

Особенности реализации общероссийской выборки трудоспособного населения по мобильным телефонам

В статье представлены методические описания организации и проведения выборочного исследования по мобильным телефонам. Даны практические рекомендации по мониторингу качества исследования, регламенту раскрытия параметров выборки и показателей достижимости. Предложенный способ презентации методических материалов соответствует международным стандартам качества по проведению телефонных опросов и может рассматриваться как необходимое и достаточное условие информирования заинтересованной аудитории о методологии эмпирического исследования.

Ключевые слова: география опроса, общероссийская выборка, опрос по мобильным телефонам, коэффициенты достижимости, подходы к перевзвешиванию, ремонт выборки, стандарты качества

Тимур Османов

ведущий специалист Фонда "Общественное мнение" (Москва)

Дмитрий Rogozin

директор Центра методологии федеративных исследований РАНХиГС (Москва)

В середине 2000-х годов исследования по мобильным телефонам были довольно редки. Лишь немногочисленные опросные компании разрабатывали такого рода выборочные дизайны. Однако, нет сомнения, что в ближайшем будущем объем рынка исследований по мобильным телефонам будет существенно расширен, а в перспективе ближайшего десятилетия вытеснит на периферию исследовательского интереса привычные формы опросов (поквартирный, уличные опросы, только стационарные телефоны). На рынке западных исследований доминирующая роль опросов по мобильным телефонам уже не вызывает сомнения, хотя еще в начале 2000-х мало кто мог прогнозировать столь широкое их распространение. Повсеместно ведущие исследовательские компании отказываются от иных опросных технологий в пользу более гибкой, гибридной и дешевой мобильной связи [Vehovar, Berzelak, Manfreda, 2010, p. 304]. Во-первых, мобильная связь позволяет покрыть практически все группы населения, независимо от пространственной (удаленность территории), социальной (места с социальной напряженностью, например, республики Северного Кавказа) или статусной достижимости (высокодоходные или высоко-статусные группы) абонентов. Безусловно, некоторая коррекция проходит по расположению к общению, и отдельные группы неминуемо выпадают из выборки из-за отказов участия в опросе. Вместе с тем есть возможность учитывать отказы, проводить редактирование данных, осуществлять дополнительные мероприятия по снижению уровня неответов. Во-вторых, плотность покрытия территории России мобильной связью постоянно возрастает. В 2011 году по данным Фонда "Общественное мнение" более 90% населения пользовались мобильной связью [Османов, 2011]; растёт доля тех, кто имеет более одной сим-карты. В результате удешевления мобильной связи постепенно снижается доля пользователей стационарными телефонами. Если экстраполировать наблюдения Рэндала Цуваллака в 2012 году в США доля населения полностью отказавшегося от стационарной связи в пользу мобильной может приблизиться к трети:

В 2003 году уровень замены стационарной связи мобильной был низким: только 2,8% взрослого населения США, проживающего в домохозяйствах, имели лишь мобильные телефоны и 1,6% вовсе не пользовались телефонной связью [Blumberg, Luke, 2007]. В первой половине 2007 года в Национальном опросе по здоровью (National Health Interview Survey — NHIS) было зафиксировано уже более 12% граждан, пользующихся дома исключительно мобильной связью. Доля нетелефонизированных домохо-

зяйств осталось на том же уровне — 1,6%. Никто не может предсказать рост населения, владеющего только мобильными телефонами. Если опираться на данные Национального опроса по здоровью, то текущий прирост составляет 15-20% каждые полгода [Zuwallack, 2009, p. 389].

На рынке телекоммуникаций Россия, как правило, повторяет международные тренды с пяти-семилетним запозданием. Сейчас покрытие мобильной связью в России позволяет реализовывать случайные выборки, основанные лишь на префиксах операторов мобильной связи, и гарантировать репрезентацию не только пользователей мобильной связи, но и в целом населения страны по большому количеству исследуемых вопросов. В-третьих, современные технологии проведения телефонных опросов позволяют радикально улучшить качество собираемых данных и усилить контроль над возможными систематическими смещениями, вызванными недобросовестной или непрофессиональной работой интервьюеров. Тотальная запись всех телефонных разговоров с последующей выборочной проверкой качества интервью позволяет контролировать реальную коммуникативную ситуацию, а не только получаемые частотные распределения. Последние, как известно социальными исследователям, представляют собой весьма ненадежный источник для выявления вероятных отклонений от общей процедуры опроса. В-четвертых, запись всех интервью, позволяет освободиться от устаревшего методологического деления исследовательских технологий на количественные и качественные и заявить о реализации смешанного подхода (mix research) в рамках единичного исследовательского дизайна. Анализ частотных распределений, выявление корреляционных зависимостей и построение дисперсионных моделей уже может соседствовать с детальным дискурсивным анализом телефонных разговоров, выявлением стратегий ответа, особенностей аргументации и смысловых конструкций, применяемых в том или ином контексте. В-пятых, с распространением ip-телефонии, опрос по мобильным телефонам, даже при ужесточении способов контроля и усложнении анализа обработки данных, существенно дешевле традиционных личных интервью в расчете на единицу наблюдения, или респондента. Это связано с двумя факторами: первый — более низкая стоимость связи с абонентом, нежели, например, физический приезд к месту проживания, второй — дешевизна повторных обращений, независимо от установленного их количества (обычно ограничиваются пятью-семью дозвонами, поскольку в дальнейшем результативность опроса перестает заметно возрастать)¹. В-шестых, после законодательно установленного запрета взимать

¹ Американские исследователи (см., например: [Zuwallack, 2009, p. 397]) в качестве основной проблемы развития опросов по мобильным телефонам называют разную тарифную политику операторов по входящим звонкам.

плату за входящие звонки, продолжительные разговоры по мобильным телефонам уже не воспринимаются негативно абонентами и не интерпретируются как способ незаконного обогащения (в отличие, например, от обмена смс-сообщениями).

