры необходимо для этого организовать? Каким данным можно доверять? Какие способы анализа использовать? Определённо, при ответе на эти вопросы следует учитывать прежде всего потенциальную пользу этого мероприятия и реалистичность времени внедрения технологий. Несмотря на то, что сегодня основным источником данных являются панели, созданные независимыми от отрасли компаниями, успех в новом мире Больших Данных потребует сочетания традиционной

модели медиарейтингов и информации о поведении аудитории, получаемой в реальном времени. Последняя включает в себя ответы на вопросы о том, как, где и какой именно контент был просмотрен на каждом из используемых аудиторией устройств.

Тем временем привычки телесмотрения меняются: зрителями востребованы гибкость в просмотре программ и персонализация услуг. Также растёт и озабоченность тем, как используется «цифровой отпечаток» потребителей медиаконтента, данные об их поведении. Оба эти фактора дают представителям теле- и медиаиндустрии все основания по-новому посмотреть на схемы использования данных.

Многие влиятельные игроки медиа и телевидения уже поделились тем, как Большие Данные влияют на их стратегии и внутренние операции. Так, ценность Больших Данных вне вопроса, однако определение, сбор и использование данных как сулит большие возможности, так и чревато сложностями.

Почему Большие Данные и почему именно сейчас?

Поиск верного применения Больших Данных становится обязательным атрибутом любой компании в мире телевидения, и такие попытки вызывают настоящие цепные реакции внутри индустрии. От вещательных компаний, платных каналов и стриминговых сервисов до поставщиков аналитических систем и инструментов, — и даже до человеческих ресурсов, необходимых для интерпретации данных, — кажется, вопрос о том, какие данные действительно полезны и как их применить, беспокоит каждый уголок индустрии.

«Зачастую Большие Данные считают панацеей. Однако в действительности наиболее заметный толчок для развития бизнеса провоцируют не красивые грандиозные идеи, а общий рост культуры понимания данных», — утверждает Том Вайс (Tom Weiss), генеральный директор Джиниус Диджитал (Genius Digital), компании, занимающейся анализом данных, в которую инвестирует GfK.

Применив анализ данных для одной рыночной сферы, Джиниус Диджитал обнаружили целый социодемографический сегмент, заинтересованный в просмотре трансляций матчей по регби. Местный оператор вещания даже не подозревал о том, что этот сегмент

может интересоваться регби. Переубеждённый данными, оператор перестроил маркетинговую схему, чтобы захватить этих зрителей.

Поистине тип и качество собираемых и анализируемых данных чудесным образом преобразуются. Не так давно первые лица в индустрии говорили о «данных, исходящих из активов» — базовых фактах, таких как число подписчиков или число показов за определённый период. Сейчас же собираемые данные уходят от «основанных на активах» в сторону данных о поведении: от данных панелей к индивидуализированным данным значительно большего масштаба, которые собираются в режиме реального времени.

«Шесть месяцев назад мы занимались самими основами. Телевизионные рейтинги, статистика по веб-сайтам и ведущим социальным медиа,— говорит Филип О'Феррал, старший вице-президент компании Viacom International Media Networks (VIMN).— Теперь же у нас есть команда бизнес-аналитиков, и её глава — доктор наук в математике. Каждая цифра, которую они добывают, повышает ценность нашего контента».

«Как владельцы, издатели и вещатели медиаконтента, мы должны признать данные не менее ценными, чем отличное информационное содержание. Мы должны свести эти два мира вместе, при этом аудитории не стоит беспокоиться о данных. Вероятно, мы в VIMN зашли далеко в стремлении увеличить непосредственные продажи, совершая сделки с компаниями вроде Нетфликс по всему миру. Это по-прежнему важно, но все усилия тщетны без резонанса вокруг наших брендов. Контент — это катализатор потока данных».

В 2014 году Астро (Astro), ведущая малазийская компания в сфере платного телевидения, предприняла ряд важных шагов на пути к накоплению и использованию Больших Данных. Одним из них является проприетарная система измерения аудитории, собирающая данные с абонентских телеприставок и агрегирующая просмотр отложенных трансляций, видео по запросу и интерактивного телевидения.

«Я была бы рада сказать, что мы хорошо управляемся с нашими данными, но, к сожалению, не могу, несмотря на то, что наши усилия сейчас устремлены в этом направлении,— говорит Роана Рожан (Rohana Rozhan), генеральный директор Астро.— Мы развиваем-

ся и учимся. Мы поддерживаем отношения с нашими клиентами, которые проживают в четырёх миллионах домохозяйств по нашей базе, что уже очень хорошо». Рожан утверждает, что Астро начинают судить о Больших Данных в более «организованном и целостном» ключе. «Мыслить в терминах домохозяйств уже нельзя: в действительности данные индивидуализированы и рассредоточены по ряду используемых аудиторией устройств».

Ключевая задача бума Больших Данных в компании Астро — повысить успешность использования данных для заключения сделок: удержание клиентов, продажа расширенных пакетов услуг текущим клиентам. Впрочем, Астро не собирается останавливаться на этом и планирует сделать данные сердцем всей операционной деятельности. «Мы делаем упор на психографические, демографические и поведенческие аспекты данных» — говорит управляющий директор по производству Астро Генри Тэн (Henry Tan).

Координирование различных источников данных крайне важно. Например, Астро создаёт две панели для отслеживания поведения зрителей. Первая обеспечивает их инсайтами, используя 70 000 городских домохозяйств высокого достатка, которые представляют «наиболее ценного клиента», в то время как вторая панель насчитывает 4200 домохозяйств и репрезентирует всё множество подписчиков Астро. Последняя панель позволит маркетологам таргетировать рекламу не только по интересам просмотра аудитории, но и по их профилю потребления товаров FMCG.

Рауль Таппа, вице-президент департамента анализа данных в Астро, дополняет: «В сегментированном, подвижном мире контента Большие Данные открывают ценные сегменты зрителей, которые ранее обобщались в более крупные сегменты или вообще не были измерены привычными системами оценки. Большие Данные позволяют оценить успех нашего контента для каждого формата, платформы и экрана. Наша команда может наслаждаться превосходной разрешающей способностью, которую нам принесли данные, получаемые от устройств просмотра. Теперь мы можем разработать лучшие решения в области контента и рекламных моделей, повысив эффективность как своего бизнеса, так и бизнеса наших клиентов из мира маркетинга».

Дело не только в наращивании объёма данных; скорее в том, чтобы понять, как люди используют ваши сервисы, например, когда и с каких устройств происходит просмотр. Для крупных вещательных компаний, которые раньше могли лишь незначительно взаимодействовать с аудиторией, это много меняет.

«Я глубоко убеждён в том, что единственный повод к сбору данных — сформировать более чёткое понимание предметной области, чтобы принимать лучшие решения», — говорит Тони Муни (Tony Mooney), управляющий директор Скай Айкью (Sky IQ) подразделения Скай Тиви (Sky TV). «Прежде чем собирать гору данных, необходимо сформулировать, какие решения на их основании вы хотите принять, и какова ценность этих решений. И абсолютно естественно, что мне немного сложно с аналитиками Больших Данных, которые говорят: «Мы можем собрать кучу всего отовсюду, сложить воедино, и всё получится. Я не сторонник такого подхода».

«Как часто происходит в мире данных, секретный ингредиент — это обработка и интерпретация данных. Когда на вас буквально направлен пожарный шланг, фонтанирующий информацией, очень легко утонуть в абсолютно нерелевантных деталях», — добавляет Тони.

С ним соглашается Джонатан Сайкс (Jonathan Sykes), бывший глава Би-Скай-Би (BSkyB) и сооснователь и исполнительный председатель Джиниус Диджитал (Genus Digital). «В настоящий момент доступно море данных, но больше — совершенно необязательно лучше. Сбор абсолютно всех доступных данных усложняет доступ к ним, в то время как понимание только поведенческих данных может помочь во всех областях бизнеса, от бизнес-стратегии и производственных операций до рекламы и от привлечения клиентов до контроля оттока клиентов».

Как повысить эффективность рекламы

Многие частные вещательные компании используют Большие Данные в первую очередь для того, чтобы повысить продажи рекламы. «Рекламные продажи на развитых рынках — величина известная, на них легко равняться», — говорит Том Вайсс из Джиниус Диджитал.

В Великобритании Скай (Sky) и Ченнел Фор (Channel 4), пользуясь своей стратегической компетенцией, одними из первых использовали Большие Данные, чтобы усовершенствовать текущие бизнес-процессы и повысить выручку. Они сначала реализовали подобную модель для видео по заказу с расчётом реализовать её и для кабельного телевидения позже.

В данный момент Ченнел Фор благодаря Большим Данным получает пятнадцать процентов от продаж рекламы, таргетированной по демографическим характеристикам в сегменте видео по заказу. Ожидается, что в течение двух лет эта цифра вырастет до 50%. Сандживан Бала (Sanjeevan Bala), глава планирования и аналитики данных в Ченнел Фор, утверждает, что такие изменения повлекли «рост выручки на 30–55%, в зависимости от рекламного продукта».

«Новые рекламные продукты будут более эффективны и экономичны для брендов и рекламодателей, которые стремятся воспользоваться более точным таргетингом аудитории видео на заказ».

Фокус в том, чтобы использовать Большие Данные для проектирования и создания того контента, который будет действительно интересен конечным потребителям. В конечном итоге это поможет привлечь больше рекламодателей. Тэн называет это стратегией «множества победителей». «Реклама — это попытка угадать, какая лошадь придёт первой, поэтому мы стараемся дать рекламодателям как можно больше лошадей-победителей».

Некоторое время назад было тяжело найти аргументы в пользу Больших Данных: зачастую хранение данных было дорогое и сложно и до недавнего времени не хватало инструментов для их анализа. Также было тяжело сопоставить различные виды данных, чтобы получить непротиворечивое и практичное целое, при помощи которого можно было бы повысить прибыль. Однако изменения в бизнес-модели вещания — усиление прямого взаимодействия между вещателем и аудиторией — сделали Большие Данные жизненно важными для бизнеса.

Для компании Виаком, владельца MTV и Nickelodeon, брендов с юной аудиторией, Большие Данные стали основным инструментом бизнес-аналитики, позволяющим затронуть одну из самых демографически сложных и трудноуловимых аудиторий.

«Все дороги ведут к хранилищу данных — именно это оказывает постоянное влияние на нашу компанию», — объясняет Филип О'Ферралл (Philip O'Ferrall), старший вице-президент Виаком Интернешнл Медиа Нетворкс (Viacom International Media Networks), ВИМН. «Будь то просмотр вебсайта, рейтинг телепрограммы или потоковое видео по запросу — мы отчётливо видим, какие единицы контента работают лучше. Это помогает нам точнее принимать решения о запуске тех или иных передач и инвестициях, поддерживать определённые таланты».

О'Ферралл из ВИМН считает, что всё дело в «смеси» данных: «Мы используем смесь из сырых данных об использовании мобильных приложений и мобильного веб-сёрфа, о том, кто и что смотрит, когда он это делает и как долго. Это, безусловно, поведенческие данные. А затем мы добавляем вспомогательные данные из сторонних источников, таких как Фейсбук, Твиттер, Youtube, а также данные с телеприставок. Но также мы используем и традиционные данные, такие как данные панелей и фокус-груп. Несмотря на то, что сами по себе данные никогда не смогут дать ответы на все вопросы, всё же следует к ним относиться как к всестороннему инструменту».

Скай АйКью собирает данные о поведении потребителей на всех платформах Скай, включая онлайн-платформы, отслеживает профили подписчиков и их взаимодействия. Они также объединяют свои данные с данными сторонних поставщиков, включая публично доступную статистику, социодемографию и данные о розничной торговле. Также нередко партнёрства, например, Скай Медиа работает с исследовательским агентством Даннхамби (Dunnhumby) из Лондона, которое в настоящий момент принадлежит Теско (Tesco). Они отслеживают покупки 17 миллионов держателей клубных карт Теско.

Основной источник данных Скай АйКью — 500 000 домохозяйств, ежедневно предоставляющих данные, в то время как БАРБ (BARB, национальный измеритель аудитории вещания в Великобритании) отслеживает всего 5 000 домохозяйств в Великобритании. За последние 18 месяцев данные Скай АйКью нашли применение у нового внутреннего клиента, программная дирекция каналов Скай, включающая развлекательные, новостные и спортивные ка-

налы. «Коллеги из программной дирекции каналов и со стороны создания контента раньше использовали БАРБ в качестве основного источника данных. При помощи него они формировали расписание программ и оценивали успех передач»,— говорит Тони Муни, управляющий директор Скай АйКью. «Теперь у них появился дополнительный источник данных значительно большего масштаба. С его помощью можно проводить намного более глубокий анализ контента. Впервые появилась возможность провести связь между поведением подписчиков при просмотре программ и их взаимодействием и операциями с нами. Таким образом удаётся построить намного более строгие зависимости между тем, что люди смотрят и тем, за что они платят».

Муни считает, что компаниям нужны базовые социодемографические данные, включая состав домохозяйства, уровень дохода, готовность тратить, история платежей и взаимодействий. «Но на самом деле вам необходимы частота и «ожидаемость»: как часто и когда в следующий раз. Пока мы не занялись трекингом просмотров, у нас не было возможности построить полную картину пользовательского поведения и потребления, поэтому это усовершенствование было исключительно мощным прорывом».

В 2014 году ПроЗибенСат 1 (ProSiebenSat.1) положили Большие Данные в основу корпоративной программы, расширив её на новый телевещательный бизнес Максдоум (Maxdome), чтобы собрать данные от различных подразделений в соответствии с постановлением о защите данных.

«Максдоум всегда занимались бизнес-аналитикой на производстве, однако эта инициатива радикально усовершенствовала наши системы. Мы увеличили число поставщиков данных и нарастили команды для работы с ними»,— говорит Лесли МакКензи (Lesley MacKenzie), генеральный директор Максдоум, компании, принадлежащей ПроЗибенСат 1. «Мы отслеживаем анонимные данные о поведении клиентов и используем прогнозирование в контроле оттока клиентов. В то же время мы формируем понимание прошлых взаимодействий пользователей благодаря ретроспективным данным. Таким образом нам удаётся выстроить индивидуальный CRM-маркетинг с узко-таргетированными сообщениями».

После перезапуска бренда и обновления контента в 2014 году Максдоум, сервис компании ПроЗибенСат, вырос на 40% по уровню пользования.

Дейв Буторак (Dave Butorac), генеральный директор ОуЭсЭн (OSN), ведущего поставщика платного телевидения на Среднем Западе, активно продвигал использование Больших Данных внутри компании в силу ненадёжности и неполноты данных о телевещании. Он утверждает, что большинство данных телерејтингов на Среднем Западе вызывают только смех.

ОуЭсЭн увеличил затраты в четыре раза на исследования и анализ получения данных за последние четыре года. «Мы тратим многие миллионы на данные ежегодно. Такие затраты оправдаются улучшением торгового предложения и ростом базы подписчиков», — говорит Буторак.

Помимо этого ОуЭсЭн также уже четыре года проводит трекинг-овое исследование с опросом раз в два месяца. Целевой группой являются как клиенты ОуЭсЭн, так и люди, не пользующиеся их услугами. В конце 2014 появился доступ к данным с телеприставок 80% базы подписчиков компании, насчитывающей 1.1 миллиона человек — в основном, на рынках, где распространено широкополосное соединение, включающим такие страны как Кувейт, Саудовскую Аравию, ОАЭ и Катар. Эти новые данные помогают составлять вещательную сетку, принимать решения по маркетингу и составлению программ.

«В современном контексте нам также необходимо понимать, в какое время дня люди нас смотрят, так как одного-единственного «прайм-тайм» больше не существует. Также важно понимать использование разных экранов. Какова связь между различными устройствами, и когда подписчик переключается на планшет?» — говорит Буторак.

Для компаний, производящих контент, привлекательность Больших Данных состоит в том, что с их использованием можно внедрять новые бизнес-модели. Эрнст Фейлер (Ernst Feiler), глава технологии драматических сериалов в ЮЭфЭй (UFA) в Германии, говорит, что контакт с аудиторией для вещателей является новым, однако для всего бизнеса создания телепрограмм и всего телевидения это важнейшее условие здоровья отрасли.

«У нас нет опыта работы с данными, и придётся учиться. Вещательные компании умеют обращаться с данными, но связываются с ними только после того, как мы им предоставляем готовый продукт. Это не самый разумный путь дальнейшей работы. Мы считаем, что необходимо сделать процесс производства более цифровым, так, чтобы на выходе мы располагали видео со всеми возможными данными о нём. Это поможет в создании новых бизнес-моделей».

Филип О'Ферралл, старший вице-президент Виаком Интернешнл Медиа Нетворкс, считает, что индустрия уходит от беспрекословной веры телевизионным рейтингам на основе панелей. В качестве примера он приводит ЭмТиВи'с Хоттест (MTV's Hottest), мероприятие, спровоцировавшее 330 миллионов взаимодействий в Твиттере и Фейсбуке. «Это (активность в социальных медиа) происходило вне наших вещательных сетей, однако если мы загоним всё это в аналитический дэшборд, то сможем осмыслить с точки зрения финансовой выгоды и обратить в выручку». В 2014 году ЭмТиВи Хоттест стала второй по успеху активацией после Кубка мира в Бразилии.

Будущее Больших Данных

Мощь Больших Данных огромна. Очевидно, что вещательные компании и операторы затронули только малую часть того, что возможно в принципе.

Это заставляет задаться вопросом о необходимости непрерывного участия человеческого фактора. «В сегодняшнем мире машины разговаривают с машинами, а скорость определяет всё, а это значит, что на самом деле миром правят алгоритмы, которые вольны решать будущее без участия человека», — говорит Тони Муни из Скай АйКью. — Поэтому главный вопрос в том, какую часть этого машинного принятия решений мы сможем подкрепить данными? Ответить на этот вопрос — это следующая большая задача, стоящая перед нами».

Понимание того, как лучше использовать Большие Данные, также находится на весьма ранней стадии развития: будет придумано множество прекрасных приложений Больших Данных, до которых пока никто не додумался. Впрочем, и новые препятствия тоже появ-

вятся. В то время как должный анализ Больших Данных определённо становится всё более значимым для телекомпаний и не выполняется в исследовательских департаментах, актуальным остаётся вопрос, кто должен выполнять этот анализ и как исследовательская деятельность рассеяна внутри организации.

С этим связана ценность аналитиков больших данных, или data scientists, тех людей, которые могут надлежащим образом проводить анализ данных, получая результаты, имеющие высокую практическую ценность.

Другое ключевое направление — совмещение данных о телесмотрении с данными из других источников. Здесь наиболее преуспели такие компании, как Скай, которые используют смешение данных, полученных с телеприставок и данных об использовании карт лояльности. Ключевые направления — это устранение преград, оберегающих проприетарные данные, и защита данных от искажений, которые нам уже приходится наблюдать в совсем молодом, неоперившемся мире аукционов рекламы в реальном времени, или RTB, где алгоритмы, более известные как «боты», «просматривают» рекламные сообщения вместо людей.

«Все бизнес-модели, строящиеся вокруг данных, до сих пор находятся в фазе мышления, отталкивающегося от «своих» данных, а не от целостной экосистемы данных. В конце концов, действительно, эффективное применение найдут системы, сочетающие данные из различных источников», уверен Нико Вешче, международный глава подразделения медиа и развлечений в ГФК.