Кроме того, с выходом на массовый рынок мобильной связи смартфонов существенно расширены возможности опросов. Для пользователей современных приложений уже нет ограничений по форме коммуникации с интервьюером (визуальный контакт, обмен документами, заполнение электронных форм), что революционным образом изменяет представления об оффлайн и онлайн исследованиях [Lane, Miluzzo, Hong, et al., 2010; Vehavar, Berzelak, Manfreda, 2010; Сапонов, 2011, 2012].

В преддверии экспансии опросов по мобильным телефонам чрезвычайно важно определиться со стандартами качества их проведения. Иначе мы имеем все шансы получить низкокачественные исполнения опросных дизайнов, дискредитирующее целое направление исследовательской активности.

Стандарты качества

На этапе становления опросной технологии от исследователя требуется не только проводить методические экспериментальные планы по тестированию технологических нововведений, но и осуществлять детальную регистрацию методических параметров текущих проектов. Американская ассоциация исследователей общественного мнения (AAPOR) задаёт минимальные стандарты качества опросного инструмента, основной акцент в которых делается на контроль выборки и регистрацию систематических смещений, связанных с недостижимостью определенных групп населения [Standard definitions, 2011]². В минимальных требованиях к раскрытию информации AAPOR требует в обязательном порядке предоставление информации о следующих методических характеристиках выборочного исследования:

Описание исследуемого населения, географического расположения, и определение основы выборки, необходимой для идентификации этого населения. Если основа выборки предоставляется третьим лицом, поставщик должен быть назван. Если при формировании выборки не использовалась основа или список, это должно быть отражено в тексте.

Описание дизайна выборки, точное и понятное представление метода отбора респондентов, включая любое квотирование или дополнительные элементы отбора в ходе опроса или после его завершения. В описании выборки должно быть однозначно представлено где применялся случайный и неслучайный отбор респондентов.

Размер выборки и обсуждение точности измерения, включая ошибку выборки для случайного отбора и описание переменных, используемых для процедур взвешивания или оценки. Следует отдельно отметить, какие обнаруженные ошибки выборки, с учетом дизайн-эффекта, были скорректированы за счет кластеризации или взвешивания [Standards for minimal disclosure, 2010].

Итак, согласно международным нормам проведения массовых опросов необ-

ходимо уделять особое внимание трем методическим областям. Во-первых, особенностям формирования выборки, расписанным до единичных процедур, задающих правила организации основы выборки, последующих механизмов отбора единиц наблюдения и критериям оценки качества и редактирования данных. Во-вторых, географии опроса, или территориальному покрытию исследования. В-третьих, уровню достижимости, или особенностям реализации выборки. Наиболее важными показателями достижимости признаются уровень ответов, коэффициенты кооперации, отказов и контактов (подробнее см. ниже). Раскрытие информации по трем методическим областям является необходимым условием для оценки адекватности проведенного исследования. Только после получения такой информации можно приступать к анализу данных и построению моделей, не подвергая себя неоправданному риску производства ложного знания.

Цель настоящей работы — продемонстрировать минимальные требования к раскрытию методической информации, которым должны придерживаться все исследовательские коллективы при написании отчетной документации и научных статей, основанных на эмпирических материалах. Только обращаясь к реальному опыту можно получить вменяемые и точные представления о границах собственных методических возможностях и внешних требованиях к их реализации. Поэтому мы отказываемся от сухого изложения стандартизированных требований, и будем рассматривать данные конкретного исследования. Для исполнителей, исследовательских коллективов, мы предлагаем пример грамотного составления методической отчетности полевого этапа исследования. Для заказчиков, или лиц, формирующих технические задания на социальные исследования, настоящая работа может быть полезной в качестве реферативного руководства к оценке проделанной субконтракторами работы. Отсутствие внятных объяснений по каждому из обозначенных пунктов должно сигнализировать о ненадлежащем исполнении исследовательской процедуры и ставить под сомнения все результаты опроса.

Методика формирования выборки

Центром методологии федеративных исследований РАН-ХиГС совместно с АНО "Социальная валидация" в ноябре 2012 года проведено исследование представлений трудоспособного населения Российской Федерации о текущей пенсионной реформе. Исследуемой совокупностью являются граждане РФ

Таблица 1. Расчетное распределение выборки по федеральным округам

Федеральный округ	Количество респондентов, чел.	Доля в выборке, %
Дальневосточный	91	4,55
Приволжский	416	20,80
Северо-Западный	193	9,65
Северо-Кавказский	128	6,40
Сибирский	271	13,55
Уральский	172	8,60
Центральный	539	26,95
Южный (с 2010 года)	190	9,50
ИТОГО	2000	100,00

² Вплоть до 1970-х отсутствовали какие-либо стандарты качества по реализации выборочного исследования, при детальной разработке норм и правил должного дизайна выборочного исследования [Kviz, 1977]. Впервые вопрос о стандартизации коэффициентов ответов был поднят Американской статистической ассоциацией и только почти через десятилетие, в 1982 году, Советом американских организаций проводящих массовые опросы (Council of American Survey Research Organizations) была предпринята первая попытка стандартизации [Frankel, 1983]. Рабочий вариант стандартов качества выборочного исследования предложен в 1998 году Американской ассоциацией исследователей общественного мнения. И до сих пор это единственный полный и фундированный перечень подходов к измерению качества выборочного исследования.

Таблица 2. Особенности пользования мобильными телефонами в России,
% от ответивших, данные без взвешивания, репрезентируются
телефонные номера / репрезентируется население*.

		Скажите, пожалуйста, вы всегда один (одна) пользуетесь этим мобильным телефоном, или им пользуется несколько человек? если несколько, то сколько именно?					
		1	2	3	4	5 и более	Всего
Сейчас вы постоянно (не реже раза в неделю) пользуетесь <u>одним</u> номером мобильного телефона, сим-картой или несколькими? если несколькими, то сколькими именно?	Одним	56,8/64,8	3,4/7,8	0,6/2,1	0,1/0,6	0,2/1,4	61,3/76,9
	Двумя	26,3/15,1	2,6/3,0	0,5/0,9	0,2/0,4	0/0	29,7/19,5
	Тремя	5,7/2,2	0,3/0,2	0,2/0,2	0,1/0,1	0/0,1	6,4/2,9
	Четырьмя и более	1,5/0,4	0,0/0,0	0/0	0/0	0/0	1,6/0,5
	Затрудняюсь ответить	0,7/0,1	0,2/0,1	0/0	0/0	0,0/	0,9/0,2
	Всего:	91,2/82,7	6,7/11,2	1,4/3,3	0,4/1,2	0,3/1,5	100/100

*При репрезентации телефонных номеров за основу выборки принимается совокупность всех мобильных телефонных номеров. При репрезентации населения учитываются особенности пользования номерами телефонов (основа перевзвешивается с учетом количества людей, имеющих несколько мобильных номеров или пользующихся одним номером с другими членами семьи) - полученные цифры говорят о распределении параметров среди населения.

от 18 лет до возраста выхода на пенсию. Статистическая погрешность выборочного исследования с объемом выборки 2000 респондентов составит не более 2,3%. Опрашивались мужчины 18-59 лет и женщины 18-54 лет. Если человек данной возрастной группы был пенсионером, то он исключался из исследуемой совокупности. Для этого в самом начале интервью задавался скрининговый вопрос о профессиональном статусе абонента.

Выборка строилась посредством случайной генерации номера телефона по DEF кодам мобильных операторов. Первоначально вся выборка (2000 респондентов) распределялась по Федеральным округам (ФО) пропорционально численности населения исследуемых половозрастных групп (табл. 1).

Далее в каждом Федеральном округе из всей совокупности задействованных номеров мобильных телефонов проводился систематический отбор номеров в выборку с равным шагом (в пределах ФО) и случайной стартовой точкой. Такой механизм отбора номера равномерно распределял выборку по субъектам РФ в ФО, по операторам мобильной связи, по DEF диапазонам и, следовательно, по населенным пунктам. Смещения по данным параметрам могли возникнуть в основном из-за различий уровней достижимости. Отметим, что при таком дизайне выборки до проведения опроса нельзя определить точное количество респондентов, которые будут опрошены в определенных регионах и населенных пунктах.

На следующем этапе отбора опрашивался человек, который взял трубку. Предполагается, что в случае опроса по мобильным телефонам, такой отбор является случайным на индивидуальном уровне, следовательно, можно говорить о репрезентативности населения без дополнительных процедур отбора (например, как это номинально принято в поквартирных опросах, когда проводится отбор внутри домохозяйства).

Каждый номер в пределах ФО имел равную вероятность попасть в выборку. Однако вероятность отбора респондентов не была равной, а зависела от количества мобильных телефонов (точнее — сим-карт), которыми постоянно пользуется респондент, и количества людей, постоянно пользующихся мобильным телефоном, который попал в выборку. Чем больше мобильных телефонов у респондентов, тем больше у него веро-

ятность попасть в выборку, напротив, чем больше людей пользуется телефоном, который попал в выборку, тем меньше вероятность каждого из них быть опрошенным, поскольку по одному номеру опрашивается один человек. Для компенсации разной вероятности отбора респондентов в ходе опроса респонденту задавались вопросы о количестве мобильных номеров, сим-карт которыми он постоянно пользуется, (N_{моб}) и количестве людей, постоянно пользующихся "этим" телефоном (N_{респ}) (табл. 2).

Больше половины (около 57%) номеров использовались одним человеком, причем этот человек использовал лишь один номер. Каждый третий (34%) телефонный номер в России принадлежит человеку, который пользуется несколькими мобильными те-

лефонами. Каждый десятый номер (9%) используется несколькими людьми. В итоге вероятность отбора респондента считалась по следующей формуле:

$$p = \frac{N_{\text{моб}}}{N_{\text{респ}}}$$

где N_{моб} — количество мобильных телефонов, N_{респ} — количество респондентов.

Каждому респонденту присваивался вес обратно пропорциональный вероятности отбора, чтобы компенсировать разную вероятность отбора респондентов. Для исключения влияния на результаты опроса особенностей пользования мобильными телефонами в различных ФО, веса нормировались в пределах каждого ФО. В результате взвешивания исходное распределение выборки по ФО сохранилось.

По всей видимости, развитие мобильных технологий, расширение рынка планшетов и смартфонов, приведет к значительному увеличению числа владельцев нескольких сим-карт. Текущее состояние по количеству сим-карт на одного человека (77% пользуется одной сим-картой, 23% двумя и более), скорее всего, отражает ситуацию пяти-шести летней давности в странах Западной Европы и весьма близко к ситуации в странах Восточной Европы. Опираясь на разные источники, Марио Галлегаро с коллегами приводят следующие оценки пользователей нескольких сим-карт по странам Западной и Восточной Европы: в Италии на 2002 год 20-25% населения; в Словении на 2004 года 12-16%, на 2001 в Финляндии года 7% и в Португалии 4% [Callegaro, Steeh, Buskirk, et al, 2007, p. 650]. Маловероятно, чтобы количество пользователей нескольких сим-карт кардинально не увеличилось за последние годы и в этих странах. Поэтому необходимость коррекции вероятности отбора респондента в зависимости от количества сим-карт со временем будет только возрастать. Поскольку с ростом количества телефонов в домохозяйстве, практики совместного использования одного мобильного телефона становятся все реже [Tucker et al., 2005], можно предположить, что все меньше людей со временем будут делить свой телефон с кем-либо из членов семьи (в

нашем опросе доля таких людей 17%). Однако для контроля равной вероятности попадания в выборку отобранных респондентов следует в каждом опросе по мобильным телефонам задавать подобные вопросы, пока доля абонентов, относящихся к мобильному телефону как разновидности коллективной связи, не станет ничтожно мала.