Классификация должностей: переход к парадигме Big Data

Лебедев Павел Андреевич,
Шурыгина Дарья Алексеевна
(*Superjob.ru*)

Введение

В рамках своей деятельности исследовательский центр рекрутингового портала *superjob.ru* занимается анализом тенденций на рынке труда (спрос и предложение рабочей силы, средние уровни зарплатных предложений и ожиданий) на основе динамического, постоянно пополняемого массива резюме (более 15 млн. на конец июля 2015 г.) и вакансий (около 270–300 тысяч активных вакансий ежемесячно). Основой для классификации вакансий и резюме является двухуровневый каталог сфер деятельности, сформированный на основе Общероссийского классификатора видов экономической деятельности (ОКВЭД). Пользователи (соискатели и работодатели), а затем модераторы портала могут «привязать» каждую вакансию/резюме максимум к пяти сферам деятельности.

Анализ структуры рынка труда, основанный на каталоге сфер деятельности, не дает возможности фокусировать анализ на конкретных профессиях, должностях и т.п. Например, в подразделе «растениеводство» (раздел «сельское хозяйство») находятся следующие вакансии: инженер-эколог, сборщик урожая, овощевод, агроном и т.д. При этом вакансия сборщика урожая также может быть отнесена в раздел «рабочий персонал».

С учетом возможности различного написания названия одной и той же должности (разный порядок слов, синонимы и т.п.), в рамках традиционного подхода для анализа структуры спроса и предложения на уровне конкретных должностей нужно проводить «ручное» кодирование, приводя написание должностей к единой образной форме. Для упрощения этой трудоемкой задачи исполь-

зуется выборочный подход: или классифицируется ограниченное количество позиций (на основе имеющегося словаря синонимов должностей), или анализ проводится на небольшой выборке из всего объема данных. Получившийся в итоге результат обладает целым рядом ограничений.

В данной статье будет описан эксперимент по формированию алгоритма автоматической классификации должностей (позиций) с помощью кластерного анализа методом k-средних. Предлагаемый алгоритм не только может быть использован для решения задач по анализу структуры спроса и предложения на рынке труда, но также способен послужить основной при решении других задач, например, для автоматического расчета средних зарплатных предложений/ожиданий по наиболее распространенным позициям. Кроме этого, подобный алгоритм может быть использован для повышения точности и полноты поисковой выдачи на портале, для внедрения подсказки пользователям при заполнении вакансии (функционал, навыки, требования, условия работы и т.п.).

Безусловно, решение разных задач требует разных характеристик проектируемого решения. Например, для аналитических задач приоритетом будет высокая (близка к 100%) точность отнесения позиции к определенной группе. Высокая степень точности необходима для правильного определения редких позиций (доля вакансий для каждой из которых составляет менее 1% базы): на них приходится более 90% от всех позиций и около 45% от всех вакансий в базе. В то же время задачи, связанные с развитием основного сервиса, требуют скорее полноты кластеров (задачи по поиску) и возможности быстрой работы алгоритма в режиме реального времени (задачи по подсказкам).

В данной статье мы сконцентрируемся на решении именно аналитических задач на примере классификации вакансий.

Классификация профессий: теоретическая рамка

Задачу по классификации должностей, выявления профессиональных групп необходимо рассматривать в связке с содержательным аспектом, связанным не только с названием позиции, но с сутью выполняемой работы. Здесь мы с одной стороны тесно

соприкасаемся с социологией профессий и занятости. Начиная с классиков (М. Вебера, Э. Дюркгейма, К. Маркса) исследователи обращали пристальное внимание на процессы разделения труда, формирования и воспроизводства профессиональных групп, нормы и правила, определяющие единство (солидарность) представителей одной профессии. Обзор текущего положения дел в российской социологии профессий см. [7].

Наиболее распространенный в настоящее время неовеберинский подход к изучению профессий рассматривает профессионализм в качестве особого культурно-исторического конструкта, связанного с легитимацией профессионалами своего права на монополию в рамках определенной сферы деятельности. Профессия — это не только определенные стандарты компетенций (навыки, знания), позволяющие выполнять определенную работу, но еще и социально-политический подтекст — формирование этих стандартов (гласных или нет) для поддержания монополии на выполняемую функцию [3; 4].

Прикладные исследования в течение длительного времени были сфокусированы на таких устойчивых и «идеально-типических» профессиональных сообществах, как врачи, учителя, юристы, и зачастую носили этнографический и антропологический характер [13; 15].

Традиционно исследователи выделяют пять сторон, определяющих/формирующих представления о том, что такое профессия:

- работники, профессионалы, которые, обладая определенными навыками и знаниями, стремятся к поддержанию монополии на свою деятельность;
- пользователи услуг, чьи потребности и ожидания (спрос на услуги) влияют на требования к профессионалам;
- бизнес (работодатели), задача которого — объединить профессионалов разных сфер и с минимальными издержками сформировать потоковое производство продукта/услуги;
- образовательные учреждения, отвечающие за процесс легитимации знаний и навыков и воспроизводства профессионалов;
- государственные структуры, которые либо служат гарантом автономии и возможности для саморегуляции профессиональных ассоциаций (англо-американская модель), либо

активно участвуют в процесс лицензирования и регуляции (европейская модель)¹.

Безусловно, каждая из сторон имеет свои собственные представления о содержании конкретной профессии и стремится к получению определенных выгод. Рассматриваемый в данном случае массив данных — вакансии работодателей — отражает представления только одной из сторон. При этом каждая конкретная вакансия соотносится не с профессией, а с позицией (более узким понятием). Однако, мы исходим из того, что в ситуации поиска работы/сотрудника главные действующие лица нацелены, прежде всего, на достижение позитивного результата. Т.е. вакансия будет составлена таким образом, чтобы привлечь максимально подходящих кандидатов с релевантным опытом. Кандидаты в свою очередь будут создавать свои резюме наилучшим образом. В результате язык описания позиции, функционала и т.п. и в вакансиях и резюме будет очень близок.

Особенности и структура анализируемых данных

Исходную базу для анализа составили 275 433 актуальные вакансии, размещенные работодателями в течение месяца до выгрузки, за вычетом дублей — 66 379 уникальных названий позиций. Массив анализируемых данных (список должностей) обладает определенными особенностями.

Во-первых, одного названия должности недостаточно для однозначного определения верной позиции. Например, позиция «менеджер по рекламе» может подразумевать специалистов с разным функционалом: организация рекламных кампаний или продажа рекламных площадей. В зависимости от функционала позиции будут отличаться требования к кандидатам, схема оплаты труда (в первом случае это чаще всего фиксированный оклад, во втором —

¹ Наша задача по классификации должностей тесно связана с различными реестрами, классификаторами деятельности, утвержденными на государственном уровне (ОКВЭД; ОКПДТР (Общероссийский классификатор профессий рабочих, должностей служащих и тарифных разрядов), Реестр профессиональных стандартов и т.п.). На международном уровне одним из основных классификаторов является The International Standard Classification of Occupations (ISCO).

оклад и процент с продаж), среднее зарплатное предложение. Подобные вакансии невозможно правильно отнести к той или иной позиции только на основании названия должности: необходим анализ должностного функционала. Однако в данном эксперименте мы предполагаем, что общее число подобных вакансий невелико, и не учитываем это ограничение.

Во-вторых, в зависимости от контекста одно слово может обладать разным смыслом. Например, «сотрудник отдела продаж» и «менеджер отдела продаж» — два названия одной и той же позиции, т.е. сотрудник и менеджер в данном случае синонимы. В то же время сотрудник ресторана и менеджер ресторана — позиции разные: в первом случае имеется в виду рядовая позиция (сотрудник на раздачу блюд и т.п.), во втором — управленец среднего звена. Т.е. здесь синонимичность слов «сотрудник» и «менеджер» отсутствует.

В-третьих, в зависимости от задачи требуется разная детализация проводимой сегментации. При анализе конкретной сферы занятости, например медицинской, необходимо различать врачей разных специализаций (врач-оториноларинголог, врач-терапевт, врач-офтальмолог и т.д.). Аналогичный анализ по основным позициям в разных сферах не требует такой степени детализации: врачей можно рассматривать как единую группу. Таким образом, необходимо проведение кластеризации каждой отдельной отрасли с последующим формированием макро-кластеров, созданием иерархии должностей, где отдельные позиции (определенные максимально точно, вплоть до специализации) объединены в более крупные кластеры (например, врачи различных специализаций объединены в кластер «врачи»).

Обзоры существующих алгоритмов кластеризации, приведенные в [8; 4; 9; 2; 11] позволяют выявить 3 группы решений, потенциально подходящих для поставленной задачи:

- статистические алгоритмы (иерархический кластерный анализ, метод k -средних и т.д.);
- тематическое моделирование (латентное размещение Дирихле — LDA²);

² Latent Dirichlet Allocation

- машинное обучение (нейронные сети, генетические алгоритмы).

В данной статье мы описываем эксперимент по кластеризации позиций с помощью метода k-средних. Алгоритм был выбран как наиболее простой в реализации [11]. Кроме того, несколько предварительных экспериментов с использованием размещения Дерихле показали невозможность однозначного определения тематической принадлежности большей части вакансий только на основе ее названия. Каждая вакансия была отнесена в среднем к 6 разным темам. Неоднозначность отнесения вакансий может быть связана с большим количеством слов, имеющих равные шансы появиться разных позициях (темах): инженер, менеджер, руководитель и т.д. Такие результаты не подходят для решения нашей задачи, хотя следует признать потенциальную эффективность метода для анализа более крупных текстов (например, блока требований в вакансиях).

Для проведения эксперимента из общего массива вакансий была сформирована случайная выборка размером в 14 000 единиц (примерно 5% от всего объема активных вакансий). Для обеспечения возможности расчета точности кластеризации выборка вакансий была предварительно классифицирована вручную. Данная выборка будет выполнять роль контрольной группы, в ней выделено 817 кластеров.

Таблица № 1. Топ-5 кластеров по размеру, присвоенных вручную

Кластер (присвоено вручную)	Количество позиций в кластере
разнорабочий (грузчик, комплектовщик, упаковщик, маркировщик, укладчик) ¹	1504
продавец-консультант	896
мерчандайзер, работник торгового зала	740
кассир	675
менеджер по продажам	497

Самый большой по размеру кластер — «грузчик, комплектовщик, упаковщик, маркировщик, укладчик» содержит 1504 вакансии. В данном случае разные позиции часто совмещаются в одной вакансии (должности): грузчик-комплектовщик, комплектовщик-

упаковщик и т.п. Фактически все это выражается одним понятием «разнорабочий на склад». Число кластеров, состоящих из 1 вакансии, составляет 312 штук, или 2,2% от выборки.

Подготовка данных для проведения кластеризации

Все названия должностей для дальнейшего анализа были подвергнуты предварительной обработке:

- удалены знаки препинания,
- буква «ё» заменена на «е» для обеспечения единства написания,
- все буквы приведены к нижнему регистру,
- в русских словах удалены окончания,
- удалены короткие слова (к коротким словам относятся слова, состоящие из 1-й или 2-х букв, за исключением отдельного списка значимых двухбуквенных сочетаний: pr, 1с, с#, cd, dj, gr, hr, it, js, ms, ui, ux, зп, ит, пк.

Таблица № 2. Пример предварительной обработки названий должностей

№	Должность до обработки	Должность после обработки
1	Менеджер по продажам интернета и цифрового телевидения (удалённая работа)	менеджер продаж интернет цифров теле-виден удаленн работ
2	PR-менеджер	pr менеджер
3	Продавец-консультант в магазин CHARUEL (МЕГА Белая Дача г. Котельники)	продавец консультант магазин charuel мер бел дач котельник
4	Инженер-проектировщик / Заведующий группой раздела Водоснабжение и канализация	инженер проектировщик заведующ групп раздел водоснабжен канализац

Для проведения кластерного анализа была создана частотная таблица размером 14000*3408, где каждая строка — название должности из вакансии, а каждый столбец — слово из списка слов, встречающихся в названиях должностей. В ячейках таблицы — частота, с которой данное слово встречается в названии данной долж-

ности. Таким образом, в результате каждая должность оказывается представленной в виде числового вектора, где каждое число — частота встречаемости определенного слова в этой должности. Именно эти вектора и подвергаются кластеризации, т.к. алгоритм *k*-средних работает только с числовыми данными, и не может работать напрямую с текстом. Подробнее об исходных данных для кластеризации [9], о векторной модели представления текстов [1].

Определение количества кластеров

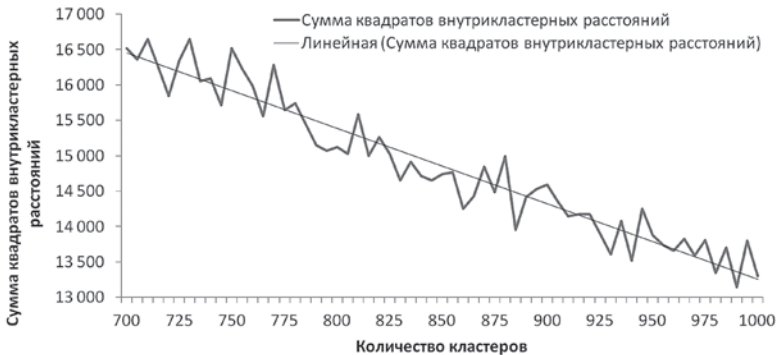
Наличие контрольного массива данных, обработанного вручную, позволяет нам задать искомое количество кластеров, однако первоначальная задача предполагает поиск алгоритма автоматического определения количества кластеров для массивов разного размера.

Первоначальная гипотеза состоит в том, что при увеличении количества групп в кластерном решении сумма квадратов внутренних расстояний в кластерах уменьшается. Таким образом, проведя серию кластеризаций с разным количеством решений можно найти «точку перегиба», после которой снижение суммы квадратов внутренних расстояний становится несущественным³. Эта точка должна соответствовать оптимальному количеству кластеров для данного массива.

Зная, что количество определенных вручную кластеров в выборке равно 817, мы выбрали для эксперимента 61 вариант кластерного анализа *k*-средних с количеством кластеров от 700 до 1000 и шагом 5 (т.е. 700, 705, 710 ... 990, 995, 1000 кластеров). На рисунке № 1 представлено распределение суммы квадратов внутрикластерных расстояний в зависимости от их количества.

³ Подобные способы определения оптимального количества кластеров являются стандартными для метода *k*-средних (См. [Воронцов, 2007]).

Рисунок 1. Распределение суммы квадратов внутрикластерных расстояний в зависимости от количества кластеров



Из графика видно, что очевидная точка перелома на ожидаемом количестве кластеров (817) отсутствует. Таким образом, подобный подход вряд ли можно считать приемлемым для определения числа кластеров.

*Вторая гипотеза*⁴ состоит в том, что чем больше уникальных (неповторяющихся) слов используется в названиях должностей, тем на большее количество кластеров объективно делится выборка. Таким образом, количество кластеров является линейной функцией от количества уникальных слов в названиях позиций.

Чтобы проверить эту гипотезу, из рабочего массива в 14 тысяч позиций было сгенерировано 1500 случайных выборок разного размера (количество элементов от 100 до 14 000). Затем, используя контрольный массив, был посчитан коэффициент корреляции между количеством уникальных слов в выборке и количеством кластеров определенных вручную. Коэффициент корреляции Пирсона между двумя переменными составляет 0,997 (уровень значимости 0,01).

Линейная регрессия также показывает высокий $R^2=0,99$ и значимые коэффициенты регрессии. Однако график распределения остатков относительно независимой переменной показывает, что их дисперсия неравномерна, т.е. нарушено одно из ограничений

⁴ Статистические гипотезы проверялись и оценивались в соответствии с критериями, описанными в [Крыштановский, 2006].

регрессионной модели (требование гомоскедастичности). Такая модель не может считаться надежной, поэтому появилась необходимость скорректировать гипотезу.

Вводить в уравнение регрессии размер выборки в качестве второй независимой переменной также невозможно из-за корреляции (коэффициент Пирсона между этой переменной и количеством уникальных слов составляет 0,964 при уровне значимости 0,01). Учитывая, что в выборке могут встречаться повторяющиеся названия должностей, правильнее говорить не о размере выборке, а о количестве уникальных должностей. Поэтому в качестве независимой переменной было использовано отношение числа уникальных слов к количеству уникальных должностей.

На рисунке № 2 видно, что существует нелинейная зависимость между отношением числа слов к размеру выборки и количеством кластеров. Подгонка кривых показывает, что экспоненциальная модель покрывает большее количество данных, чем степенная. В то же время при небольших значениях независимой переменной ни одна из моделей не дает хорошего результата.

Оценка всех используемых моделей для расчета оптимального количества кластеров представлена в таблице № 3.

Рисунок 2. Распределение отношения количества уникальных слов в названиях должностей к числу уникальных должностей в зависимости от количества кластеров

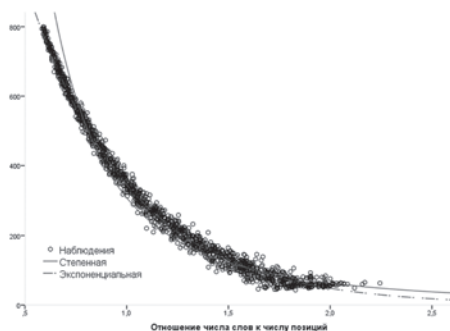


Таблица № 3. Сводные данные по моделям расчета количества кластеров

	Линейная модель	Степенная модель	Экспоненциальная модель
Параметры модели			
Зависимая переменная	Количество кластеров		
Независимая переменная	Количество уникальных слов в названиях должностей	Отношение количества уникальных слов к количеству уникальных должностей	
Оценка качества модели			
R ² скорректированный	0,994	0,958	0,975
Уровень значимости модели	0,000	0,000	0,000
Проверка ограничений модели			
Нормальное распределение остатков	Близко к нормальному	—	Близко к нормальному
Гомоскедастичность (примерно одинаковая дисперсия остатков при разных значениях независимой переменной)	—	—	+
Отсутствие автокорреляции остатков	+	+	+

Показатель R^2 у экспоненциальной модели несколько лучше, что позволяет использовать ее в качестве основной. Кроме того, экспоненциальная модель успешно проводит проверку на все ограничения.

Получена формула для расчета количества кластеров:

$$y = 2490,505 * e^{-1,973x}, \text{ где}$$

y — количество кластеров,

x — отношение количества уникальных слов в названиях позиций к размеру выборки.

Проведение кластерного анализа и оценка результатов

На вход в алгоритм кластеризации подается матрица размером $m \times n$, где m — количество уникальных слов в названиях должностей, n — количество вакансий в массиве. В ячейках матрицы находятся 0 и 1 в зависимости от наличия того или иного слова в определенной должности. Выдержка из матрицы приведена в таблице.

Таблица № 4. Пример матрицы для кластерного анализа

Должность	бухгалтер	главный	механик	младший	помощник
главный механик	0	1	1	0	0
Механик	0	0	1	0	0
главный бухгалтер	1	1	0	0	0
помощник главного бухгалтера	1	1	0	0	1
помощник механика	0	0	1	0	1
младший бухгалтер	1	0	0	1	0

Традиционно при проведении кластеризации текстов для повышения точности словам, имеющим **большую** значимость, приписывают **большие** веса. Применение весов происходит за счет умножения соответствующих 1 из вышеописанной матрицы на нужный вес.

Основная задача — найти формальные критерии для выбора семантически более значимых слов. Под семантической значимостью мы подразумеваем возможность с большой вероятностью определить позицию, опираясь на наличие этого слова в должности.