Для проведения опроса задействовано два колл-центра (из Воронежа и Новосибирска). Поскольку предполагалось, что мобильный телефон респондент практически всегда держит под рукой, то время проведения опроса не ограничивалось стандартными вечерними часами в будние дни; опрос проходил с 11 до 21 часа по местному времени. Вместе с тем, после весьма краткого представления, все разговоры начинались с вопроса "Вам удобно говорить?" или "Вам удобно разговаривать?". Кроме формы вежливости, такая конструкция начала разговора, необходима и для обеспечения безопасности абонента [Zuwallack, 2009, p. 396]. Звонок может приходиться на время, когда абонент ведет автомобиль, занят каким-то делом, находится в публичном месте или ведёт весьма важный разговор. Переключение внимания на входящий звонок, может не только нарушать естественное развитие ситуации, но и угрожать жизни и здоровью абонента. Поэтому в дальнейшем важно как установить статистические границы наиболее комфортного времени разговоров, так и разделить скрининговый звонок (отбор респондента) от основного интервью. Последнее требует как большего времени, так и внимания, а значит должно различаться по основным параметрам ведения разговора (время, внимание, уместность в данном контексте).

Отклонение по социально-демографическим параметрам

Одна из угроз репрезентативности опроса по мобильным телефонам — неполный охват исследуемой совокупности, ведь шанс быть опрошенными имеют лишь пользователи мобильных телефонов. Однако согласно данным Фонда "Общественное мнение" летом 2011 года у 91% людей в возрасте от 18 до 60 лет был личный мобильный телефон. Следовательно, из выборочной совокупности исключается лишь 9% исследуемой группы. Это позволяет сделать экспертное предположение, что ошибка измерения, вызванная неполным покрытием целевой группы, не превысит 3-4% при наличии существенной зависимости исследуемых параметров от пользования мобильным телефоном. Если же такой зависимости нет, то неполное покрытие исследуемой группы не вносит систематической ошибки.

Судя по данным Фонда "Общественное мнение", уровень мобильной телефонизации снижается с возрастом. Например, среди людей в возрасте 18-44 лет 94% имеют мобильный телефон, среди людей 45-54 лет — уже 86%, а среди мужчин 55-59 лет — 82%. Данное явление должно вызывать занижение доли представителей старшего возраста в выборке. Как показывают расчеты, это занижение должно быть незначительным — старшая группа занижается на 1,3 процентных пункта, а самая молодая завышается на 0,9 процентных пункта. Однако в опросе этот эффект проявился значительно сильнее (табл. 3).

Самая молодая группа оказалась завышена относительно данных Росстата на 2,5 процентных пункта, люди от 25 до 34 лет завышены на 6,7 процентных пункта, люди от 35 до 44 занижены на 1,1 про-

Таблица 3. Распределение результатов опроса пользователей мобильными телефонами (ФОМ) и населения (Росстат) по возрастным группам, % по столбцу

Возрастные группы	Результат опроса	Пользователи моб.тел., ФОМ	Население, Росстат
18 - 24 лет	21,2	19,6	18,7
25 - 34 лет	33,9	28,0	27,2
35 - 44 лет	21,9	22,9	22,9
45 - 54 лет	20,5	25,7	26,0
муж 55-59 лет	2,6	3,9	5,2

центных пункта, люди от 45 до 54 занижены на 5,5 процентных пункта и мужчины от 55 до 59 лет занижены на 2,6 процентных пункта. Столь серьезные расхождения можно попытаться объяснить неоднозначной зависимостью между наличием мобильного телефона и вероятностью принять участие в опросе по этому телефону. Другими словами, гипотеза о том, что люди, у которых есть мобильные телефоны, имеют равную вероятность быть опрошенными по ним (с учетом наличия нескольких телефонов и нескольких людей, пользующихся телефоном) поставлена под сомнение. Для более глубокого понимания данного феномена требуется изучение различий практик пользования мобильным телефоном среди представителей различных социальных групп. Второй гипотезой, объясняющей столь значительные расхождения данных по возрасту, может служить предположение о неравномерном распределении отказов по социальным группам (в том числе, возрастным). Наконец, третья гипотеза — различия связаны с ошибками в регистрации населения, допущенными в последней переписи населения.

Если не принимать во внимание последнюю гипотезу, одним из выходов из сложившейся ситуации может быть наложение квотных условий на отбор респондентов. Однако, на наш взгляд, данная мера является поверхностной, поскольку в результатах опроса мы получим "правильное" распределение по полу и возрасту, не понимая истинной природы полученных смещений. Квотное задание также не гарантирует отсутствие систематических смещений по неквотируемым переменным. Помимо этого, наложение квотных ограничений не позволяет считать выборку случайной, а, следовательно, и репрезентативной в строгом смысле.

В данном же опросе, если предполагается сильная корреляция изучаемых переменных с возрастом, рекомендуется проводить расчеты с использованием весовых коэффициентов, выравнивающих распределение выборочной совокупности по возрасту. Исходя из предварительных результатов, в которых наблюдались статистически значимые зависимости между возрастом и большинством зависимых переменных, выборка была перевзвешена в соответствие с данными Росстата.

Использование таких весов не сильно сказалось на основ-

Таблица 4. Сопоставление результатов проведенного опроса с данными Росстата по образованию.

СКАЖИТЕ, ПОЖАЛУЙСТА, КАКОЕ У ВАС ОБРАЗОВАНИЕ?	Результат опроса	Результат опроса с весами на возраст	Данные Переписи населения 2010 года (м. 18-59, ж. 18-54)
Неполное среднее	3,8	3,8	1,3
Среднее общее	18,7	18,9	24,9
Начальное профессиональное	-	-	6,3
Среднее специальное	38,2	39,1	32,7
Незаконченное высшее	8,2	7,3	6,0
Высшее	31	30,8	25,7
Затрудняюсь ответить	0,1	0,1	3,1

Таблица 5. Непользователи интернета, % от опрошенных

Тип населенного пункта	Никогда не приходилось пользоваться интернетом		
	Результат опроса	Результат опроса с весами на возраст	Данные ФОМ (м. 18-59, ж. 18-54, пользователи моб.тел)
город с населением 1 млн и более	8,7	10,5	13,1
город с населением от 500 тыс. до 1 млн	16,2	20,0	18,9
город с населением от 250 до 500 тыс.	15,5	16,9	19,6
город с населением от 100 до 250 тыс.	18,1	21,3	19,2
город с населением от 50 до 100 тыс.	26,6	31,4	24,5
город с населением менее 50 тыс.	21,5	24,3	25,7
посёлок городского типа	33,3	37,5	28,5
село	39,5	42,6	39,5
нет данных	26,8	30,7	-

ных социально-демографических распределений: распределение выборки по Федеральным округам, типам населенного пункта изменится не более чем на 0,4 процентных пункта. Заметнее перевзвешивание сказалось на переменных, которые сильно коррелируют с возрастом. Например, суточная аудитория интернета уменьшилась на 3,2 процентных пункта; доля людей со средним специальным образованием увеличилась на 1 процентный пункт в основном за счет уменьшения доли людей с неоконченным высшим образованием (табл. 4).

Существенной характеристикой человека в наше время является частота пользования интернетом. Фонд "Общественное мнение" регулярно проводит замеры аудитории интернета, задавая вопрос: "Приходилось ли Вам лично пользоваться интернетом? Если да, то когда Вы лично последний раз пользовались интернетом?". В текущем опросе использована иная операционализация этого же концепта: "Пользуетесь ли вы интернетом? Если да, то каждый день, несколько раз в неделю или несколько раз в месяц и реже?". Сравнение распределений двух, столь разных вопросов, некорректно. Однако одно из закрытий, а именно процент людей, никогда не пользующихся интернетом, должно быть схожим в обоих опросах. Летом 2012 года у ФОМа (с фильтром по возрасту и наличию мобильного телефона) 24,8% опрошенных никогда не пользуются интернетом, в текущем опросе — 21,8%, с весами на возраст, для коррекции возможных смещений — 24,8%, то есть полное совпадение. Весьма близки результаты двух опросов и при разбивке выборок по типам населенных пунктов (табл. 5). Недельная аудитория: ФОМ — 67%, текущий опрос — 68,7%, весами на возраст — 65,5%. Суточная аудитория: ФОМ — 55,1%, текущий опрос — 49,9%, с весами на возраст — 46,7%.

Данные по пользователям интернета, независимой социально-демографической переменной, практически, совпадают в двух типах опроса. Этот показатель требует дальнейшего наблюдения и в случае обнаружения его устойчивости к разным типам интервьюирования, он может служить критерием оценки внешней валидности исследования, что чрезвычайно важно в условиях весьма ненадежных данных, предоставляемых Росстатом.

География опроса

Особенностью телефонного опроса является тот факт, что при его проектировании не обязательно строить кластерную выборку, т.е. группировать респондентов по кластерам (населенным пунктам, районам). Благодаря этому выборка получается "распыленной" по всей территории России, что позволяет включить в неё недостижимое для обычного опроса по месту жительства количество субъек-

тов РФ и населенных пунктов. Помимо этого, в выборку без дополнительных затрат попадают труднодоступные районы крайнего Севера, удаленные районы и районы с неспокойной криминогенной обстановкой (например, некоторые районы Северного Кавказа), которые обычно исключаются из опроса по месту жительства.

В результате проведения исследования были опрошены респонденты из 82 субъектов РФ (всех субъектов РФ, кроме Ненецкого автономного округа). В выборку при общем объеме 2000 респондентов попало 386 городов, 44 поселка городского типа и 358 сельских населенных пунктов

(у 159 респондентов не удалось идентифицировать населенный пункт). Стоит отметить, что название населенного пункта интервьюер записывал со слов респондента, поэтому не во всех случаях удалось его распознать. Населенный пункт восстанавливался по двум вопросам: "в каком регионе вы сейчас живете?" и "вы живете в региональном центре, другом городе или в сельской местности?" Если респондент говорил, что живет в региональном центре, то по региону восстанавливался город, если респондент говорил, что живет в другом городе, интервьюер записывал название города со слов респондента. Однако как показало прослушивание аудиозаписей интервью, значительная часть респондентов не различают понятия региональный и районный центры, что может вносить неконтролируемые смещения в данные по этим вопросам.

Интересно сравнить распределение выборки по типам населенных пунктов с официальными данными Росстата. К сожалению, Росстат не публикует данные о численности населения в разбивке по возрасту и типу населенного пункта, поэтому сравнить распределение опросных данных можно лишь с официальным распределением всего населения России по типам населенных пунктов (напомним, что опрашивались мужчины в возрасте 18-59 лет и женщины 18-54 года) — табл. 6.

В данном сравнении выявились систематические отклонения — выборка смещена в сторону более крупных населенных пунктов. Такое смещение можно объяснить, во-первых, существующей миграцией взрослого населения в крупные населенные пункты из сел и маленьких городов, что не отображается в распределении всего населения по типам населенных пунктов. А во-вторых, более высоким проникновением мобильной связи в крупных городах. Также стоит отметить, что у ряда демографов вызывает сомнение качество проведения переписи населения в России, на основе которой рассчитываются офи-

Таблица 6. Сопоставление распределения по типу населенного пункта, полученного в ходе опроса, с данными Росстата

Тип населенного пункта	опрос 18-54(59), %	росстат 0+, %	расхождение, %
более 1 млн	20	20	-0,2
от 500 тыс до 1 млн	15	11	3,8
от 250 до 500 тыс	12	9	3,0
от 100 до 250 тыс	8	10	-1,6
от 50 до 100 тыс	7	8	-0,6
менее 50 тыс	9	12	-2,9
пгт	2	5	-3,5
село	19	26	-7,0
нет данных	9	0	8,9

циальные данные Росстата.