Были сделаны следующие предположения:

- К семантически значимым можно отнести слова, которые отдельно от других слов могут представлять собой название должности (т.е. существуют должности, состоящие только

из одного такого слова): повар, бухгалтер, сметчик и т.д. Всего в анализируемой выборке было выделено 230 таких слов. Далее в тексте такие слова называются одиночными.

- К семантически неоднозначным (т.е. встречающимся в названиях разных позиций) относятся слова, указывающие на уровень позиции (например, директор, руководитель, ведущий, старший, менеджер и проч.). Всего в выборке было выделено 17 таких слов. Им было решено дать меньшие веса.
- Учет синонимов может увеличить точность кластеризации. К синонимам относятся слова, имеющие одинаковое значение в рамках решения задачи по кластеризации: синонимы в лингвистическом смысле (юридический — правовой); русское и английское слово с одинаковым значением (разработчик — developer); написание одного слова кириллицей и латиницей (ит — it); близкие по смыслу должности, которые имеет смысл объединять в одну позицию: упаковщик — сортировщик). Всего была составлена 61 пара вида «каноническое слово — синоним». При определении синонимических пар мы опирались на использующиеся в рамках других проектов Superjob списки синонимов. Такие списки синонимов составлялись вручную на основе экспертных заключений для различных задач, в частности, для оптимизации поиска и подсказок пользователям при заполнении вакансий/резюме.

Для проверки перечисленных предположений было проведено 6 кластеризаций с различными вариантами весов (например, для одиночных слов брались веса 2, 4, 10). Центры кластеров определялись случайным образом. Для обеспечения возможности сравнения результатов различных вариантов кластеризаций методом k-средних, мы использовали функцию `set.seed`, определяющую, с какого числа начинается генерация случайных чисел, и обеспечивающую повторение того же набора случайных чисел в той же последовательности⁵.

Показатель точности рассчитывался по стандартной формуле (например, [2;14]):

⁵ Подробнее о работе функции — <http://www.inside-r.org/r-doc/base/set.seed>

$$P=x/n,$$

где P — точность,

x — количество правильно определенных позиций,

n — объем кластеризуемого массива

Правильность определения позиции проверялась по следующей схеме:

- для каждого автоматически определенного кластера рассчитана мода вручную присвоенных позиций,
- такая мода считается названием автоматически определенного кластера,
- если несколько автоматических кластеров имеют одинаковое название, то при расчете первого показателя точности (T_1) правильность определения считается только по вакансиям, отнесенных к самому большому кластеру с таким названием; остальные кластеры с аналогичными названиями считаются определенными неправильно; при расчете второго показателя точности (T_2) кластеры с одинаковыми, вручную присвоенными модальными названиями сливаются в один большой кластер и точность считается суммарно по всем позициям.

Более наглядно схему определения точности кластеризации можно разобрать на условном примере — см. таблицу № 4.

Таблица № 5. Пример определения правильности кластеризации

№ кластера (присвоено алгорит- мом кластеризации)	Название позиции в кластере	Правильно определено (Точность 1)	Правильно определено (Точность 2)
<i>Столбец 1</i>	<i>Столбец 2</i>	<i>Столбец 3</i>	<i>Столбец 4</i>
1	Продавец	1	1
1	Продавец	1	1
1	Продавец	1	1
1	Бухгалтер	0	0
2	Продавец	0	1
2	Продавец	0	1
2	Водитель	0	0
Итого Точность		42%	71%

Название кластеров № 1 и № 2 — «Продавец» (модальное значение — столбец № 2). В случае расчета Т 1 анализируется только кластер *большого* размера. В случае Т 2 — кластеры с одинаковыми названиями сливаются в один и точность считается уже по обобщенному кластеру (Т 2).

Рассчитанные по данному алгоритму показатели точности Т 1 и Т 2 сравнивались для кластерных решений с различными весами отдельных слов. В качестве основного показателя использовался показатель Т 1.

Таблица № 6. Результаты кластеризации методом k-средних при разных весах

Веса / Слова	Показатель	Без синонимов	С синонимами
Без весов	T 1	42%	
	T 2	77%	
Вес одиночных слов-должностей = 2	T 1	45%	52%
	T 2	79%	81%
Вес одиночных слов-должностей = 4	T 1		52%
	T 2		81%
Вес одиночных слов-должностей = 10	T 1		51%
	T 2		80%
Вес одиночных слов-должностей = 2; вес слов-показателей уровня *0,5	T 1		51%
	T 2		80%

Из таблицы видно, что хорошее увеличение точности дало введение синонимов: T 1 увеличилась с 45% до 52%. Введение весов для одиночных слов повышает уровень точности с 42% до 45%. Дальнейшее увеличение веса одиночных слов к улучшению результата не привело, также как и пониженный вес для слов, определяющих уровень позиции. Итоговая модель строилась с синонимами и повышенным весом (равным 2) для одиночных слов.

Автоматическое определение названий кластеров

Точность определения позиции к конкретному кластеру на уровне 80%, получаемая на контрольном массиве, с вручную присвоенными модальными названиями кластеров в целом можно считать приемлемым результатом даже для аналитических задач. Однако в дальнейшей работе желательно добиться более высокой точности при полной автоматизации процесса кластеризации, вплоть до выбора названия.

Для определения автоматических названий был разработан следующий алгоритм:

- для каждого кластера, определенного в результате машинной кластеризации, рассчитывается среднее количество слов в названии должности. Полученные цифры округляются;
- для каждого кластера рассчитываются наиболее часто встречающиеся в названиях должности слова;
- названием кластера считаются наиболее частотные слова из названий должностей в количестве, определенном на первом этапе.

Среди автоматически присвоенных названий кластеров встречались, например, «консультант. продавец. мтс.», «продаж. менеджер.», «разнорабоч. завод.», «монтажник. сист.», «фрезеровщик.», «консультант. продавец. молл.», «сборщик. техник. вахт. бытов. зеленодольск.», «официант.», «продавец. магазин. магнит.», «сборщик. слесар.», «оператор. чпу.» «заместител. начальник. перевозк. речн. служб.», «электрик.», «грузчик.», «продавец.» и т.п.

Проведя серию из более чем 1300 экспериментов по определению названия кластеров на основе автоматически присвоенных названий на случайных выборках разного объема (от 100 до 8850 позиций в каждой) мы получили среднюю $T_1 = 56\%$; среднюю $T_2 = 40\%$. Другими словами, процедура автоматического сливания кластеров с двумя одинаковыми названиями приводила к ухудшению результата. Это связано с тем, что определение название по модальному значению предполагает большое количество шума в данных. Позитивный эффект от слияния кластеров, получаемый на контрольном массиве с модальными значениями, проставленными вручную, оказывается недостижим.

Повышение точности при кластеризации последующих выборок

Проведя серию экспериментов по классификации должностей (вакансий) методом k-средних, мы можем сделать вывод о том, что требуемый уровень качества кластеризации на текущий момент полностью автоматизированным подходом получить не удалось.

Средний уровень точности колеблется на уровне 52–56%. Хотя полученный результат нельзя назвать идеальным, его все же можно использовать в дальнейшей прикладной работе. В частности, полученный уровень точности кластеризации позволяет существенно снизить объем ручной работы, неизбежной для решения поставленной задачи. Из исходного массива в 14 тысяч вакансий было сформировано около 800 кластеров. При этом ручная проверка 800 кластеров — существенно менее затратная задача, чем классификация всего массива вручную. Особенно если учитывать, что для аналитических целей параметр полноты не настолько значим и неправильно определенные позиции можно просто исключить из массива.

Полностью автоматизировать задачу не удалось, что в целом соотносится с исходными теоретическими посылами о сложности однозначного и формализованного определения конкретных позиций. В любом случае необходимо человеческое решение о требуемом уровне детализации профессиональных границ.

В рамках эксперимента был отработан следующий алгоритм действий:

- **подготовка массива** — стемминг, исключение коротких слов, приведение слов к одному регистру;
- **формирование числовой (векторной) матрицы** на основе обработанного массива;
- **автоматическое определение оптимального числа кластеров** как экспоненциальной функции от отношения числа уникальных слов к размеру анализируемого массива вакансий;
- **кластеризация** векторной матрицы методом k-средних исходя из оптимального количества кластеров;
- **присвоение автоматического названия кластера на основе** наиболее часто встречаемых слов в названиях должностей, входящих в кластер;
- **ручная проверка кластерного решения** — вычистка неправильно отнесенных позиций, переименование кластеров в «читаемый» вид.

Для оптимизации процесса ручной проверки результатов кластеризации более продуктивной представляется схема, при которой

проводится серия автоматических кластеризаций отдельно по каждой сфере деятельности. Количество кластеров в рамках одной сферы будет очевидно существенно меньше, чем в ситуации обработки единого массива вакансий из разных сфер занятости. Кроме того, работа с каждой сферой по отдельности обеспечивает возможность распараллеливания ручной проверки (над разными сферами работают разные специалисты) и оптимизации решения задачи во времени (возможность обрабатывать в первую очередь те сферы деятельности, которые наиболее нужны с точки зрения текущих бизнес-задач).

По итогам ручной проверки мы получаем унифицированную таблицу соответствия, в которой каждому названию должности приписано свое «правильное», «каноническое» ее написание. Благодаря ручной проверке, точность определения позиции близка к 100%. Помимо позиций, в таблице обозначаются сфера деятельности (по внутреннему каталогу) и уровень должности (специалист-руководитель-директор, определяется по наличию ключевых слов).

Таблица № 7. Выдержка из итоговой таблицы позиций

Название должности	Позиция	Сфера деятельности	Уровень должности
Менеджер по продажам салона связи МТС (г. Санкт-Петербург)	продавец-консультант	Продажи	Специалист
Продавец-консультант (м. Рязанский пр-т, пл. Карачарово)	продавец-консультант	Продажи	Специалист
Слесарь-сборщик	слесарь	Рабочий персонал	Специалист
Фельдшер-лаборант	фельдшер	Медицина, здравоохранение	Специалист
Руководитель корпоративного университета (переезд в г. Владимир)	руководитель учебного центра	Образование	Руководитель

Полученная таблица соответствия является основой алгоритма присвоения канонического названия должности для новых массивов вакансий. Данный алгоритм работает по следующей схеме:

- для каждой вакансии из нового массива алгоритм ищет такое же название должности в таблице соответствия;
- если такое название находится, то данной вакансии автоматически присваиваются соответствующее каноническое название, сфера и уровень;
- если такое название не находится, алгоритм ищет ближайшие названия должностей в таблице (содержащие такие же слова);
- алгоритм выдает не поддающиеся однозначной категоризации должности пользователю, предлагая выбрать из канонических названий похожих должностей или присвоить новое название;
- разнесенные вручную варианты попадают в таблицу соответствия и используются в дальнейшем для присвоения канонических названий новым позициям.

В рамках эксперимента по применению таблицы соответствий было автоматически кластеризовано 11 000 вакансий из сферы «ИТ, информационные технологии». Они были автоматически разбиты на 139 кластеров, каждый из которых был проверен вручную. В результате была создана таблица соответствий из 4 588 должностей. Следует отметить, что снижение числа должностей с 11 000 до 4 588 обусловлено двумя факторами. Во-первых, во многих вакансиях указаны одинаковые должности. Во-вторых, некоторые должности, по-разному представленные в вакансиях, оказались одинаковыми после предварительной обработки для кластеризации (стемминг, приведение к одному регистру и т.п.).

Через месяц была выгружена база активных вакансий из сферы «ИТ, информационные технологии». Поскольку регламентом портала Superjob предусмотрено, что вакансия остается активной в течение месяца, предполагается, что в выборке оказались новые вакансии (за исключением тех, для которых работодатели специально продлили срок действия на следующий месяц). Из 11 000 новых вакансий 75% было найдено в таблице соответствий. Остальные 25% были кластеризованы с помощью описанного выше алгоритма.

Направления для дальнейших исследований

Предполагается, что при дальнейшем ежемесячном анализе базы вакансий доля должностей, однозначно определенных с помощью таблицы соответствий, будет постоянно увеличиваться, что позволит максимально снизить ручную часть работы.

Более того, при накоплении достаточного количества данных (при ежемесячной категоризации всех активных за месяц вакансий в течение года будет накоплена база из более чем 3 млн. вакансий) станет возможным построение нейронной сети, что позволит исключить ручную часть работы и сделать работу по категоризации вакансий более быстрой и полностью автоматической.

Рассмотренный в статье алгоритм является примером перехода от полностью ручной классификации вакансий к инструментам Big Data, каким является нейронная сеть.

При этом возможности применения полученного алгоритма не ограничиваются рамками описанного проекта. Алгоритм может быть адаптирован под любую задачу, где требуется классификация небольших текстовых значений, в т.ч. наименования товаров, услуг и т.д., причем важен высокий уровень точности.

ЛИТЕРАТУРА

1. Andrews N.O., Fox. E.A. Recent Developments in Document Clustering // Department of Computer Science, Virginia Tech, Blacksburg, 2007. // URL: <<http://eprints.cs.vt.edu/archive/00001000/01/docclust.pdf>>. Дата обращения 30.09.2015.
2. Daud A., Li J., Zhou L., Muhammad F. Knowledge discovery through directed probabilistic topic models: a survey // Frontiers of Computer Science in China. 2010. Vol. 4. Issue 2. P. 280-301.
3. Faulconbridge J.R. Muzio D., The rescaling of the professions: towards a transnational sociology of the professions. // International Sociology. 2012. No 27 (1). P.109-125.
4. Jain A., Murty M., Flynn P. Data Clustering: A Review. // ACM Computing Surveys. 1999. Vol. 31. No. 3. P. 264-322.
5. Saks M. Defining a Profession: The Role of Knowledge and Expertise // Professions and Professionalism, 2012. Vol. 2. No. 1. P. 1-10.
6. Steyvers M, Griffiths T. Probabilistic topic models. // Landauer T, Mcnamara D, Dennis S, Kintsch W (Eds), Latent Semantic Analysis: A Road to Meaning / Lawrence Erlbaum, 2007. // URL: <http://psiexp.ss.uci.edu/research/papers/Steyvers_Griffiths_LSA_Book_Formatted.pdf>. Дата обращения 30.09.2015.
7. Абрамов Р.Н. Социология профессий и занятий в России: обзор текущей ситуации // Социологические исследования. 2013. № 1. С. 99-108.

8. Воронцов К.В. Лекции по алгоритмам кластеризации и многомерного шкалирования. МГУ 2007. // URL: <<http://www.ccas.ru/voron/download/Clustering.pdf>>. Дата обращения 30.09.2015.
9. Котов А., Красильников Н. Кластеризация данных. 2006 // URL: <<http://logic.pdmi.ras.ru/~yura/internet/02ia-seminar-note.pdf>>. Дата обращения 30.09.2015.
10. Крыштановский А.О. Анализ социологических данных. / Издательский дом ГУ-ВШЭ. Москва. 2006.
11. Кушнарев, Д.А. классификация алгоритмов кластеризации текстовых документов // Карповские научные чтения: сб. науч. ст. Вып. 5: в 2 ч. Ч. 1 / редкол.: А.И. Головня (отв. ред.) [и др.] — Минск: «Белорусский Дом печати», 2011. — С.179-183. // URL: <<http://elib.bsu.by/bitstream/123456789/52559/1/Кушнарев%20Д.А.%20Классификация%20алгоритмов%20кластеризации%20текстовых%20документов.pdf>>. Дата обращения 30.09.2015.
12. Попова И.П. Профессии и профессионализм в международной дискуссии // Социологические исследования. 2009. №8. С 52-57.
13. Романов П., Ярская-Смирнова Е. Мир профессий как поле антропологических исследований // Этнографическое обозрение 5. 2008. С. 3–17.
14. Уланов А. Обработка текстов на естественном языке. Лекция 5, «Кластеризация текстов» // URL: <https://compscicenter.ru/media/slides/nlp_2014_spring/2014_03_24_nlp_2014_spring.pdf>. Дата обращения 30.09.2015.
15. Щепанская Т.Б. Сравнительная этнография профессий: повседневные практики и культурные коды (Россия, конец XX — начало XXI в.). / СПб.: «Наука». 2010. С. 338.

Автоматизированные методы анализа больших массивов интернет-текстов

Кольцова Олеся
(*Национальный исследовательский университет*
«Высшая школа экономики»)

Введение

Автоматизированные методы анализа текстов становятся все более актуальными в связи с нарастающим потоком интернет-данных, представляющих интерес как для социальных наук, так и для бизнеса. Сейчас наблюдается взрывное развитие таких методов, и в то же время для большинства языков, включая русский, они находятся в стадии становления. Это ограничивает возможности их применения конечными пользователями — академическими исследователями и практическими аналитиками, в том числе и для анализа интернет-текстов. В виду своей сложности и незавершенности успешно эти методы применяются в основном в крупных IT-компаниях — например, в поисковых сервисах — для их собственных нужд, и это знание остается закрытым.

Исходя из вышесказанного, существующие исследования в области автоматического анализа интернет-текстов можно классифицировать по нескольким основаниям. Во-первых, их можно разделить на интернет-центрированные и общество-центрированные. Первый тип — это исследования, в которых интернет-тексты являются самостоятельным объектом, а знание о них — конечной целью; например, это исследования тематики политических онлайн-дискуссий или тональности отзывов на товары. Второй тип — это исследования, в которых интернет-тексты служат средством для объяснения или прогнозирования явлений офлайн-реальности — например,

прогнозирование результатов выборов по текстам твитов или эпидемий гриппа по поисковым запросам. Оба типа пока крайне слабо представлены в России за исключением группы работ по анализу тональности отзывов на товары.

Это подводит нас ко второму основанию классификации, по которому исследования можно разделить на «содержательные» и методологические, последние из которых доминируют. Исследователи в гораздо большей степени заняты разработкой методов, чем их применением и интерпретацией конкретных результатов. Методы же по понятным причинам разрабатываются в основном в обществе специалистов в области программирования, математики и компьютерной лингвистики, а не социальных наук. Социальным исследователям зачастую сложно даже сформулировать запрос, не говоря уже о разработке и применении методов современного уровня, что и определяет некоторую оторванность методов от задач социальной науки. Этот разрыв намного меньше в коммерческой среде, где для его преодоления существуют более сильные стимулы и большие ресурсы.

Наконец, сами методы автоматизированного анализа текстов (которые не всегда специфичны именно для интернет-текстов) можно условно разделить на три больших группы:

1. классификация, кластеризация и тематическое моделирование;
2. sentiment-анализ и извлечение мнений;
3. извлечение объектов и фактов.

Это не полная и не идеальная типология, методы могут комбинироваться в зависимости от задач и, более того, сами могут рассматриваться как задачи, требующие комбинации инструментов. Так, для оптимизации информационного поиска могут сочетаться кластеризация и извлечение объектов. Классификация может использоваться для задач sentiment-анализа и т.д. В этой статье мы сначала даем общую характеристику названных трех групп методов, а затем рассказываем об имеющихся российских исследованиях, в которых эти методы применяются для анализа интернет-текстов. Под российскими исследованиями понимаются работы с участием русскоязычных авторов, либо авторов, живущих в России, либо сделанные на русском или русскоязычном материале.

Автоматический анализ текстов для социальных наук

Подготовка текстов к анализу

Подготовка или препроцессинг текстов — с одной стороны, техническая задача, а с другой — требующая методологически выверенных и трудоемких процедур. Основная цель — сообщить компьютеру, из каких элементов состоит обрабатываемый текст, для того чтобы его потом можно было автоматически анализировать. Для большинства задач требуется, во-первых, расчленение текста на слова (что скорее представляет проблему в иероглифических языках, чем в русском). Во-вторых, обучение компьютера узнавать словоформы одного и того же слова, для чего проводится лемматизация или стемминг, и это уже отдельная большая проблема в случае флективных языков, включая русский. Чаще всего отдельные слова становятся признаками, которые потом подвергаются анализу. Но так бывает не всегда; поэтому — и это в-третьих — может потребоваться расчленение текста на последовательности слов (n -граммы), а также обучение компьютера узнаванию и вычленению частей речи, членов предложения, синтаксических конструкций, целых предложений и других элементов текста. Ни один исследователь не может осуществить препроцессинг «с нуля» и вынужден опираться на существующие инструменты: словари, программные продукты и т.д., в том числе лемматизаторы (например, MyStem (MyStem 2015) для русского языка) и парсеры (например, томиита-парсер (томиита-парсер 2015) для русского языка). Поэтому качество автоматического анализа текстов во многом зависит от развитости лингвистических инструментов для данного языка.

Классификация, кластеризация и тематическое моделирование

Под классификацией обычно понимают разделение элементов на заранее известные классы; в случае текстов самым очевидным примером является разделение их по языкам. При автоматической классификации обычно применяются методы машинного обуче-

ния «с учителем»: машине подаются тексты, разделенные на классы людьми, и она «обучается» на них как на примерах. Анализируя частоту различных слов в поданных примерах с помощью различных алгоритмов, машина относит новые тексты к одному из классов: так, текст со словами “*bonjour*” и “*merci*” будет в нашем примере отнесен к классу «французский», поскольку в поданных примерах эти слова встречались в текстах именно этого класса. К наиболее распространенным алгоритмам, используемым для классификации, относятся: метрические (например, метод ближайших соседей), вероятностные (например, наивный байесовский классификатор), линейные (например, машина опорных векторов (SVM), а также алгоритмы на основе регрессий и нейронных сетей) [1, р. 163-222; 28].

Однако зачастую состав классов не известен заранее: это особенно характерно для таких задач, как определение тематического состава корпуса текстов. В таких случаях применяется кластерный анализ, не требующий «учителя» (такие алгоритмы часто называют обучением без учителя). Кластерный анализ в целом хорошо знаком социальным исследователям; кластеризация интернет-текстов отличается большими объемами — не только большим количеством объектов, но и большим количеством признаков: число уникальных слов в больших коллекциях обычно измеряется несколькими сотнями тысяч. Это требует особых, быстро работающих алгоритмов [1, р. 77-128] и подходящего программного обеспечения. Например, целое семейство алгоритмов для кластеризации данных больших размерностей (т.е. в частности текстов) было предложено Дж. Кариписом и воплощено в его программном обеспечении Cluto [5]; алгоритмы включают специальные модификации иерархических и неиерархических моделей, включая графовые. Отдельную и очень важную задачу представляет интерпретация и присвоение имен кластерам. Кроме того, такие объекты как тексты зачастую воспринимаются как принадлежащие сразу к нескольким тематическим кластерам, а однозначное отнесение их только к одному кластеру обедняет представление о корпусе текстов.

Для решения этой проблемы используется либо нечеткая кластеризация, либо — все чаще — тематическое моделирование. Последнее — это группа вероятностных алгоритмов, сходных с факторным анализом [4; 18]. Предполагается, что тексты являются порождени-

ем латентных переменных, называемых темами, чье распределение по словам и по текстам неизвестно, и его предстоит восстановить. В ходе работы такие алгоритмы, во-первых, приписывают каждому слову вероятность принадлежности к каждой теме, в результате чего получаются списки наиболее характерных для каждой темы слов, по которым можно судить о содержании тем, присущих корпусу текстов. Сумма вероятностей всех слов по теме служит индикатором значимости темы в корпусе. Во-вторых, вероятность принадлежности к каждой теме приписывается и каждому тексту; таким образом, выполняется задача нечеткой кластеризации текстов. Среди множества имплементаций тематического моделирования можно выделить TopicMiner как интерфейсный пакет, разработанный специально для русского языка [20].

Классификация, кластеризация и тематическое моделирование текстовых корпусов могут выполняться с использованием не только отдельных слов, но и других признаков, наиболее распространенными из которых являются биграммы (последовательности из двух слов). Это заметно улучшает качество результатов, но утяжеляет данные и усложняет работу. Поэтому исследователи пока редко идут дальше работы с отдельными словами, скорее, это дело будущего. Общей проблемой для всех перечисленных методов является вопрос выбора количества классов, кластеров или тем. Теоретически его можно автоматизировать, задав параметры оптимизации, но вопросом остаются как раз эти параметры, и, в частности, критерии качества работы данных алгоритмов. Наиболее известное свободное программное обеспечение, объединяющее все три подгруппы рассмотренных алгоритмов — Mallet [12].

Сентимент-анализ и извлечение мнений

Сентимент-анализ, или анализ тональности текста [15], проще всего помыслить на примере анализа отзывов на товары и услуги: его задача — автоматически разделить отзывы на негативные, то есть содержащие отрицательное отношение автора к объекту, и позитивные (положительное отношение). Существующие подходы к оценке тональности можно условно разделить на словарные [19] и не ис-

пользующие словарь. При словарном подходе алгоритм присваивает тексту оценку тональности на основании сравнения его словарного состава с тезаурусом, слова в котором вручную отнесены к позитивным или негативным. Преимущество такого подхода в простоте самого алгоритма, но качество его зависит от того, насколько удастся подобрать такие слова, которые будут приводить систему к правильному решению. Альтернативой является классификация на основе обучения с учителем, здесь, однако, потребуются вручную размечать тексты, которые будут служить образцами. В отношении отзывов это часто не проблема: отзывы оставляются на специальных сайтах, где текст сопровождается выставленной автором оценкой. Она и служит меткой тональности.

Совсем другая ситуация с записями в блогах и социальных сетях. Они не только не размечены авторами, но понятие «позитивности» и «негативности» в них не так однозначно. Во-первых, в отличие от отзывов, они могут содержать множественные объекты, и отношение к ним может быть разным. Во-вторых, структура мнений в таких текстах может быть вообще гораздо сложнее: так, во время катастроф в социальных сетях растет число записей, наполненных негативными словами («ужасно», «кошмар»), однако большинство этих текстов выражают не негативное отношение к событию или его участникам, а соболезнование пострадавшим. Отношение типа «соболезнование» не совсем укладывается в шкалу «негативное — позитивное» и требует иной типологии мнений. Для политических текстов также характерны пары отношений типа «согласие»–«несогласие» или «поддержка»–«оппонирование», которые могут быть выражены как с использованием эмоционально окрашенной лексики, так и без нее. Инструменты тональной оценки в целом очень чувствительны к предметной области — так, инструменты, настроенные на отзывы на автомобили, будут плохо работать даже на отзывах на фильмы. За пределами отзывов на товары и услуги качество сентимент-анализа вообще пока чрезвычайно низко. Пока можно наблюдать отдельные попытки классификации текстов в зависимости от мнений — например, Лин и Хауптманн [11] предлагают метод автоматического разделения текстов СМИ на «произраильские» и «проарабские». Наиболее известным свобод-

но доступным программным обеспечением для сентимент-анализа является [17], однако оно пока не работает с русским языком.

К выявлению мнений примыкают задачи, связанные с поиском различных социальных проблем через пользовательский интернет-контент: нахождение экстремистских текстов или речи ненависти на основании расовой, этнической или религиозной принадлежности; выявление суицидальных настроений по текстам социальных сетей, выявление текстов, пропагандирующих употребление наркотиков или содержащих обращения к детям с целью их сексуального использования. Такие виды контента не совсем подходят под категорию «мнение», но методы их выявления очень схожи с сентимент-анализом и предполагают работу со словарями или машинное обучение. Также с сентимент-анализом их роднит значимость для качества результатов различных признаков текста, не сводимых к отдельным словам — синтаксических конструкций, словосочетаний, n-грамм, особенностей использования знаков препинания, заглавных букв и эмотиконов. Все эти признаки гораздо менее важны для качества тематического моделирования. Для обнаружения сложных признаков в алгоритмы вводятся правила: например, искать глагол желания 1 л. ед.ч + инфинитив глагола смерти («хочу умереть», «мечтаю сдохнуть»), возможно, на расстоянии 1–3 слова («мечтаю поскорее сдохнуть»). Использование правил часто сочетается со словарными подходами и иногда с машинным обучением.

Извлечение объектов, аспектов и фактов

Множественность объектов, на которые может быть направлено отношение — одна из проблем, решением которой занимаются методом извлечения объектов. Простейшим объектом являются персоны, которые могут быть обозначены разными способами, например: «Путин В.В.», «Владимир Путин», «президент РФ» или просто «он». Задача извлечения объекта — определить, какие слова и словосочетания указывают на один и тот же объект. После такого определения можно найти способы автоматического соотнесения выявляемых отношений или тем именно с интересующим объектом. Другие примеры объектов — организации или географические названия.

Можно выделить три основных подхода к извлечению объектов. Первый основан на онтологиях — специальных словарях, содержащих связи между элементами (например, о том, что президент РФ — должность Путина). Онтологии достаточно трудоемки в составлении и, как и все словари, чувствительны к предметной области. Другой подход — так же, как и в сентимент-анализе, — машинное обучение, для которого, впрочем, требуются тексты не просто маркированные одной меткой, но детально размеченные. И третий подход предполагает ручное составление правил. Он, пожалуй, является основным для задачи извлечения фактов, то есть групп объектов и связей между ними. Примером такой задачи является автоматический поиск сообщений о терактах или о протестах. Такие тексты содержат схожие структуры, которые можно описать правилами: так, в описании теракта обычно участвуют глаголы или существительные, указывающие на взрыв; рядом с ними присутствуют объекты — устроители взрыва, не являющиеся военными, и объекты — места взрыва, как правило, публичные городские пространства вне зоны военных действий. Кроме того, помимо объектов, распространенной задачей является извлечение аспектов этих объектов — то есть параметров, по которым авторы текстов оценивают объекты в своих текстах. Это важно, потому что оценки одного и того же объекта по разным аспектам могут быть разными. Например, типичными аспектами отелей являются следующие: местоположение, номерной фонд, питание, персонал и некоторые другие; однако аспекты других товарных категорий менее очевидны и требуют извлечения, которое также обычно основано на методах машинного обучения. Извлечение объектов, аспектов и фактов [2, р. 11–42] — область, чрезвычайно развитая в коммерческом секторе, так как эти техники используются для оптимизации поисковых выдач и для маркетинговых исследований.

Российские исследования с применением автоматического анализа текстов

Перейдем к обзору российских исследований, использующих вышеописанные методы. Как уже говорилось, основные работы в области исследований интернета с использованием методов авто-

матического анализа текстов делаются в основном вне социальных наук. У этой группы работ нет единого «места» в научном сообществе. Часть работ докладывается на ежегодной конференции «Диалог» [31] и затем публикуется в ее сборнике «Компьютерная лингвистика и интеллектуальные технологии». Работы по интернет-данным получили здесь развитие благодаря многолетнему проекту РОМИП [33] — Российскому семинару по оценке методов информационного поиска. Это некоммерческое партнерство десять лет проводило соревнования алгоритмов классификации, кластеризации, извлечения объектов, сентимент-анализа и извлечения мнений в рамках конференции «Диалог» на своих «дорожках» — тестовых коллекциях текстов, часть из которых извлекались из интернета. Это были как статичные сайты, так и новости и отзывы пользователей на товары. Еще одна существенная часть работ в этой области докладывается на конференциях Российской летней школы по информационному поиску [16] — здесь больше работ, связанных с интернетом, но они не публикуются (что, однако, позволяет увидеть здесь результаты коммерческих исследований). Часть работ можно встретить в трудах конференции АИСТ («Анализ изображений, сетей и текстов») [22]. Они носят в основном методологический характер, не представляя конечных эмпирических результатов, а предлагая способы их получения. Отдельные эмпирические работы разбросаны по разным изданиям и сообществам. В данном обзоре мы пытаемся сфокусироваться на этом последнем типе работ.

Классификация, кластеризация и тематическое моделирование в российских исследованиях

В группе методов классификации и кластеризации уже сейчас можно выделить некоторое количество ярких работ, сделанных математиками совместно с социальными исследователями. Так, Якушев и соавторы [35] исследовали возможности моделирования наркотизации населения по данным социальных сетей. Для этого ими была использована база знаний, состоящая из словаря и правил, и на основе выраженности присутствия признаков из базы в текстах аккаунтов ВКонтакте «хозяева» аккаунтов были классифицированы

по степени вовлеченности в наркокультуру. Доли разных типов вовлеченности, определенные по материалам ВКонтакте, совпали с долями этих типов, смоделированными на офлайновых данных. Статья не лишена ограничений: в ней не описана база знаний и нет никаких сведений об офлайновой модели. Тем не менее эта работа является наиболее продвинутой из ряда работ, использующих словарный подход; нередко социальные исследователи ограничиваются мониторингом упоминаемости нескольких ключевых слов с использованием таких коммерческих систем мониторинга социальных медиа как IQBuzzили Wobot. Так, Биккулов [26] исследует динамику обсуждения самоубийств в социальных сетях, изучая распределения ключевых слов во времени, по типам источников и пользователей на данных сервиса IQBuzz. Бершадская и Чугунов [25] сходным образом, на основе того же сервиса анализируют динамику обсуждения в социальных сетях темы электронного правительства, а также проекта «Российская общественная инициатива».

Хотя метод классификации с машинным обучением широко используется для опознавания разных социально значимых видов текстов за рубежом, среди российских исследований удалось найти только исследования, направленные на решение коммерческих задач. Так, Гречников и соавторы [30] тестируют алгоритм классификации TreeNet для задачи опознавания неестественных текстов (спам, сгенерированного ботами). Коршунов и соавторы [32] применяют алгоритм классификации, известный как машина опорных векторов, для определения демографических характеристик пользователей Твиттера по текстам их записей. Метод применяется к шести европейским языкам, включая русский; он разбивает твиты пользователей на слова, биграммы и триграммы и приписывает пользователя к той демографической группе, для которой характерно наиболее сходное с данным пользователем сочетание этих признаков. Выборки составили по 500 пользователей (из них по 450 составили обучающую коллекцию) для каждого языка и для каждого исследуемого социально-демографического атрибута. Следует отметить, что это исследование может иметь применение не только в маркетинге, но и в социологии, где исследователям часто не достает демографических признаков изучаемых пользователей социальных сетей. Метод демонстрирует большую точ-

ность, однако он не описан подробно, а используемые выборки довольно малы.

Тематическое моделирование было в последнее время использовано в нескольких социологических работах. Кольцова и Кольцов [10] показали применимость тематического моделирования для исследования динамики тематики блогов на примере русскоязычного Живого Журнала, исследованного в «спокойный» и «предвыборный» периоды. Это первое в России социологическое исследование с применением тематического моделирования сравнивало две выборки постов популярных блоггеров, примерно по 25 тысяч постов за август и декабрь 2011 года, и показало вытеснение темой выборов других социально-значимых тем во второй период по сравнению с первым, в то время как рекреационно-потребительские темы остались неизменными.

Алексеева, Кольцова и Кольцов [23] сравнивали тематику топовых (популярных) и случайных блоггеров Живого Журнала и не обнаружили значимой разницы, которая заключалась лишь в количестве оставляемых постов и получаемых комментариев. Также Кольцова и соавторы [9] изучали принципы формирования сообществ совместного комментирования — групп пользователей, которые комментируют примерно одну и ту же группу постов. Анализ проводился на материале одной недели, включившей все посты топ-2000 блоггеров и все комментарии к ним. Авторы обнаружили, что лишь в некоторых сообществах посты близки лексически (а в большей степени они близки по авторству), а тематическое моделирование показало, что в сообществах, где комментируются лексически близкие посты, можно вычленивать, как правило, несколько доминирующих тем. Таким образом, нет убедительных свидетельств того, что люди связывают совместные дискуссии о постах на основании их темы. Воскресенский и соавторы [21] изучили обсуждения в группах многоквартирных домов Петербурга в сети ВКонтакте и показали, что тематика открытых и закрытых групп отличается и связана также с членством их участников в других группах. В частности, открытые группы больше ассоциируются с членством в градозащитных группах и с правозащитной тематикой в противовес «бытовой»; тематическое сравнение открытых и закрытых групп — совершенно неизведанная

область, способная многое сказать о влиянии настроек приватности на поведение пользователей. Горгадзе и Александров [7] провели мэппинг групп «кавказцев» ВКонтакте с помощью сетевого анализа, куда вошло 887 активных групп, и тематического моделирования, и 287 групп по тремужно-кавказским этническим группам и псевдоэтнониму «кавказцы». Авторы выявили, что среди моноэтнических и полиэтнических групп последние являются посредниками; большинство групп не политизированы; религия и ислам, в частности, не являются центральной темой.

Все эти работы, однако, объединяет недостаток «наивного» использования алгоритма латентного размещения Дирихле, при котором его параметры (гиперпараметры альфа и бета, количество тем, количество итераций) никак не подбираются и не оцениваются. Кроме того, алгоритм запускается один раз. Кольцов и соавторы [8] показали, что даже при разных запусках с одними и теми же параметрами алгоритм дает радикально разные результаты, затрудняющие их социологическую интерпретацию, а о разнице результатов с разными параметрами остается только догадываться, так как эта тема практически не изучена. Николенко и соавторы [14] предлагают меру качества работы алгоритма и подтверждают ее успешность с помощью экспериментов на людях. Однако еще только предстоит разработать меры качества, оценивающие не отдельные темы, а целые решения с разных точек зрения, и выработать рекомендации по подбору параметров для различных задач.

Сентимент-анализ и извлечение мнений в российских исследованиях

Сентимент-анализ широко представлен в российских работах, тестирующих новые методы анализа отзывов на книги, фильмы, компьютерные игры, мобильные телефоны, цифровые камеры — во многом благодаря конкурсу РОМИП. Большинство таких работ не представляют никаких содержательных выводов по анализируемым товарным категориям, а те, которые представляют, крайне малочисленны. Например, Бойко [27] предлагает не столько новый алгоритм анализа отзывов, сколько интегрирует уже имеющийся

в схему конкретного маркетингового исследования, проводя анализ отзывов на банки. В качестве объектов используются четыре российских банка; в ходе исследования автоматически вычлняются аспекты для этих банков (персонал, кредит, депозит, карта, кассовое обслуживание) и рассчитываются рейтинги удовлетворенности каждым из аспектов в каждом из банков. Так, самые низкие оценки получают кредиты Альфа-банка, а самые высокие — депозиты его же и ВТБ24. Ценность этой работы — в попытке преодолеть разрыв между лингвистическими наработками в области методов и встраиванием их в программы конкретных исследований. Андреева и Никитина [24] анализируют записи блогов об iPad, продукте Apple, однако их методика не совсем ясно описана. Вероятнее всего, для анализа русскоязычной платформы F5 используется ручное кодирование сообщений, а при анализе англоязычных твитов применяется встроенный в Твиттер тональный анализатор. При неясности описания исследования выводы авторов представляются не совсем надежными.