Помимо сравнения данных опроса с данными Росстата по типам населенных пунктов можно сопоставить распределение по субъектам РФ. В большинстве регионов распределение выборки по субъектам РФ и распределение населения от 18 лет и старше по регионам по данным Росстата слабо отличаются друг от друга. Заметно (более, чем на 10 респондентов) выборка занижена в Москве, Санкт-Петербурге и Ростовской области, заметно завышена выборка в Тюменской, Оренбургской областях и Чечне. Отчасти такое расхождение можно объяснить различием в уровне отказов, высоком для первой группы (67%) и низком для второй группы (53%) при среднем уровне отказов в 62%.

Достижимость

Уровень достижимости в данном опросе принципиально зависит от колл-центра, назовем их условно Воронеж и Новосибирск по городам, в которых они расположены. Данная зависимость объясняется, во-первых, тем, что в Воронеже дозвонились интервьюеры, а в Новосибирске дозвонились автоматически и переключались на интервьюера, в случае, если респондент брал трубку. Во-вторых, программа дозвона Новосибирска не регистрировала несуществующие номера. Таким образом, при расчете коэффициентов достижимости Новосибирска учитывались все телефонные номера, на которые были проведены дозвон, а для Воронежа — только существующие телефонные номера. Фактически, работа обоих колл-центров разбита на две части: актуализация базы данных случайно сгенерированных номеров и звонок, направленный на разговор. Если идентификация несуществующих номеров проводится предельно корректно и количество попавших в этот список временно недоступных номеров несущественно, подобное разделение существенно улучшает не только коэффициенты достижимости, но и производительность работы колл-центра. Проблем несопоставимости данных вызвана разным программным обеспечением, задействованным для организации автоматической обработки данных. Чтобы не происходило подобных недоразумений рекомендуется работать с одинаковыми аппаратными ресурсами, что позволит своевременно идентифицировать коммуникативные проблемы, возникающие в том или ином колл-центре.

В соответствие со стандартами качества опросов, разработанных Американской ассоциацией исследователей общественного мнения [Standard definitions, 2011; Стандартные определения, 2005], рассчитаны четыре коэффициента, отражающих результативность выборки. Коэффициент ответов (RR3), или отношение полностью взятых интервью ко всем номерам, включенным в выборку, составил 13%. Коэффициент кооперации (COOP1), или доля опрошенных, ко всем единицам наблюдения, удовлетворяющим условиям отбора, с которыми удалось связаться по телефону, составляет 24%. Коэффициент отказов (REF3), или отношение отказавшихся от участия в опросе людей от всех потенциально удовлетворяющих условиям выборки единицам наблюдения, равен 40%. На-

Таблица 7. Коэффициенты достижимости в зависимости от Федерального округа и колл-центра

Колл-центр	Федеральный округ	RR3	COOP1	REF3	CON2
Воронеж	ПФО	0,19	0,25	0,54	0,77
	СЗФО	0,16	0,23	0,51	0,70
	СКФО	0,18	0,24	0,51	0,76
	СФО	0,22	0,31	0,47	0,72
	УФО	0,18	0,27	0,45	0,65
	ЦФО	0,17	0,23	0,54	0,73
	ЮФО	0,17	0,23	0,53	0,74
	Все	0,18	0,24	0,52	0,73
Новосибирск	ПФО	0,05	0,23	0,18	0,23
	СФО	0,08	0,26	0,23	0,31
	УФО	0,07	0,27	0,19	0,26
	ДВФО	0,06	0,26	0,18	0,25
	Все	0,07	0,26	0,19	0,26
Общий результат		0,13	0,25	0,39	0,54

конец, коэффициент контактов (CON2), или доля телефонных номеров, по которым удалось дозвониться ко всем работающим номерам, включенным в выборку, составил 55% (табл. 7).

Коэффициент COOP1, в расчет которого не входит количество несуществующих номеров, у обоих колл-центров находится на одном уровне. Остальные коэффициенты (RR3, REF3, CON2) значительно различаются между колл-центрами, но у каждого колл-центра достаточно устойчивы по Федеральным округам. Для опросов по мобильным телефонам типична ситуация большой вариативности коэффициентов ответов и отказов. Это происходит из-за большей вариативности условий и внешних факторов, сопутствующих соединениям по мобильным телефонам, нежели стационарным. Так, проводя сравнительное исследование по одному опросному инструменту в США, Словении и Финляндии, Марио Каллегаро с коллегами зафиксировали коэффициент ответов в США 10%, в Словении — 27%, а в Финляндии — 79% [Callegaro et al., 2007, p. 666]. Столь сильные различия отчасти объясняются разным выборочным дизайном и техникой проведения опроса: в США и Словении была реализована случайная генерация номеров, а в Финлян-

Таблица 8. Характеристики работы колл-центров в разбивке по Федеральным округам

Колл-центр	Федеральный округ	Доля номеров, на которые дозвонились	Доля непрошедших по фильтру (от согласных на опрос)
Воронеж	ПФО	25%	25%
Новосибирск		13%	34%
Воронеж	СЗФО	22%	29%
Воронеж	СКФО	22%	18%
Воронеж		28%	29%
Новосибирск	СФО	17%	31%
Воронеж	УФО	24%	28%
Новосибирск		16%	35%
Воронеж	ЦФО	23%	22%
Воронеж	ЮФО	24%	26%
Новосибирск	ДВФО	14%	36%
Воронеж	Все	24%	24%
Новосибирск	Все	15%	33%
Общий результат		20%	20%

дии опрос проводился по базе данных, доступной через интернет. Кроме того, в Финляндии всем попавшим в выборку абонентам были предварительно разосланы письма. Различия между США и Словенией можно объяснить тем, что в США были задействованы все операторы мобильной связи, а в Словении выбран лишь один, покрытие сети которого составило 80% абонентов. Авторы не проделали специальный детальный анализ расхождений, основанный на экспериментальных планах, однако уже приведенных объяснений достаточно для того, чтобы не браковать массивы с сильно различающимися коэффициентами достижимости, а пытаться сконструировать релевантные интерпретации зафиксированных расхождений.