Герданович и Герданович [29] анализируют отношение к образованию в англоязычных твитах, имеющих геотэги, с целью определить регионы, в которых отношение наиболее позитивное и которые, как предполагают авторы, таким образом, могут быть наиболее вероятными потребителями белорусских образовательных услуг. Авторы классифицируют твиты на пять классов (от -2 до +2) с помощью готового метода, предложенного Брином (Breene), однако далее они затрудняются связать оценки с их причинами (объектами и аспектами). Построенные ими облака слов для положительных и отрицательных отзывов состоят в основном из общеупотребительных слов, что они ошибочно объясняют недостатками выборки (на самом деле в облаках представлены просто самые частотные для классов слова, а не слова с наибольшей дискриминационной силой). Тем не менее интересной является попытка построения карт распределения сентимента.

Этлинг [6] сравнивает отношение к украинским протестам 2013—2014 годов (вплоть до отставки Януковича) в англоязычных и русскоязычных интернет-источниках, выделяя среди последних российские и украинские. Данная работа демонстрирует возможности метода сентимент-анализа для исследования самых горячих политических

тем. Автор использует данные и инструменты анализа компании CrimsonHexagon. Он показывает, что в целом англоязычные источники более негативны, чем русскоязычные, вопреки его ожиданиям, однако если выделить среди всех источников только блоги и социальные сети и разделить их на украинские, американо-британские и российские, то негативность будет нарастать от первых к последним. Следует отметить, что автор сначала характеризует выделяемые им классы текстов как поддерживающие и осуждающие протесты, затем как негативные и позитивные, не объясняя, как эти типы классов соотносятся и какие именно в итоге он выделяет. Кроме того, размеры размеченных обучающих коллекций (30–45 постов для каждой тональной категории на каждом языке) вызывают некоторые сомнения, несмотря на то, что автор ссылается на рекомендации CrimsonHexagon.

Следует отметить, что работы с применением сентимент-анализа для русскоязычного материала (а не предлагающие методы) — единичны, а работы с применением вычленения сущностей отсутствуют.

Заключение

Методы автоматического анализа текстов стремительно развиваются и, несмотря на то, что они еще слабо освоены в социальных науках, в будущем можно ожидать прорывов в социальных исследованиях именно на основе этих методов. Наиболее интересным в этой связи представляется сочетание разных методов. Так, после получения представления о тематической структуре корпуса текстов не менее важным оказывается вычленить сущности (людей, организации, события), характеризующих темы, аспекты этих сущностей, отношения авторов текстов к этим аспектам и их распределение во времени и пространстве. Коммерческие сервисы, предлагающие упрощенные варианты описанной системы аналитики, уже существуют, а когда такие системы разовьются в полную силу, исследователи получат действительно мощные инструменты сбора и анализа текстов данных.

ЛИТЕРАТУРА

1. Aggarwal C.C., Zhai C. A Survey of Text Classification Algorithms. In: C. Aggarwal, C. Zhai (eds) Mining Text Data. Springer 2012. P. 163–222.

2. Aggarwal C.C., Zhai C. A Survey of Text Clustering Algorithms. In: C. Aggarwal, C. Zhai (eds) Mining Text Data. Springer 2012. P. 77-128.
3. Aggarwal C.C., Zhai C. Information Extraction from Text. In: C. Aggarwal, C. Zhai (eds) Mining Text Data. Springer 2012. P. 11-42.
4. Blei D. Probabilistic topic models. Communications of the ACM, 55(4). P. 77-84, 2012.
5. Cluto 2007. URL: <http://glaros.dtc.umn.edu/gkhome/cluto/cluto/download> (доступ осуществлен: 17.09.2015).
6. Etling, B. Russia, Ukraine, and the West: Social Media Sentiment in the Euromaidan Protests. The Berkman Center for Internet and Society at Harvard University. Research Publication. No. 2014-13. September 25. 2014. URL: <http://cyber.law.harvard.edu/publications/2014/euromaidan> (доступ осуществлен: 17.09.2015).
7. Gorgadze A., Alexandrov D. Virtual Transnational Movements in the Caucasus. Paper presented at Social Media and Social Movements conference. St. Petersburg. 2014. URL: <http://linisevents.hse.ru/data/2014/09/01/1313572271/Gorgadze%20Alexandrov%20SMSM%202014%20three%20pages.pdf> (доступ осуществлен: 17.09.2015).
8. Koltsov S., Koltsova O., Nikolenko S. I. Latent Dirichlet Allocation: Stability and Applications to Studies of User-Generated content // Proceedings of WebSci '14 ACM Web Science Conference, Bloomington, IN, USA. June 23 – 26. 2014. NY : ACM. 2014. P. 161-165.
9. Koltsova O., Koltcov S., Nikolenko S. Communities of co-commenting in the Russian LiveJournal and their topical coherence // Internet Research. Vol 26. Iss. 3. 2016 (forthcoming).
10. Koltsova O., Koltcov S. Mapping the Public Agenda with Topic Modeling: The Case of the Russian LiveJournal // Policy & Internet. 2013. Vol. 5. No. 2. P. 207-227.
11. Lin W.-H., Hauptmann A. Are these documents written from different perspectives? A test of different perspectives based on statistical distribution divergence. In Proceedings of the International Conference on Computational Linguistics (COLING) / Proceedings of the Association for Computational Linguistics (ACL). pages 1057–1064. Sydney. Australia. July 2006. Association for Computational Linguistics.
12. Mallet 2002. URL: <http://mallet.cs.umass.edu/> (доступ осуществлен: 17.09.2015).
13. Mystem 2015. URL: <https://tech.yandex.ru/mystem/> (доступ осуществлен: 17.09.2015).
14. Nikolenko S., Koltsov S., Koltsova O. Topic modelling for qualitative studies 2016 (forthcoming).
15. Pang, B., Lee, L. Opinion mining and sentiment analysis // Foundations and Trends in Information Retrieval. 2008. Vol. 2. No 1-2. P. 1–135.
16. RuSSIR 2015 URL: <http://romip.ru/russir2015/> (доступ осуществлен: 17.09.2015).
17. SentiStrength URL: 2012. <http://sentistrength.wlv.ac.uk/> (доступ осуществлен: 17.09.2015).
18. Steyvers M., Griffiths T. Probabilistic topic models. In: T. Landauer, D McNamara, S. Dennis, and W. Kintsch (eds), Latent Semantic Analysis: A Road to Meaning. Laurence Erlbaum. 2007.
19. Taboada, M., Brooke, J., Tofiloski, M., Stede, M. Lexicon-Based Methods for Sentiment Analysis // Computational Linguistics. 2011. Vol. 37. No. 2. P. 267-307.
20. Topic Miner 2013 URL: <http://linis.hse.ru/en/soft-linis> (доступ осуществлен: 17.09.2015).
21. Voskresenskiy V., Sukharev K., Musabirov I., Alexandrov D. Online Communication in Apartment Buildings, in: Lecture Notes in Computer Science Vol. 8852. SocInfo 2014.

International Workshops. Barcelona. Spain. November 11. 2014. Revised Selected Papers. Springer International Publishing. 2015.

22. АИСТ 2015. URL: <http://aistconf.org/> (доступ осуществлен: 17.09.2015).

23. Алексеева С., Кольцова О., Кольцов С. Общественное мнение онлайн: сравнение структуры и тематики постов «обычных» и «популярных» блогеров Живого Журнала // Supplementary Proceedings of the 3rd International Conference on Analysis of Images, Social Networks and Texts (AIST 2014) / Ed. By D. I. Ignatov, M. Y. Khachay, A. Panchenko, N. Konstantinova, R. Yavorsky, D. Ustalov. Vol. 1197: Supplementary Proceedings of AIST 2014. CEUR-WS.org. 2014. P. 177-181.

24. Андреева А.Н., Никитина М.С. Сентимент-анализ брендов в Российской блогосфере как инструмент маркетинговых исследований // Бренд-менеджмент, 04 (65) 2012.

25. Бершадская А.А., Чугунов А.В. Востребованность услуг электронного правительства: анализ дискуссий в социальных сетях // Интернет и современное общество: сборник научных статей. Труды XVI Всероссийской объединенной конференции «Интернет и современное общество» (IMS-2013), Санкт-Петербург, 9 – 11 октября 2013 г. СПб: НИУ ИТМО, 2013. С. 67 – 71.

26. Биккулов А.С. Подростковые самоубийства в обсуждениях блогосферы // Интернет и современное общество: сборник научных статей. Труды XVI Всероссийской объединенной конференции «Интернет и современное общество» (IMS-2013). Санкт-Петербург, 9 – 11 октября 2013 г. СПб: НИУ ИТМО. 2013. С. 72 – 76.

27. Бойко М.В. Исследование удовлетворенности потребителей в банковской сфере на основе анализа текстовых отзывов // Вестник УГАТУ. 2014. Т. 18. № 5 (66). С. 139–145.

28. Воронцов К.В. Лекции по метрическим алгоритмам классификации. 2008. URL: <http://www.ccas.ru/voron/download/MetricAlgs.pdf> (доступ осуществлен: 17.09.2015).

29. Гедранович, Б.А. Гедранович А.Б. Отношение к высшему образованию: сентимент-анализ данных микроблогов // Инновационные образовательные технологии. 2013. № 1 (33). С. 46—54.

30. Гречников Е.А., Гусев Г.Г., Кустарев А.А., Райгородский А.М. Поиск неестественных текстов. Труды 11й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». RCDL'2009. Петрозаводск. Россия. 2009.

31. Диалог 2015. URL: <http://www.dialog-21.ru/> (доступ осуществлен: 17.09.2015).

32. Коршунов А., Белобородов И., Гомзин А., Чуприна К., Астраханцев Н., Недумов Я., Турдаков Д. Определение демографических атрибутов пользователей микроблогов // Труды Института системного программирования РАН. Том 25. 2013. С. 179-194.

33. РОМИП 2014 URL: <http://romip.ru/> (доступ осуществлен: 17.09.2015).

34. Томита-парсер 2015. URL: <https://tech.yandex.ru/tomita/> (доступ осуществлен: 17.09.2015).

35. Якушев А.В., Митягин С.А., Бухановский А.В. Имитационное моделирование наркотизации населения по данным мониторинга социальных сетей // Современные исследования социальных проблем (электронный научный журнал). 2012. № 2 (10). С. 133-151.

WAVE 8¹: контент и его ценность.

Новый язык общения производителя и потребителя

*Бородин Даниил Александрович
(UM Russia)*

¹ Wave

В 2006 году агентство UM запустило исследовательский проект Wave, предназначенный для измерения масштаба и влияния социальных медиа в мире. В ходе исследования было показано, что это не просто история беспрецедентного роста, но и общественная эволюция. Таким образом, с каждой новой волной исследования мы получаем возможность расширить и углубить наше понимание социальных медиа и получить выводы, которые помогут компаниям-производителям успешно ориентироваться на рынке. Важно отметить, что Россия присутствует в исследовании с самого начала.

Вот краткая история исследований Wave:

Wave 1 (2006) показало, что социальные медиа оправдывают ожидания пользователей: в интернете коммуницирует большое и активное сообщество.

Wave 2 (2007) показало, насколько социальные медиа ушли от простого текстового формата блоггеров и тех, кто комментирует их посты, к формату, наполненному аудио и видео-контентом, который создают и распространяют пользователи.

Wave 3 (2008) показало демократизацию взаимного влияния социальных сетей и их пользователей и то, как социальные медиа предоставляют большие средства и возможности для взаимного влияния потребителей.

Wave 4 (2009) изучало причины, стоящие за бурным ростом потребления социальных медиа, пытаясь понять мотивации использования различных платформ. Исследование показало, что пользователи вовлекаются в ту или иную платформу потому, что она соответствует его специфическим потребностям, причем платформы в разной степени отвечают этим потребностям.

Wave 5 (2010) рассказало нам, что существует огромный спрос на взаимодействие в социальных сетях с компаниями-производителями. Причина и степень этого взаимодействия сильно варьируется в зависимости от потребителя и категории бренда. Компании-производители, которые преуспели в этом, выиграли очень много, получив лояльность, одобрение и рост продаж.

Wave 6 (2011) показало, что путь навстречу нуждам потребителя приносит разный эффект для бренда. Например, кому-то удастся добиться лояльности, тогда как другим — роста продаж. Таким образом, четкое понимание того, что нужно потребителю и как компания-производитель может достичь поставленных целей, есть ключ к долгосрочному успеху в социальных сетях.

Wave 7 (2013) продемонстрировало, что наблюдения, которые мы извлекаем из понимания поведения в социальных сетях, не просто важны в контексте пространства социальных медиа, но и являются отражением свойственных любому человеку черт, которые обосновывают отношения потребителя с продуктами и брендами.

Wave 8 (2014) — Методология. Исследование сохраняет одну и ту же методологию от первой волны до восьмой, предоставляя возможность сравнивать данные. Для Wave 8 — The Language of Content — агентством опрошено 50021 активных интернет-пользователей в возрасте от 16 до 54 лет в 65 странах — тех, кто посещают интернет хотя бы раз в два дня. Все анкеты респонденты заполняли самостоятельно, и собранные данные, в целом, носят количественный характер. Эта выборка репрезентирует генеральную совокупность 1,29 миллиарда активных интернет-пользователей по всему миру.

Активные интернет-пользователи — это те, кто использует интернет каждый день или через день, в том числе социальные медиа. Благодаря им адаптируются коммуникативные платформы и инструменты. Именно эти люди определяют, какие инструменты и платформы станут основными.

Исследование UM² Wave 8 — The Language of Content («Волна 8 — Язык контента») — позволяет продемонстрировать важность контента как нового языка общения с потребителем, дает наглядное описание нового нарциссического поколения «Селфи» и выявляет истинные мотивы потребителей как распространителей контента.

Подводя итоги опроса среди более чем 50 000 человек в 65 странах мира, включая Россию, данные, полученные в ходе проведенного исследования Wave 8, убедительно демонстрируют, что понимание силы контента как «социального продукта», которым люди хотят делиться, — это ключ к созданию более сильных социальных стратегий в цифровой среде. Из этого следуют следующие возможности развития брендов, основанные на трех основных наблюдениях:

- Лучший способ продвинуть контент — адаптировать его под нужды целевой аудитории. Исследование Wave 8 показывает, что мотивы распространения контента сильно различаются среди разных категорий продуктов, рынков и возрастных групп и уже становится недостаточным создавать смешной или развлекательный контент без дополнительной персонализации. Исследование Wave 8 помогает определить эти мо-

² UM (www.umwww.com) — глобальное медиа-агентство, для которого движущим принципом достижения креативных преимуществ бренда через медиа являются любознательность и креативность. Философия любознательности UM (Curiosity) вдохновляется знаниями, идеями и ориентацией на поведение потребителей.

UM — подразделение IPG Mediabrands, которое насчитывает более 130 офисов в 100 странах мира. Более 4800 сотрудников компании занимаются разработкой креативных, стратегических и индивидуальных предложений для клиентов компании.

В 2012 году AdNews (<http://www.adnews.com/>) назвал UM первым в списке самых креативных медийных агентств в мире, а MediaMagazine (<http://www.mediapost.com/>) отметил лидерство UM в продвинутой аналитике данных.

Список глобальных клиентов UM включает: Brown-Forman, Chrysler, Coca-Cola, ExxonMobil, Hershey's, Johnson & Johnson, L'Oreal Paris, MasterCard, Microsoft and Sony.

В 2012 году соотношение по развитию нового бизнеса и сохранению текущего принесло глобальному UM оценку А по классификации RECMA.

В России UM является частью АДВ (Группа маркетинговых коммуникаций АДВ (Группа АДВ)) — одна из крупнейших коммуникационных групп в России и СНГ. Адрес в интернете: <http://www.advglobal.com/>) и работает с Burger King, Дымов, Промсвязьбанк и др.

тивы и имплементировать эти данные в коммуникационную стратегию для дальнейшего продвижения бренда.

- Существует значительная разница между тем контентом, который потребители ценят, контентом, который получает большое количество просмотров на YouTube, и контентом, которым потребители делятся в социальных сетях. Только 11% считают «что-то, что отражает их точку зрения» ценным контентом, но вдвое больше людей говорят, что они поделились бы этим контентом с остальными.
- Пользователи интернета становятся все более эмоционально вовлеченными в процессе потребления контента, который они ищут, смотрят или читают. Когда мы делимся чем-то, мы таким образом заявляем что-то о себе, соответственно, контент становится связан с нашей репутацией в Сети. Эта сильная эмоциональная связь имеет опосредованное воздействие, в том числе и на бренды.

После многих лет исследований — с первого исследования Wave 1 по последнее на настоящий момент Wave 8 — был наглядно подтвержден вывод, что у интернет-пользователя есть только пять основных фундаментальных потребностей, которые они пытаются удовлетворить онлайн — обучение, построение отношений, развлечение, развитие и признание. Таким образом, в переполненном информацией интернет-пространстве, в котором хотя бы выделить не только бренды, но и потребители, последние для достижения своей цели используют контент как социальный продукт. И этот контент, которым делятся с нами потребители, может помочь определить их персональные, социальные и профессиональные потребности. И именно по этой причине для нас как для рекламного агентства становится важнее побудить целевую аудиторию поделиться нашим контентом, чем просто получить их просмотры, даже если не говорить об очевидной выгоде от увеличения охвата аудитории. Мы используем контент, чтобы решить коммуникационные задачи для нашего бренда, и если контент помогает нам в этом, то бренд также выигрывает.

Восьмая волна исследования убедительно демонстрирует, что пять вышеперечисленных потребностей потребителя можно раз-

бить по различным запросам аудитории, и эту классификацию бренды должны учитывать во время кампаний в социальных медиа.

Несколько примеров такого рода приведены ниже.

- В то время как потребители люксовых товаров и/или услуг хотят видеть вдохновляющий контент от своих брендов, вдохновение имеет мало значения для потребителей на рынке продуктов для ухода за детьми.
- В товарах FMCG, например, категории товаров личной гигиены, больше всего ценится практичный контент — тот, который дает знания о бренде, помогает практическими рекомендациями и отвечает на вопрос «как?».
- 25% потребителей хотели бы, чтобы разработчики видеоигр делились с ними контентом, например, в виде головоломок-квестов, в процессе выполнения которых игроки получали бы внутриигровые или реальные бонусы.
- А если говорить о такой целевой аудитории, как топ-менеджеры компаний, то они хотят получать тот контент, который помогает им получать новые знания из области своих интересов.

Наглядную разницу между различными аудиториями и их потребностями в контенте можно увидеть на диаграммах 1 и 2.

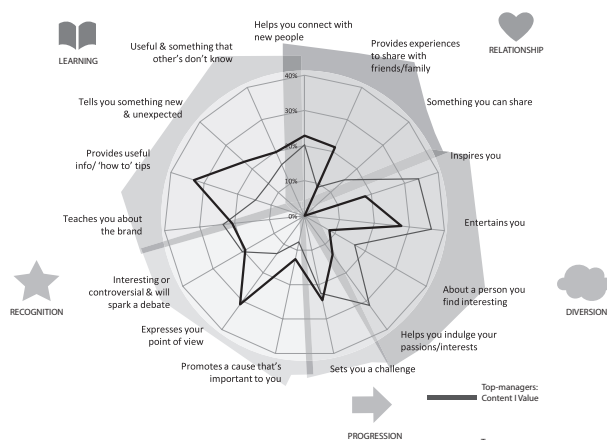


Диаграмма 1

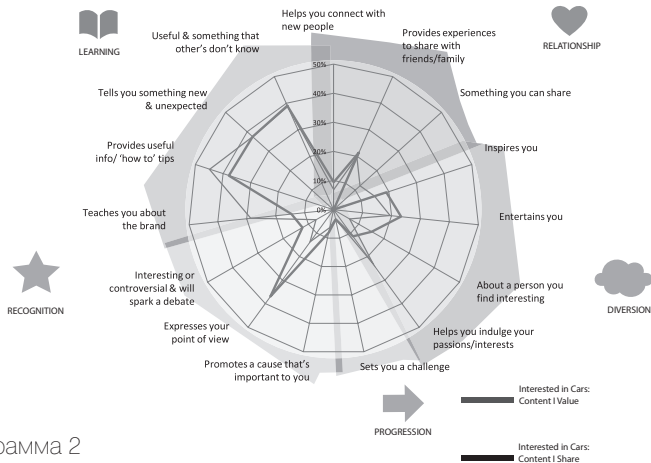


Диаграмма 2

Мотивации сильно варьируются и в зависимости от страны. Покупатели на развивающихся рынках с большей вероятностью поделятся информацией о товарах и услугах, которые они покупают: подавляющее большинство — 88% в Индии и 85% в Мексике — делают это минимум раз в месяц, тогда как в Великобритании этот показатель составляет не более 60%. В России данный показатель также высок и составляет 84,3%.

Важно учитывать, что когда мы от имени бренда предлагаем разные виды контента потребителю, то они вызывают соответствующий коммуникационный эффект. Например, развлекающий и забавляющий контент лучше всего работает, когда нужно сделать бренд желанным, в то время как контент, который дает потребителю новые знания о бренде, лучше всего подходит для того, чтобы заставить потребителя попробовать продукт.

Распространители контента

Данные, полученные по итогам исследования WAVE 8, еще раз подчеркивают, что даже среди людей старше 45 лет, заходящих в социальные сети хотя бы раз в два дня (возрастная группа, представи-

тели которой с наименьшей степенью вероятности будут размещать контент), 68% делятся фотографиями или видео как минимум раз в месяц. Возрастная группа, представители которой с наибольшей вероятностью будут делиться чем-то увиденным по телевизору как минимум раз в месяц — это 25–34-х летние. В Великобритании активные интернет-пользователи являются наиболее продуктивными создателями и распространителями контента, особенно при помощи смартфона³. В предыдущие 6 месяцев 2014 года более 23% активных британских интернет-пользователей загружали собственное видео, при этом более 10% приходится на возрастную группу старше 45 лет. В России ситуация несколько иная — всего видео-контент в интернет загружали 20,6% опрошенных респондентов, но при этом доля аудитории в возрасте 45+ составила 18,1%. Данные цифры могут говорить о том, что Россия все еще «отстает» в освоении возможностей интернет, но при этом старшее поколение более открыто для новых возможностей, нежели в тех странах, где «цифровая революция» произошла несколько ранее чем в России. Однако, как наглядно демонстрируют результаты всех волн исследования Wave, интернет-пользователи России очень быстро осваивают новые возможности Сети и где-то даже опережают в этом представителей других стран.

Это всё про меня!

Разумеется, катализатором новой эры распространения информации являются наше растущее присутствие онлайн и наша возрастающая зависимость от цифрового пространства. 57% интернет-пользователей в мире и 58,5% в России говорят, что испытывают серьезный дискомфорт, когда не имеют возможности войти в интернет. Данное обстоятельство вкупе с ростом количества портативных устройств у потребителей и с распространением бесплатных

³ При этом не стоит забывать про тот немаловажный факт, что Великобритания стала на сегодняшний день, пожалуй, самой активной страной с точки зрения потребления интернета. Многие аналитики прогнозируют, что в самом ближайшем будущем интернет станет медиа-каналом № 1 для медиа-рынка Великобритании.

точек Wi-Fi сделало большинство из нас «постоянно подключенными к Сети», что, согласно исследованию Wave 8, тоже имело влияние на наше эго.

На следующий год после того, как Оксфордский словарь английского языка объявил «селфи» словом года, 45% опрошенных согласились, что они чаще делятся фотографиями с собственным изображением, чем фотографиями, на которых изображены другие люди. Более того, 45% людей в возрасте от 16 до 24 лет в мире и 29,7% в России по итогам исследования отметили, что им бы хотелось быть знаменитыми в Сети. Среди людей старше 45 лет такие амбиции есть всего лишь у 13,1% респондентов. При этом крайне интересным является тот факт, что чем моложе респондент, тем больше и ярче у него выражена данная потребность. В России разница между возрастными группами наиболее значительна по сравнению с другими странами (см. Диаграмму 3).

Социальные медиа практически полностью изменили человеческие амбиции среди разных поколений. Положительная реакция других на то, что мы говорим или делаем, приносит нам счастье, то же касается и случаев, когда мы делимся контентом онлайн. Представители брендов, которые это поймут, дадут своим молодым



Диаграмма 3

потребителям тот контент, который удовлетворит их эго и поможет им достичь личных целей. В результате в выигрыше останутся как потребители, так и сами производители брендов.

Более того, связь между онлайн-пространством и нашей возможностью удовлетворять свои амбиции стала намного более прочной. На четверть больше активных интернет-пользователей считают социальные сети хорошим местом для заработка денег по сравнению с 2012 годом (28% против 23%, см. Диаграмму 4). При этом все более активно респонденты используют для данной деятельности профессиональные социальные сети, например, LinkedIn. Так, за один год количество их пользователей в России выросло с 34 до 39% (см. Диаграмму 5).

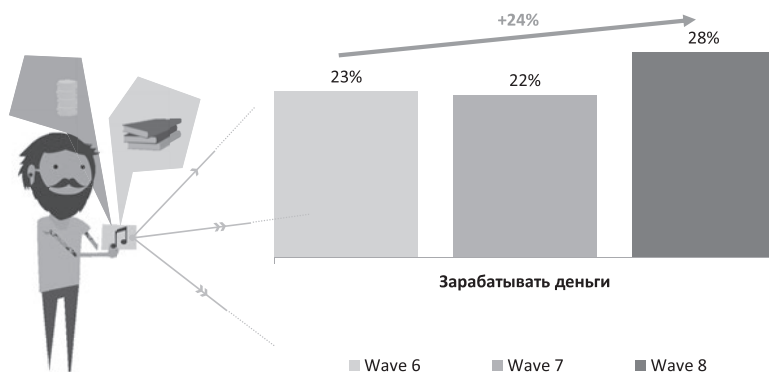


Диаграмма 4

Обеспокоенность сохранностью личных данных

В ходе исследования Wave 8 была также выявлена интересная дихотомия между стремлением пользователей интернет стать знаменитыми онлайн и их желанием сохранять анонимность. Обеспокоенность сохранением личных данных в сети постепенно растет. Пять лет назад 61% активных интернет-пользователей выражали озабоченность количеством личных данных, которые они держат

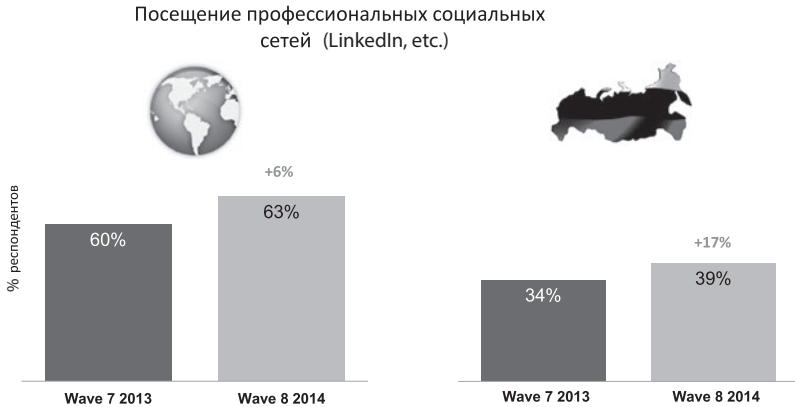


Диаграмма 5

онлайн. К 2013 году, когда произошел скандал с АНБ США⁴, эта цифра выросла до 67%, а в этом году на фоне новостей о том, что хакеры похитили личные фотографии, хранящиеся на iCloud⁵, эта цифра вновь выросла — до 69%.

Наше стремление к анонимности и в то же время к коммуникации ведут нас в разных направлениях. 58% говорят, что хотели бы

⁴ В 2013 году Эдвард Сноуден передал в СМИ документы, подтверждающие слежку АНБ за мобильными данными резидентов и нерезидентов США. Сбор данных при этом осуществлялся при помощи программного комплекса PRISM. Широкой общественности о существовании программы стало известно 6 июня 2013 года, когда отрывки из секретной презентации о PRISM были опубликованы в газетах «Вашингтон пост» и «Гардиан». По оценкам Washington Post от 2010 года, ежедневно системы сбора информации АНБ (в том числе PRISM) перехватывали и записывали около 1,7 миллиардов телефонных разговоров и электронных сообщений и около 5 миллиардов записей о местонахождении и передвижениях владельцев мобильных телефонов по всему миру

⁵ О массовом взломе учетных записей знаменитостей стало известно 31 августа 2014 года, когда почти две сотни частных снимков были размещены на ресурсе 4chan, а позже на reddit, anon-ib и tumblr. Вероятной причиной утечки стала уязвимость в сервисе iCloud, синхронизирующем фотографии с iPhone, хотя сами представители компании Apple опровергали взлом iCloud. Неизвестные выложили часть полученных фотоматериалов, а полный нецензурированный архив в течение нескольких недель пытались продать.

оставаться анонимными во всем, что они делают в сети. При этом 47% из тех, кто хочет оставаться анонимным, также соглашаются с тем, что им бы хотелось быть знаменитыми онлайн и использовать свой логин от социальных сетей для доступа к другим сайтам. По мнению Глена Паркера (глава IPG Media brands Marketing Sciences G14), это говорит о «пугающем недостатке понимания того, откуда у социальных сетей появляются личные данные».

Смартфон правит миром

Впервые за 8 лет исследований Wave самыми распространенными устройствами в мире среди частых интернет-пользователей стали смартфоны, обогнавшие стационарные компьютеры и ноутбуки. По сравнению с прошлым годом распространенность первых выросла на 15%, почти удвоившись за два года и достигнув 84% (См. Диаграмму 6). За этот же самый период в России количество владельцев смартфонов выросло на 16,5% по сравнению с прошлой волной исследования Wave 7 и в 2,5 раза (!!!) по сравнению с данными двухгодичной давности (Wave 6, см. Диаграмму 7). Выросло не толь-

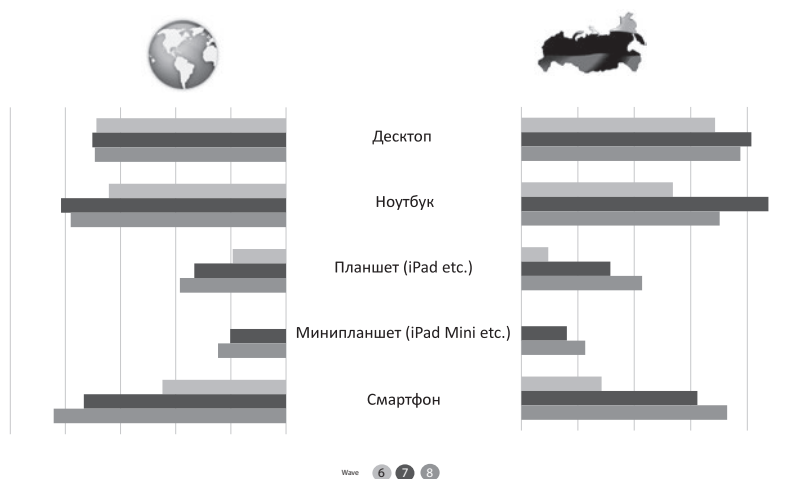


Диаграмма 6

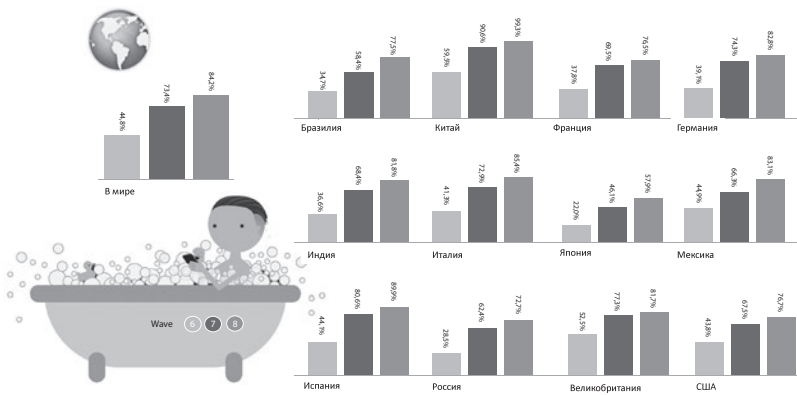


Диаграмма 7

ко количество пользователей смартфонов, но и количество людей, которые считают их удобным средством для того, чтобы делать покупки, — с 30,9% в 2012 году до 42,7% сейчас. Несмотря на выросшее с 41,6% до 51,2% проникновение планшетов, смартфоны вытесняют их в качестве первого устройства для любой деятельности. Мы видим «гравитационную» тягу к смартфонам в качестве главного устройства для создания и распространения контента. Это очень хорошо заметно на тех действиях, которые совершают пользователи при помощи различных устройств. Как видно на Диаграмме 8, большая часть социальных действий совершается при помощи мобильных устройств; например, это просмотр видео онлайн, использование сервисов микроблогов или посещение профессиональных социальных сетей.

Таким образом, в настоящий момент мы наблюдаем удивительный симбиоз мобильных технологий (прежде всего — смартфонов), распространяемого пользователями разнообразного контента и удовлетворения потребностей интернет-пользователей, изменяя при этом привычные способы коммуникаций. Данный симбиотический феномен демонстрирует нам, как меняется привычная среда общения, когда пользователи стремятся добиться уважения друг друга, используя при этом сочетание мобильных технологий и воз-

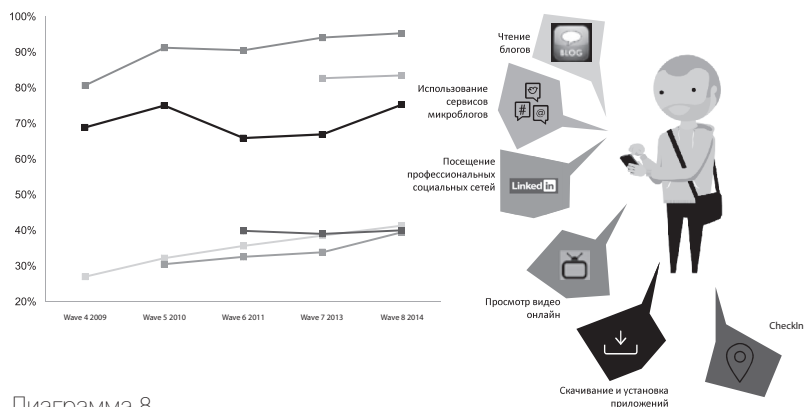


Диаграмма 8

возможностей, которые предоставляют социальные сети. Только поняв идеальное сочетание нужд и амбиций целевой аудитории, предоставив тот контент, который сам сможет стать социальным продуктом, компании-производители смогут построить сильные и взаимосвязанные отношения с потребителями.

Аннотации
и сведения
об авторах

Summaries
and information
about the authors

Русский

**От «виртуальной лаборатории» до «социального телескопа»:
метафоры тематических и методологических инноваций
в онлайн-исследованиях**

Девиатко Инна Феликсовна — д.соц.н., профессор, зав. кафедрой анализа социальных институтов департамента социологии Национального исследовательского университета «Высшая школа экономики», зав. отделом теории и истории социологии Института социологии РАН (e-mail: ideviatko@hse.ru, deviatko@gmail.com)

Ключевые слова: интернет-исследования, онлайн-исследования, методологические инновации, методы социальных наук, социальные сети.

В статье представлен аналитический обзор ключевых публикаций в социологии и сопредельных дисциплинах (прежде всего в психологии и когнитивной науке), позволяющих продемонстрировать произошедшие с начала 2000-х гг. изменения в тематическом репертуаре онлайн-исследований, в типах доступных данных, а также в понимании природы Интернета как среды для прикладных и академических исследований. Особое внимание уделяется основным методологическим инновациям, а также возможностям и вызовам, возникающим по мере превращения онлайн-исследований из выполняющего вспомогательные функции «зеркала» традиционных лабораторных экспериментов, опросов и техник включённого наблюдения в преобладающую форму исследования глобальных эффектов микросоциального взаимодействия и источник новой «социальной феноменологии» для многих исследовательских направлений, в том числе для масштабных экспериментальных исследований социального влияния, распространения информации и культурных образцов, мобилизации и т.п.

Опыт прогнозирования развития российской онлайн-индустрии

Давыдов Сергей Геннадьевич — к.ф.н., доцент, заместитель декана Факультета коммуникаций, медиа и дизайна НИУ ВШЭ (e-mail: sdavydov@mail.ru); **Логунова Ольга Сергеевна** — к.соц.н., доцент Факультета социальных наук НИУ ВШЭ (e-mail: olga.logunova@gmail.com); **Казарян Карен Размикович** — главный аналитик Российской ассоциации электронных коммуникаций, главный редактор журнала «Интернет в цифрах» (e-mail: kazaryan@taes.ru).

Ключевые слова: «Экономика Рунета», экспертный опрос, онлайн-опрос, онлайн-рынки, интернет-реклама, электронная коммерция, экономический кризис

Статья посвящена исследовательскому проекту «Экономика Рунета», совместно реализуемого НИУ ВШЭ и РАЭК с 2012 г. Подробно рассматриваются исследования второй волны (2013 г.), в рамках которой был сделан прогноз развития российской онлайн-индустрии на пять лет (2013–2018). Описаны четыре сценария развития (консервативный, стабилизационный, инновационный, кризисный) развития интернета и его четырех сегментов. Также приведены результаты четвертой волны (2015 г.), посвященные представлениям экспертов об экономическом кризисе в российской онлайн-индустрии и в стране в целом.

Дизайн исследований в стиле Бонсай

Джон Пулестон — вице-президент по внедрению инноваций Global Market Insite, INC (GMI) (e-mail: jpuleston@gmi-mr.com).

Ключевые слова: дизайн, исследования

В статье рассматриваются необходимые для дизайна исследования техники в стиле Бонсай: короткого и прекрасно сформулированного опроса.

Не всегда короче значит лучше

Инна Бурдейн — доктор наук, директор отдела анализа панельных опросов The NPD Group (e-mail: Inna.Burdein@npd.com)

Ключевые слова: длительность опросов, структура анкеты.

Цель этой статьи состоит в том, чтобы обогатить исследования по длительности опросов за счет более глубокого изучения самой концепции длительности и ее составляющих, а также описания взаимодействия, возникающего между структурой анкеты и человеком, ее заполняющим.

Новое поколение онлайн-анкет? Использование преимуществ интерактивности Интернета

Мелани Ревилла — доктор наук, Исследовательский и экспертный центр методологии опросов (RECSM) — Университет Помпеу Фабра (e-mail: melanie.revilla@upf.edu); **Карлос Очоа** — Netquest (e-mail: cochoa@netquest.com); **Альберт Турбина** — Netquest (e-mail: aturbina@netquest.com).