Стоит указать на характеристики работы колл-центров, которые не описываются стандартными определениями качества телефонных опросов, а именно: доля номеров, на которые дозвонились, от всех номеров, на которые были проведены дозвон, и доля респондентов, не входящих в целевую группу, от всех согласившихся принять участие в опросе. Эти характеристики не зависят от фиксации несуществующих номеров, следовательно, при прочих равных условиях, должны быть инвариантны относительно колл-центра. Однако во всех Федеральных округах, где работали оба колл-центра (Приволжский, Сибирский и Уральский — выборка в них разбивалась приблизительно поровну), значения этих параметров получились зависимыми от колл-центра (табл. 8).

Зависимость доли номеров, на которые дозвонились, от колл-центра можно объяснить различиями в системах дозвона (силами интервьюеров или автоматически). У Новосибирска с автоматическим дозвоном эта доля стабильно меньше, чем у Воронежа. Возможно, система дозвона запрограммирована на слишком непродолжительный период ожидания ответа, и часть респондентов просто не успевают взять трубку. Доля же непрошедших по фильтру зависит от характеристик исследуемой совокупности и не может зависеть от характеристик интервьюера, колл-центра, специфик набора номера. Однако эта доля значительно отличается у двух колл-центров. На данный

Таблица 9. Распределение выборки по регионам

Регион	Респондентов	Регион	Респондентов
Республика Адыгея	6	Костромская область	14
Республика Алтай	2	Курганская область	11
Республика Башкортостан	50	Курская область	21
Республика Бурятия	12	Ленинградская область	29
Республика Дагестан	37	Липецкая область	17
Республика Ингушетия	2	Магаданская область	1
Кабардино-Балкарская Республика	7	Московская область	108
Республика Калмыкия	6	Мурманская область	8
Карачаево-Черкесская Республика	7	Нижегородская область	46
Республика Карелия	11	Новгородская область	12
Республика Коми	9	Новосибирская область	38
Республика Марий Эл	7	Омская область	21
Республика Мордовия	20	Оренбургская область	41
Республика Саха (Якутия)	7	Орловская область	8
Республика Северная Осетия-Алания	8	Пензенская область	12
Республика Татарстан	50	Псковская область	10
Республика Тыва	5	Ростовская область	50
Удмуртская Республика	20	Рязанская область	20
Республика Хакасия	16	Самарская область	37
Чеченская Республика	30	Саратовская область	45
Чувашская Республика	16	Сахалинская область	6
Алтайский край	35	Свердловская область	51
Краснодарский край	78	Смоленская область	12
Красноярский край	43	Тамбовская область	7
Приморский край	26	Тверская область	19
Ставропольский край	35	Томская область	14
Хабаровский край	23	Тульская область	19
Амурская область	16	Тюменская область	32
Архангельская область	20	Ульяновская область	16
Астраханская область	18	Челябинская область	50
Белгородская область	26	Ярославская область	22
Брянская область	18	г. Москва	149
Владимирская область	18	г. Санкт-Петербург	61
Волгоградская область	32	Еврейская автономная область	2
Вологодская область	25	Ненецкий автономный округ	0
Воронежская область	31	Ханты-Мансийский АО - Югра	21
Ивановская область	14	Чукотский автономный округ	1
Иркутская область	27	Ямало-Ненецкий автономный округ	6
Калининградская область	11	Пермский край	38
Калужская область	9	Камчатский край	7
Кемеровская область	36	Забайкальский край	24
Кировская область	20	Нет данных	7

момент у нас нет гипотез для интерпретации этого различия.

Подробные сведения по реализации выборки (таблицы)

Обычно в публикациях присутствуют лишь агрегированные данные, скорректированные под интерпретации и объяснительные схемы исследователя. В лучшем случае эмпирический материал сопутствует теоретическим выкладкам, с детальным описанием выпадающих, необъясненных данных, определяемых в качестве казуса. В худшем — последние исключаются из рассмотрения и читателям представляется весьма рафинированное знание. Для разрушения установки на легкое чте-

ние и приятное "скольжение" по тексту, мы предлагаем подробные таблицы, представляющие особенности проведенного исследования и задающие стандарт раскрытия информации. В таблице 9 представлено распределение выборки по всем субъектам РФ.

Полученные расхождения с данными Росстата не указывают на низкое качество опроса. Во-первых, нет задачи репрезентации столь дробных территориальных образований и для последнего потребовалась бы выборка на порядок превышающая по объему реализованную; во-вторых, данные Росстата представлены в ином диапазоне и включают население моложе 18 лет и старше 55 лет для женщин и 60 лет для мужчин; в-третьих, данные полученные в ходе переписи, сами имеют весьма сомнительную достоверность и надежность. Потребность в таких данных — это мониторинг и контроль за распределением выборки в региональном разрезе и формулирование гипотез, объясняющих территориальные сдвиги. Оставаясь противниками любого немотивированного редактирования данных, мы призываем исследователей раскрывать всю возможную информацию о всех интервенциях и преобразованиях, осуществленных с массивом. Данные о полученных расхождениях по значимым переменным — весьма ценный источник для понимания потенциальных и реальных систематических смещений.

Заключение

Опросы по мобильным телефонам входят в обыденные, рутинные практики исследовательских компаний. Чрезвычайно важно, чтобы за рутинизацией опросных процедур не была потеряна методическая точность и надежность выборочного инструментария. Иначе, даже располагая всеми необходимыми техническими средствами, исследователи в очередной раз получат весьма сомнительный инструмент конструирования общественного мнения. Зарубежными коллегами накоплен значительный опыт по методическому сопровождению подобных опросов. Отталкиваясь от разработок ведущих западных исследовательских центров, частично воспроизводя, частично дорабатывая и видоизменяя экспериментальные планы, можно сформулировать несколько рекомендаций, следование которых способствует поддержанию качества проводимых опросов:

1. В России сформировалась весьма благоприятная ситуация для развития опросов по мобильной связи. Во-первых, нет тарификации входящих звонков. Во-вторых, несмотря на отсутствие привязки номерных знаков к определенной территории, операторы мобильной связи придерживаются определенных правил в распределении диапазонов по географическому признаку. Последнее позволяет строить случайные телефонные выборки, не прибегая к дополнительным процедурам актуализации номеров (например, дополнительная генерация номеров по дозвонившемуся и подходящему по условиям отбора номеру), попадающих в требуемое географическое пространство. В-третьих, по уровню мобильной телефонизации Россия не отстает от мировых тенденций и уже сейчас можно утверждать о высоком охвате населения страны этим способом связи.

2. Минимальным набором методических характеристик опроса, без которого не должно рассматриваться ни одно исследование, должны быть статистические показатели достижимости, детальное описание процедур формирования и ремонта выборочной совокупности. Кроме того, следует приводить все аргументы, которыми руководствовались аналитики, принимая то или иное методическое решение.

3. Необходимо обеспечивать максимально возможное раскрытие методической информации о проведенных опросах. На этапе становления методологии, наблюдается большая вариативность основных показателей, объяснить которую невозможно без детального анализа всех условий проведения оп-

роса. Поэтому к стандартным коэффициентам достижимости и описаниям выборки, следует добавлять табличные распределения географического покрытия опроса, особенностей работы интервьюеров и влияние любых методических интервенций (ремонт выборки) на основные зависимые и независимые переменные исследования.

4. Процедура достижения по телефону потенциального респондента распадается на три этапа: во-первых, актуализация сформированной базы данных и отсеив неработающих и незарегистрированных телефонных номеров; во-вторых, проведение скрининга и отбор отвечающим условиям абонентов; в-третьих, проведение интервью. Каждый этап характеризуется особенностями соединения и требованиями к безопасности и комфортности возможных собеседников.

Методические исследования не могут рассматриваться в качестве не подлежащих раскрытию, наносящих угрозу конфиденциальности заказчика, участника опроса или представляющие интеллектуальную собственность разработчика. Придерживаясь этических норм и текущего законодательства, следует избегать любой формы закрытия методической информации. Развивая местечковое знание, исследователь загоняет в тупик не только собственную компанию, но и существенно снижает эффективность работы коллег. Методологическое знание принадлежит к тому типу интеллектуальной собственности, отчуждение которой лишь умножает ее стоимость для всех сторон сделки дарения.

Литература

1. Blumberg, S.J., Luke, J.V. Wireless substitution: Preliminary data from the January-June 2006 National Health Interview Survey. [Online] <<http://www.cdc.gov/nchs/products/pubs/pubd/hestats/wireless2006/wireless2006.htm>>
2. Callegaro, M., Steeh, Ch., Buskirk, T.D., et al. Fitting disposition codes to mobile phone surveys: Experiences from studies in Finland, Slovenia and the USA // *Journal of the Royal Statistical Society. Series A: Statistics in society*. 2007. Vol. 170. P. 647-670.
3. Frankel, L.R. The report of the CASRO task force on response rates // *Improving data quality in a sample survey: Proceeding Marketing Science Institute workshop*. 1982. 8th November / Ed. by F. Wiseman. Cambridge: Marketing Science Institute, 1983. P. 1-11.
4. Kviz, F.J. Toward a standard definition of response rate // *Public Opinion Quarterly*. Vol. 41. P. 265-267.
5. Lane, N.D., Miluzzo, E., Lu, H., et al. A survey of mobile phone sensing // *IEEE Communications Magazine*. 2010. Vol. 48. No. 9. P. 140-150.
6. Standard definitions: Final dispositions of case codes and outcome rates for surveys. 7th ed. / The American Association for Public Opinion Research. Deerfield, IL: AAPOR, 2011.
7. Standards for minimal disclosure // *AAPOR code of professional ethics and Practice*. Part. III. Deerfield, IL: AAPOR, 2010. P. 3-5.
8. Tucker, C., Brick, M.J., Meekins, B.J., Morgenstein, D. Household telephone service and usage patterns in the US in 2004 // *Proceeding survey research methods section American Statistical Association*. 2005. P. 4528-4534.
9. Vehavar, V., Berzelak, N., Manfreda, K.L. Mobile phones in an environment of competing survey modes: Applying metric for evaluation of costs and errors // *Social Science Computer Review*. 2010. Vol. 28. No. 3. P. 303-318.
10. Zuwallack, R. Piloting data collection via cell phones: Results, experiences, and lessons learned // *Field Methods*. 2009. Vol. 21. No. 4. P. 388-406.
11. Османов Т.Э. RDD выборка номеров мобильных и стационарных телефонов. Опыт использования в России. Выступление на конференции "Кризис в методах опроса и пути его преодоления", Киев. 2011. 28-29 октября.
12. Сапонов Д.И. Мобильные опросы в социологических исследованиях // *Мониторинг общественного мнения*. 2011. № 4 (104). С. 64-73
13. Сапонов Д.И. Мобильные опросы. Влияние способа сбора данных на результат // *Мониторинг общественного мнения*. 2012. №1 (107). С. 45-51
14. Стандартные определения: систематическое описание диспозиционных кодов и коэффициентов результативности для массовых опросов. 3-изд. 2004 год / Американская ассоциация исследователей общественного мнения; Пер. с англ. Д.М. Рогозина, Е.М. Киселева // *Социологический журнал*. 2005. № 2. С. 78-119.