Ключевые слова: *онлайн-опросы, интерактивные анкеты, качество, время ответа, впечатления от участия в опросе, Netquest.*

Разнообразные методики, связанные с инновационными способами сбора данных (например, большие данные), оказывают давление на традиционные анкетные опросы. Мы уверены, что для достижения вовлеченности респондентов, необходимой для получения от них данных высокого качества в рамках онлайн-опросов, следует предложить анкеты нового поколения, позволяющие совместить приятные впечатления от участия с хорошим качеством полученной в итоге информации. Хотя разработка онлайн-анкет по образу и подобию анкет бумажных до сих пор считается нормой, имеется возможность создавать динамические опросы, которые позволяют сфокусировать респондентов на каждом выполняемом задании по очереди. Данная работа посвящена стратегии совершенствования опросных анкет за счет использования преимуществ интерактивности интернета, основанной на сочетании предварительного тестирования, лабораторного эксперимента и опроса. Полученные результаты свидетельствуют о позитивном впечатлении от участия в подобных опросах в целом, а также значимом повышении качества собранных в итоге данных. Гораздо большее число респондентов четко выполнили особые инструкции, включенные в анкету для проверки их внимательности при чтении заданий. Таким образом, подобные динамические анкеты перспективны для сбора данных в рамках дальнейших исследований.

Как влияет приглашение на отклик в онлайн-опросах?

Айгуль Маратовна Мавлетова — старший преподаватель, департамент социологии, факультет социальных наук, Национальный Исследовательский Университет Высшая школа экономики (e-mail: amavletova@hse.ru); **Наталья Геннадьевна Малошенок** — младший научный сотрудник, Институт образования, Национальный Исследовательский Университет Высшая школа экономики (e-mail: nmaloshonok@hse.ru); **Евгений Андреевич Терентьев** — ведущий аналитик, Центр внутреннего мониторинга, Национальный Исследовательский Университет Высшая школа экономики (e-mail: eterentev@hse.ru).

Ключевые слова: *онлайн-опросы, электронные приглашения, исследования.*

Исследователи с особой тщательностью подходят к составлению электронного приглашения к онлайн-опросу, поскольку именно на его основании большинство респондентов принимают решение о заполнении анкеты. Как правильно составить приглашение и что в него включить? В данной работе мы даём практические советы, опираясь на результаты экспериментальных исследований. Мы показали, что предварительное уведомление о проведении исследования, персональное обращение к респондентам по имени (или имени и отчеству), а также предоставление вознаграждения

увеличивает отклик в онлайн-опросах. Мы также указываем, какая тема письма, отправитель и длина приглашения могут увеличить количество участников опроса. Учитывая рост числа мобильных пользователей, особое внимание стоит уделить мобильной оптимизации приглашения.

Эффект знакомства при формировании приглашения онлайн-опросах: опыт эмпирического изучения

Гаврилов Кирилл Андреевич — к.соц.н., доцент кафедры анализа социальных институтов Национального исследовательского университета «Высшая школа экономики», научный сотрудник Института социологии РАН (e-mail: gavrilov@socio.msk.ru).

Ключевые слова: *онлайн-опрос, доля ответивших, эффект знакомства, профессиональная солидарность.*

Изучается влияние двух факторов на принятие приглашения поучаствовать в онлайн-опросе среди студентов вуза: фактор личного знакомства респондентов с отправителем приглашения, а также фактор, связанный с профессиональной солидарностью респондентов-социологов (случай, когда их приглашает преподаватель-социолог). Использовалась база данных из 253 студентов 2 курса двух факультетов НИУ ВШЭ. Всего на вопросы анкеты ответили 49% студентов. Проведение дисперсионного анализа показало наличие значимого эффекта для фактора знакомства: если приглашение отправлял знакомый преподаватель, то доля ответивших составила 58%, а если просьба исходила от незнакомого лица — 44%. В то же время эффект фактора профессиональной солидарности не наблюдался.

Нужно ли оставлять выборку на волю случая?

Дэбра Сантус, вице-президент Sample Strategies TNS Research International; **Питер Квок** TNS Research International; **Франк Келли**, старший вице-президент по глобальному маркетингу и стратегии Lightspeed GMI (e-mail: frankmkelly@hotmail.com).

Ключевые слова: *квотирование, выборка, опрос.*

Выбор подходящего метода квотирования является критичным, если мы хотим быть уверенными, что наши выборки позволяют обеспечивать точные, воспроизводимые и высококачественные результаты. Принимая это во внимание, мы сфокусировали дизайн нашего исследования и анализ его результатов на важности установки квот как при приглашении к опросу, так и на завершённые интервью в отличие от стандартного подхода, когда квотируются только завершённые интервью. Создание исходящей выборки, которая сбалансирована с использованием соответствующих переменных и далее скорректирована с учетом уровня отклика, позволяет быть уверенными, что группа респондентов, заполняющих опрос, будет соответствовать целевой аудитории.

Применение выборки, направляемой респондентом в онлайн-исследованиях: методологические вызовы и исследовательские перспективы.

Моисеев Станислав Павлович — аспирант Факультета социальных наук Национального исследовательского университета «Высшая школа экономики», исследователь Международной лаборатории прикладного сетевого анализа НИУ ВШЭ, руководитель исследовательских проектов ООО «инФОМ» (e-mail: spmoiseev@gmail.com); **Савинкова Юлия Константиновна** — менеджер исследовательских проектов ООО «Ипсос».

Ключевые слова: Выборка, направляемая респондентом; RDS; WebRDS; методология онлайн-исследований, выборочные методы.

Возможность использования выборки, направляемой респондентами для проведения онлайн-опросов, была представлена научному сообществу около семи лет назад. За прошедшее время зарубежные и отечественные исследователи провели ряд экспериментов по применению WebRDS, однако их результаты сложно считать однозначными. С одной стороны, некоторые авторы называют WebRDS удобной и многообещающей исследовательской технологией. С другой, часть исследователей указывает на сложности в применении выборки, направляемой респондентами, и неудачные попытки получить несмещенные оценки параметров генеральной совокупности. Обозначенное противоречие обостряется более широкой дискуссией о возможности обеспечить соблюдение строгих допущений, на которых основана реализация RDS и ее онлайн-разновидности. В связи с этим в данной статье проводится анализ экспериментов по созданию выборки, направляемой респондентами в онлайн-исследованиях, и обобщаются результаты последних методологических исследований по развитию RDS. В заключении статьи даются рекомендации по проведению методологических исследований WebRDS.

Эффект самоотбора при формировании онлайн-панели: опыт первого года лонгитюдного исследования студентов НИУ ВШЭ

Малошонок Наталья Геннадьевна — к.с.н., младший научный сотрудник Института образования НИУ ВШЭ (e-mail). **Семенова Татьяна Вадимовна** — аналитик, Института образования НИУ ВШЭ (e-mail: tsemenova@hse.ru). **Терентьев Евгений Андреевич** — ведущий аналитик Центра внутреннего мониторинга НИУ ВШЭ.

Ключевые слова: эффект самоотбора, смещения в данных, онлайн-панель, лонгитюдное исследование

В статье анализируется опыт первого года реализации лонгитюдного исследования студентов, проводящегося с помощью онлайн-опроса. Приводится описание процедуры набора респондентов в панель, а также их удержания в течение первого года исследования. Кроме того, предпринимается попытка продемонстрировать возможные смещения в данных, свя-

занные с эффектом самоотбора. Эффект самоотбора в работе определяется с помощью сравнения студентов, согласившихся на участие в лонгитюдном исследовании, с теми студентами, которые не стали участвовать в данном исследовании. Сравнение проводится по социально-демографическим характеристикам, социальному и культурному капиталу, характеристикам обучения, причинам поступления в вуз и лояльности к нему, а также по показателям исходной подготовки и академической успеваемости в университете.

Использование исследовательских игр вместо геймифицированных опросов

Влияние метода исследовательских игр на вовлеченность респондентов и вероятность их будущего участия в подобных проектах

Бетти Адаму — Research Through Gaming Ltd, Лондон, Великобритания (e-mail: betty.adamou@researchthroughgaming.com); **Дэвид Беркс** — Бизнес-школа при Университете Винчестера, Великобритания (e-mail: david.birks@winchester.ac.uk).

***Ключевые слова:** игры, исследовательские игры, Research Game, геймификация, опросы, IMPRINTS Futures, вовлеченность респондентов, сбор данных*

Геймификация используется в отрасли маркетинговых онлайн-исследований для повышения уровня вовлеченности респондентов. Высокая вовлеченность дает множество преимуществ, таких как повышение доли ответивших на все вопросы анкеты и качества полученных данных, повышение удовлетворенности от участия и возможность понять реальный образ мыслей респондентов.

Некоторые специалисты убеждены, что исследовательские игры позволяют обеспечить более высокий уровень вовлеченности респондентов по сравнению с геймификацией. По мнению ученых, в т.ч. специалистов по играм, геймифицированные задания менее увлекательны по сравнению с играми, потому что последние способны вызвать у игроков азарт и соответствующие контексту эмоции. Таким образом, данная работа посвящена не геймификации и геймифицированным опросам, а игровой механике и игровым компонентам в рамках онлайн-исследований, предполагающих создание полноценных игр, для обозначения которых используется термин «исследовательские игры».

В данной работе представлены выводы о вовлеченности респондентов и эмоциональном отклике, обусловленном контекстом, в рамках двух исследовательских игр, разработанных в целях сбора данных для научного исследования. В работе показано существенное положительное воздействие обеих исследовательских игр на участников, в результате которого 81% респондентов продолжили участие в проекте. В сочетании с другими показа-

телями, свидетельствующими в пользу предложенной методики, например, более 90% положительных откликов от участников, эти результаты свидетельствуют, что это многообещающая методика. Её имеет смысл тестировать, развивать и обсуждать, поскольку высокий уровень вовлеченности респондентов приносит пользу как самим респондентам, так и клиентам, заинтересованным в получении исследовательских результатов.

В качестве иллюстративных примеров в данной работе будут использованы две исследовательские игры, запущенные в Великобритании весной 2013 года под названиями T.E.S.S.A. Undercover AgentsTM и DubiousTM.

Предсказание будущего

Джон Пулестон — вице-президент по внедрению инноваций Global Market Insite, INC (GMI) (e-mail: jpuleston@gmi-mr.com); **Александр Уитли** — исследователь инноватор Lightspeed GMI; **Хубертус Хофкирхнер** — главный футурист Prediki Prognosedienste GmbH (e-mail: us.office@prediki.com)

Ключевые слова: прогнозирование, онлайн-исследования, опросы.

Статья посвящена изучению науки прогнозирования. В ней рассматриваются вопросы об областях прогнозирования, применении этого метода, а также факторы, влияющие на способность делать надежные прогнозы, и изменение получаемых данных при адресации вопросов в форме просьбы сделать прогноз. В ней также представлены рекомендации по лучшим способам постановки подобных вопросов и лучшим методам проведения прогностических исследований.

Качественные онлайн-методы

Рожественская Елена Юрьевна — д.соц.н., профессор департамента социологии факультета социальных наук Национального исследовательского университета «Высшая школа экономики», ведущий научный сотрудник Института социологии РАН (e-mail: erozhdestvenskaya@hse.ru)

Ключевые слова: качественные исследования в интернете, опрос, интервью, групповая дискуссия, наблюдение, анализ документов

В статье рассматривается перспектива качественных социологических исследований онлайн, с основным фокусом на опросе, наблюдении и анализе документов. Онлайн условия требуют не только использования инновационных технологических разработок для контакта с полем, но и адаптации заслуживших признание в офлайне методологических приемов взаимодействия между исследователями и респондентами.

Автор полагает, что не существует одного решения за или против качественных онлайн-методов. Скорее, речь идет о методологически диверсифицированном доступе, который комбинирует подходы на базе интернета в зависимости от текущей задачи с помощью проверенных методов онлайн-исследования.

Проблема способа проведения в качественных интервью: скайп, онлайн, лицом к лицу

Колозариди Полина Владимировна — аспирант факультета социальных наук департамента социологии кафедры анализа социальных институтов НИУ ВШЭ. (e-mail: poli.kolozaridi@gmail.com).

Ключевые слова: интервью, качественные методы, качественные онлайн-исследования.

Статья представляет обзор и анализ опыта исследователей, связанный с проведением качественных интервью с помощью онлайн-инструментов (Скайп, текстовые сообщения, электронная почта).

Обзор включает классификацию видов интервью, опосредованных цифровыми технологиями, и анализ подходов к оценке качества исследования. Выделяются несколько ключевых аспектов осмысления роли технологий в интервью: эффективность, этические вопросы (приватность, распределение власти), проблема анонимности, различия в использовании технологий (от чатов и электронной почты до скайпа и специальных программ).

Автор приводит примеры исследования разных аспектов и предлагает варианты использования разновидностей метода: интервью через системы передачи мгновенных сообщений, по электронной почте, видеоинтервью.

Онлайн-панель как инструмент телевизионных измерений.

Искра Иван Сергеевич — старший консультант отдела медиа исследований ООО «ГФК-Русь» (e-mail: ivan.iskra@gfk.com). **Кетов Сергей Валерьевич** — руководитель отдела медиа исследований ООО «ГФК-Русь» (e-mail: sergey.ketov@gfk.com).

Ключевые слова: онлайн-панель, измерение телевизионной аудитории, исследование зрительских оценок телепрограмм, панель зрительских оценок

Быстрое развитие онлайн-среды и разработка широкого набора исследовательских онлайн-инструментов в последнее десятилетие создали принципиально новые возможности для исследования телевизионной аудитории. Во многих европейских странах с начала 2000-х годов онлайн-панели используются для постоянного и оперативного изучения зрительских оценок телевизионного контента.

В основе данной статьи лежит российский опыт реализации онлайн-панели зрительских оценок, существующей с 2013 года, и ставшей для телеканалов важным дополнительным инструментом качественной оценки телепрограмм, выходящих на российском телевидении. Основной упор в статье делается на методических вопросах обеспечения качества данных онлайн-панели как инструмента регулярных телевизионных исследований.

Поиск реальной целевой аудитории своего бренда.

Ткачев Олег Викторович — директор по стратегическому планированию брендингового агентства Wellhead, автор книги «Visual Бренд: притягивая взгляды потребителей». (e-mail: tkachev@whead.ru)

Ключевые слова: *Брендинг, стратегия, интернет-исследования, онлайн-исследования, кластерный анализ, сегментация потребителей.*

В статье поднимается проблема, связанная с поиском и исследованием психологии людей, скрывающихся за термином «целевая аудитория». Автором представлен практический опыт применения онлайн-технологий и использования альтернативных методов классификации потребителей для решения стратегических задач для FMCG-компаний.

Использование методов онлайн-исследований для прогнозирования поведения потребителей в кризис

Микаелян Инга Ваагновна — старший аналитик агентства РБК. research (e-mail: imikaelyan@rbc.ru); **Хитров Сергей Сергеевич** — старший аналитик агентства РБК. research (e-mail: skhitrov@rbc.ru).

Ключевые слова: *прогнозирование, поведение потребителей, онлайн-исследования.*

В статье рассматриваются вопросы использования методов онлайн-исследований для анализа и прогнозирования ситуации на потребительском рынке, а также оперативной фиксации нарождающихся сдвигов в потребительском поведении, что в конечном итоге позволяет опережать своих конкурентов.

Практикум. Безопросные методы: Big Data, аналитика социальных сетей.

Большие вопросы — большие ответы

Том Вейс — генеральный директор Genius Digital (e-mail: tom.weiss@geniusdigital.tv); **Франк Тюро** — Руководитель стратегического развития GfK (e-mail: Franck.Thureau@gfk.com); **Ник Вееш** — руководитель направления по развитию медиа (e-mail: Niko.Waesche@gfk.com).

Ключевые слова: *большие данные, опросы.*

В отчете компании GfK рассматривается тема «больших данных», их использования в составлении прогнозов, построении бизнес-моделей и отслеживания потребительского поведения. Ключевые менеджеры и управленцы из 14 медиагрупп в формате интервью рассказывают о преимуществах использования «больших данных» в индустрии вещания.

Классификация должностей: переход к парадигме Big Data

Лебедев Павел Андреевич — к. соц. н., руководитель направления исследований портала Superjob.ru (e-mail: p.lebedev@superjob.ru); **Шурыгина Дарья Алексеевна** — старший аналитик портала Superjob.ru (e-mail: d.shurygina@superjob.ru).

Ключевые слова: *большие данные, big data, кластеризация текстов, k-means, рынок труда, классификация должностей.*

В статье описан эксперимент по классификации текстовых данных методом k-means для целей автоматической группировки названий конкретных должностей (в вакансиях и резюме). Эксперименты проводились на выборке из 14 тысяч названий вакансий портала Superjob. В статье приводится алгоритм перехода к автоматическому определению синонимов должностей, позволяющему сократить время на обработку и анализ данных. В основе алгоритма лежит совмещение механизма работы с Big Data и ручной проверки на определенных этапах, что позволяет добиться высокой точности результата при снижении временных затрат.

Автоматизированные методы анализа больших массивов интернет-текстов

Олеся Кольцова — социолог, возглавляет лабораторию интернет-исследований в Петербургском кампусе Высшей школы экономики. Ее научные интересы включают изучение онлайн-общественного мнения, интернет-сообществ и связи между онлайн- и офлайн-сторонами социальных явлений с использованием методов автоматического анализа текстов и социальных сетей на больших данных; (e-mail: ekoltsova@hse.ru).

Ключевые слова: *автоматизированный анализ текстов, новые методы, исследования интернета в социальных науках, большие данные.*

В статье представлен обзор современных методов машинного анализа текстов, имеющих большой потенциал для изучения интернета в социальных науках. Рассматриваются три основные группы: (1) классификация, кластеризация и тематическое моделирование; (2) сентимент-анализ и извлечение мнений; (3) извлечение объектов, аспектов и фактов. Вторая часть статьи содержит краткий обзор российских исследований, предпринимающих первые попытки применения этих методов в социальных науках для изучения интернета.

WAVE 8: Контент и его ценность. Новый язык общения производителя бренда и потребителя.

Бородин Даниил Александрович — директор по исследованиям UM Russia (e-mail: Daniil.borodin@umww.ru, danil.borodin@gmail.com).

Ключевые слова: *интернет-исследования, онлайн-исследования, WAVE 8, социальные сети, активные интернет-пользователи, контент, ценность контента, потребности, коммуникационные задачи, личные данные, смартфоны, мобильные технологии.*

В статье описываются результаты глобального международного исследования WAVE 8, которое проводится рекламным агентством UM. В ходе данного исследования было установлено, что контент может иметь очень большое влияние на отношение потребителя к бренду, с которым связан данный контент. А от этого во многом зависят отношения потребителя и бренда — чем лучше контент бренда отвечает потребностям потребителя, тем выше уровень доверия к бренду потребителя. Также немаловажен и правильный выбор канала коммуникации потребителя с брендом — настольного компьютера, смартфона или планшетного устройства. Таким образом, при правильном определении потребностей интернет пользователей, канала коммуникации и вида контента можно выстроить прочные отношения между производителем и потребителем.

English

From “Virtual Lab” to “Social Telescope”: Metaphors of Theoretical and Methodological Innovations in Online Research

Inna F. Deviatko, Ph.D., Professor, Head of Department of Social Institute Analysis at the School of Sociology of the National Research University Higher School of Economics, Head of Theory and History of Sociology Division at the Institute of Sociology of the Russian Academy of Sciences (e-mail: ideviatko@hse.ru, deviatko@gmail.com)

Keywords: *Internet research, online research, methodological innovations, social science research methods, social networks*

The key trends and publications in social sciences are reviewed in the paper in order to reveal and present in a systematic way the changes in dominant research topics and methods of Internet research taking place since early 2000s as well as concomitant modifications in types of available data and in our understanding of the nature of Internet communications. Particular attention is paid to leading methodological innovations and to possibilities and challenges which arise as a leading function of online research shifts from an auxiliary “mirror” of conventional lab experiments, surveys and participant observations to a prevalent mode of research of global effects of micro-interactions and source of new “social phenomenology” for many research

fields, including large scale experiments on the topics of social influence and cultural diffusion, political mobilization, etc.

Forecasting the Development of the Russian Online Industry

Sergey G. Davydov, Ph.D., Associate Professor, Deputy Dean of the Faculty of Communications, Media and Design at the National Research University Higher School of Economics (e-mail: sdavydov@mail.ru). **Olga S. Logunova**, Ph.D., Associate Professor of the Faculty of Social Sciences at the National Research University Higher School of Economics (e-mail: olga.logunova@gmail.com). **Karen R. Kazarian**, Chief Analyst at the Russian Association of Electronic Communications, Editor-in-Chief of the *Internet in Numbers* magazine.

Keywords: “*The Economy of Runet*”, *expert survey, online research, online markets, Internet advertising, electronic commerce, economic crisis*

The paper is based on the results of the Runet Economy research project, jointly implemented by the NRU HSE and RAEC since 2012. Detailed analysis of the second wave (2013) data is proposed, containing the five-year forecast of the Russian online industry development (in 2013–2018). The four development scenarios are described (conservative, stabilization, innovation, crisis) for the Internet and its four market segments. The paper also provides the fourth wave (2015) results concerning experts’ opinions on the economic crisis in Russia and its online industry in particular.

Bonsai Survey Design

Jon Puleston, Global Market Insite, INC (GMI) (e-mail: jpuleston@gmi-mr.com)

Keywords: *design, research*

This paper will explore the tools and technique needed to design Bonsai surveys: short, perfectly formed pieces of market research.

Shorter Isn’t Always Better

Inna Burdein, Ph.D., Director of Panel Analytics The NPD Group (e-mail: Inna.Burdein@npd.com).

Keywords: *survey length, questionnaire structure.*

A search on “best practices” for online surveys on various market research websites will inevitably lead to advice on survey length. The most common sentiment is: Keep the survey short. The definition of “short” varies, but the general stance is the longer the survey, the more data quality, completion rates, and survey satisfaction will suffer. Recently with the mobile revolution, even greater focus has been placed on cutting up long and arduous surveys into quick and easy ones to alleviate respondent fatigue. This push for shorter surveys seems as much guided by actual research as it is guided by presumptions. There is a general assumption that shorter is better, time is limited, and sitting through a long questionnaire is tiresome.

A New Generation of Web Questionnaires? Taking Advantage of the Interactivity of the Internet

Melanie Revilla, PhD, RECSM — Universitat Pompeu Fabra (e-mail: melanie.revilla@upf.edu); **Carlos Ochoa**, Netquest (e-mail: cochoa@netquest.com); **Albert Turbina**, Netquest (e-mail: aturbina@netquest.com).

Keywords: *Web surveys, interactive questionnaires, quality, response times, survey experience, Netquest.*

All kinds of new methodologies that aim to collect data in innovative ways (e.g. big data) are putting pressure on the traditional surveys based on questionnaires. We believe that in order to obtain from respondents the necessary effort to provide high quality data through web surveys, it is necessary to propose to them a new generation of questionnaires, capable of combining a satisfactory experience with good data quality. Although it is still the norm to design online questionnaires much like paper questionnaires, it is possible to create dynamic surveys that may help make the respondents focus on the tasks one at a time. This work explores, by combining pretesting, laboratory experiment and survey, a development strategy for questionnaires that take advantage of the interactivity of the Internet. The results show an overall good survey experience, and some significant improvement of data quality: a much larger share of respondents followed properly specific instructions added to check that they were reading carefully. Therefore, these dynamic questionnaires could be considered for future data collection.

The Impact of Invitation on Response in Web Surveys

Aigul M. Mavletova, Senior Lecturer, School of Sociology, Faculty of Social Sciences at the National Research University Higher School of Economics (e-mail: amavletova@hse.ru). **Natalya G. Maloshonok**, Junior Researcher, Institute of Education at the National Research University Higher School of Economics (e-mail: nmaloshonok@hse.ru). **Evgeny A. Terentyev**, Chief Analyst, Center for Institutional Research at the National Research University Higher School of Economics (e-mail: eterentev@hse.ru).

Keywords: *electronic invitations, web survey, research.*

Since most of respondents decide to participate in a web survey based on an invitation, researchers pay attention to each element of an electronic invitation they send to the participants. How should researchers design invitations? In the paper we give some practical advice on how to design an invitation based on the results of experimental studies. Pre-notification, personalization and incentives increase response rates in web surveys. We also show how the subject, the invitation sender, as well as the invitation length can increase response rates. Considering the increasing proportion of mobile web respondents researchers should take into consideration mobile optimization of electronic invitations.

Effect of Acquaintanceship in Formation of Online Survey invitations: an Empirical Study

Kirill A. Gavrilov, Ph.D., Associate Professor of the Department of Social Institute Analysis at the National Research University Higher School of Economics, Researcher at the Institute of Sociology of the Russian Academy of Sciences (e-mail: gavrilov@socio.msk.ru).

Keywords: online survey, response rate, the effect of acquaintanceship, professional solidarity.

In our study we analyze the influence of two factors on the acceptance of an invitation to participate in an online survey among university students. The first factor is the acquaintanceship with the invitation initiator, the second — the professional solidarity of sociology students (when they are invited by a sociologist). We used a sample of 253 second-year students from two faculties of the Higher School of Economics. The overall response rate was 49%. ANOVA showed a significant effect for the acquaintanceship with the invitation initiator factor: if the request was from a familiar lecturer the response rate is 58%, whereas if the request came from an unknown person, the rate is 44%. At the same time the effect of the professional solidarity factor was not observed.

Should Sampling be Left to Chance?

Debra Santus, VP of Sample Strategy and Standards, TNS; **Peter Kwok**, Manager of the Marketing Sciences Center, TNS; **Frank Kelly**, SVP Global Marketing & Strategy, Lightspeed GMI (e-mail: frankmkelly@hotmail.com).

Keywords: quota, sample, survey.

Choosing an appropriate technique to manage quotas is critical to ensuring our samples are designed to provide accurate, replicable and high quality results. Keeping this in mind, our primary research design and analysis focuses on the importance of managing quotas for both sample invitations and survey completes vs. only managing quotas within the survey completes. Creating an outgoing sample that is balanced using the appropriate variables and further adjusted for expected response rates ensures that a group of respondents starting a survey matches the target audience.

Implementation of Respondent-Driven Sampling in Online Research: Methodological Challenges and Research Perspectives

Stanislav P. Moiseev, Ph.D. student of the Faculty of Social Sciences at the National Research University Higher School of Economics, Researcher of the International Laboratory of Applied Network Research at the National Research University Higher School of Economics, Head of Research Projects at OOO inFOM (e-mail: spmoiseev@gmail.com). Yulia K. Savinkova, Manager of Research Projects at OOO Ipsos.

Keywords: *RDS; WebRDS; online research methodology, sampling methods.*

The possibility to use respondent-driven sampling to conduct online surveys was introduced to the scientific community seven years ago. Since that time, foreign and domestic researchers conducted a series of experiments with WebRDS. However, their results have been ambiguous. On the one hand, some authors see WebRDS as a convenient and promising research technology. On the other hand, some researchers bring attention to the difficulty of applying respondent-driven sampling and unsuccessful attempts to obtain unbiased estimates of the target population parameters. The contradiction is further escalated by broader discussion of the possibility to follow basic RDS assumptions. Therefore, this paper analyzes implementation of respondent-driven sampling in online research and summarizes results of recent methodological discussion on the development of RDS. In conclusion, the paper provides recommendations on conducting methodological studies of WebRDS.

The Self-Selection Effect in Online Panel Formation: First Year Overview of Longitudinal Research of HSE Student Population

Natalya G. Maloshonok, Ph.D., Junior Researcher, Institute of Education at the National Research University Higher School of Economics (e-mail: nmaloshonok@hse.ru). **Tatyana V. Semenova**, Analyst, Institute of Education at the National Research University Higher School of Economics (e-mail: tsemenova@hse.ru). **Evgeny A. Terentyev**, Chief Analyst, Center for Institutional Research at the National Research University Higher School of Economics.

Keywords: *self-selection, bias, online panel, longitudinal research, research of student population*

The paper analyzes the experience of the first year of students longitudinal study conducted via online survey and presents the description of some procedures for respondent recruitment into the panel and their retention during the first year of research. Moreover, we attempted to demonstrate possible bias in the data derived from the effect of self-selection. The effect of self-selection is determined by comparing the students who agreed to participate in the longitudinal study, with those who refused to participate, with the comparison based on their socio-demographic characteristics, social and cultural capital, educational characteristics, reasons for university enrollment and university loyalty level, as well as indicators of initial attainment level and academic performance at the university.

ResearchGames as a Methodology: Moving Away from Gamified Surveys The Impact of Online ResearchGames Upon Participant Engagement and Future ResearchGame Participation

Betty Adamou, Research Through Gaming Ltd, London, United Kingdom (e-mail: betty.adamou@researchthroughgaming.com); **David Birks**, University of Winchester, Business School, Winchester, United Kingdom (e-mail: david.birks@winchester.ac.uk).

Keywords: *Games, ResearchGames, Gamification, Surveys, IMPRINTS Futures, Participant Engagement, Data Collection.*

Gamification is used within the online market research industry to grow participant engagement. Through increased engagement, an array of benefits are contended to be produced, such as increased response rates and data quality, increased participant enjoyment, and the ability for researchers to understand the true mindsets of participants. It is contended that ResearchGames (fully-fledged games made for online research) can produce higher levels of engagement than gamification, as ludologists and other academics assert that gamified activities are not as engaging as games, as games can induce 'flow' and in-context emotional states in players. Therefore this paper is focused on game mechanics and game components with online research when applied as a fully-fledged game, labeled as a 'ResearchGame', not applied as Gamification research or 'gamified surveys'. This paper offers explanations and results about the impact that two ResearchGames (that were created to provide data for an academic study), had on participant engagement as well as emotionally-driven, in-context participant responses. This paper highlights the significant positive impact both ResearchGames had on participants where an 81% continuation rate was produced. Combined with supportive evidence such as over 90% positive participant feedback, these results show that ResearchGames are a promising methodology that would benefit from further experimentation, development and discussion as high levels of participant engagement can benefit both the participant and the client requiring research results. Two ResearchGames, launched in the United Kingdom in Spring, 2013 named T.E.S.S.A. Undercover Agents™ and Dubious™ will be used as illustrative cases.

Predicting the Future: Primary Research Exploring the Science of Prediction

Jon Puleston, GMI (e-mail: jpuleston@gmi-mr.com); **Hubertus Hofkirchner**, Chief Futurist Prediki Prognosedienste GmbH (e-mail: us.office@prediki.com); **Alexander Wheatley** serves as Research Innovator, Lightspeed GMI.

Keywords: *prediction, online research, quiz.*

This paper is an exploration of the science of prediction. It examines what we are good and bad at predicting and why; what influences our ability to make reliable predictions and what differences arise in the data when we ask

questions predictively. It provides advice on the best ways to ask predictive questions and the best techniques for conducting predictive research. It is based on the learnings from over 30 research on research experiments conducted over a 3-year period, the meta analysis of over 500 predictions and a review of the prominent academic literature published on this topic. We hope this will act as a guide for the market research industry wishing to exploit prediction protocols in their research.

Qualitative Methods of Online Research

Elena Yu. Rozhdestvenskaya, Ph.D., Professor of the School of Sociology, Faculty of Social Sciences at the National Research University Higher School of Economics, Leading Researcher at the Institute of Sociology of the Russian Academy of Sciences (e-mail: erozhdestvenskaya@hse.ru)

Keywords: *qualitative research in the Internet, the survey, interviews, group discussions, observation, analysis of documents*

The paper discusses the prospects of qualitative sociological research online, with the main focus on survey, observation and analysis of documents. Online conditions require not only the use of innovative technological developments for contact in the field, but also the adaptation of recognized offline methods of interaction between researchers and respondents. The author believes that there is no single solution for or against online qualitative methods. Rather, it is a methodologically diversified access, which combines Internet-based approaches adequate to the current task with the help of proven methods in online research.

The Problem of the Qualitative Interview Method: Skype, Web-Based, Face-to-Face

Polina V. Kolozaridi, Ph.D. student of the Faculty of Social Sciences, School of Sociology, Department of Social Institute Analysis at the National Research University Higher School of Economics (e-mail poli.kolozaridi@gmail.com).

Keywords: *interview, qualitative research, qualitative online research.*

The paper provides an overview and analysis of the researchers' experience associated with the qualitative interviews using online tools (Skype, text messages, e-mail).

The review includes the digitally-mediated interview classification and analysis of research properties. There are several key aspects for technology role interpretation: effectiveness, different technologies usage (messaging, e-mail, Skype, etc.) and ethical issues, which are confidentiality, privacy, power relations, anonymity.

The author observes various research approaches and suggests how online mediated interviews might be used in the most productive way.

Online Panel as a Tool for TV Audience Measurement

Ivan S. Iskra, Senior Consultant of the Media Research Department, OOO GfK-Rus (e-mail: ivan.iskra@gfk.com). **Sergey V. Ketov**, Head of Media Research Department, OOO GfK-Rus (e-mail: sergey.ketov@gfk.com).

Keywords: *online panel, TV audience measurement, appreciation survey, appreciation panel system.*

A rapid development of the online environment in general together with a wide set of online research tools in particular developed in the last decade have created new opportunities for television audience research. Online panels have become a popular instrument for appreciation survey in many European countries since the early 2000s.

An experience in the implementation of an online Appreciation Panel System on request of a Russian TV channel is at the core of this paper. The panel was set in 2013 and has proven to be an important and valuable instrument for a qualitative assessment of TV programs broadcasted on the Russian TV. Methodological issues and data quality in regards to a continuous television audience research study, based on an online panel, are in the focus of the paper.

Finding the Real Target Audience for Your Brand

Oleg V. Tkachev, Strategic Planning Director a Branding Agency Wellhead, author of VisualBrand: Drawing the Eyes of Consumers. (e-mail: tkachev@whead.ru)

Keywords: *branding, strategy, internet research, online research, cluster analysis, consumer segmentation.*

The paper addresses the problem of search and research of psychology of the people that are designated with the term “target audience”. The author presents his practical experience with application of online technologies and use of alternative consumers’ classification methods for the solution of FMCG companies’ strategic tasks.

Application of Online Research Methods in Forecasting of Crisis-Time Consumer Behavior

Inga V. Mikaelyan, Senior Analyst at RBC.research (e-mail: imikaelyan@rbc.ru); **Sergey S. Khitrov**, Senior Analyst at RBC.research (e-mail: skhitrov@rbc.ru).

Keywords: *forecasting, consumer behavior, online research.*

The paper discusses the application of online research methods in analysis and forecasting of consumer market situation, and immediate identification of emerging shifts in consumer behavior, which ultimately offers a competitive advantage.

Big Questions, Big Answers. Will Harnessing Smart Data for Audience Analytics Save the Broadcast Industry?

Tom Weiss, CEO Genius Digital (e-mail: tom.weiss@geniusdigital.tv); **Franck Thureau**, Head of Business Development Media, Audience Measurement & Insights (e-mail: Franck.Thureau@gfk.com); **Niko Waesche**, Global Lead Media and Entertainment (e-mail: Niko.Waesche@gfk.com).

Keywords: *big data, survey.*

Big Data is still in its infancy. Companies are offering more and more subscription-based or registration-based services, across more devices, and are collecting mountains of data about their users and audiences. But most of this potential is currently going unused. Rohana Rozhan, CEO of Malaysia's leading pay-TV company Astro agrees, stating, "I would love to say we are very good with data but the reality is we are not yet and we are working hard to change that." Yet the industry can be dramatically reshaped by the insights Big Data has to offer. Broadcasters, content curators and advertisers can utilize this information and make predictions that will fundamentally change business models and revenue streams. But according to Jonathan Sykes, Chairman of Genius Digital, "more doesn't necessarily mean better... The bottom line is understanding that the behavioral data helps in all areas of the business, from strategy to operations, advertising, acquisition to churn management." In this paper key decision makers and executives from 14 media groups were interviewed to explore the benefits of Big Data in broadcast. Together, these companies serve over 70 million subscribers and deliver content that reaches nearly a billion people every day.

Classification of Job Positions: Shift to Big Data Paradigm

Pavel A. Lebedev, Ph.D., Head of Research, Superjob.ru (e-mail: p.lebedev@superjob.ru); **Daria A. Shurygina**, Senior Analyst, Superjob.ru (e-mail: d.shurygina@superjob.ru).

Keywords: *big data, text clusterization, k-means, labor market, classification of job positions,*

The paper observes an experiment on text classification with k-means clusterization method. The purpose is automatic grouping of job positions titles (vacancies and resumes). The experiment is based on a sample of 14,000 vacancies from Superjob recruitment portal. The paper presents the shift to automatic detection of positions synonyms, reduces the time for data analysis. The algorithm is based on the combination of Big Data techniques and manual inspection that leads to high accuracy of the result with reduced time expenses.

Automated Analysis of Big Internet Text Corpora

Olessia Koltsova is a sociologist and Director of the Laboratory for Internet Studies at St. Petersburg campus of the Higher School of Economics, Russia. Her research interests include studies of online public opinion, internet communities and the relation between online and offline aspects of social phenomena with the methods of automatic text analysis and social network analysis of big data; (e-mail: ekoltsova@hse.ru).

Keywords: *text mining, new methods, internet research for social science, big data.*

The paper presents a review of new methods for automatic text analysis that have a great potential for internet studies in social science. Three major groups of methods are being considered: (1) classification, clustering and topic modeling; (2) sentiment analysis and opinion mining; and (3) extraction of object, aspects and facts. The second part of the paper briefly reviews Russian studies that make the first steps applying these methods to social science Internet research.

WAVE 8: The Content and Its Value. The New Language of Brand — Consumer Communication

Daniil A. Borodin, Director for Research, UM Russia (e-mail: Daniil.borodin@umww.ru, daniil.borodin@gmail.com).

Keywords: *Internet research, online research, WAVE 8, social networks, active Internet users, content, value content requirements, communications objectives, personal data, smart phones, mobile technology.*

The paper describes the results of a global international study called WAVE 8, carried out by UM Advertising Agency. It was found in the study that the content may have a very important influence on the attitude toward the brand associated with the content. It has a direct impact on the consumer — brand relations — the better the content of the brand meets the needs of the consumer, the higher is the level of consumer confidence in the brand. Also of considerable importance is the right choice of channel for the consumer — brand communication: desktop, smartphone or tablet device. Thus, with proper identification of the Internet users needs, communication channel and type of content, it becomes possible to build a strong relationship between the manufacturer and the consumer.

Онлайн-исследования в России: тенденции и перспективы
Под редакцией Шашкина А.В., Девятко И.Ф., Давыдова С.Г.

Дизайн: Светлана Годубушина
Компьютерная верстка: Юлия Левашова
Корректор: Ольга Пружинина

Подписано в печать ???
Формат 60*90/16 Бумага офсетная. Печать офсетная.
Усл. печ. л. 26,25. Заказ № ???. Тираж 1000 экз.

ООО «Издательский дом «Кодекс»,
129626, г. Москва, Графский пер., д. 9, стр. 2.

Отпечатано в типографии ООО «ТИПОГРАФИЯ КЕМ»
129626, г. Москва, Графский пер., д. 9, стр. 2, тел. (495) 363-5612
E-mail: sekret@i-kem.ru