

DOI: [10.14515/monitoring.2021.4.1855](https://doi.org/10.14515/monitoring.2021.4.1855)



Е. В. Князева

RESPONSE RATE — ПОКАЗАТЕЛЬ КАЧЕСТВА ОПРОСА ИЛИ МЕТОДИЧЕСКАЯ ПРОБЛЕМА? КАК МЫ СЧИТАЛИ ДОСТИЖИМОСТЬ РЕСПОНДЕНТОВ В КОМБИНИРОВАННОМ ИССЛЕДОВАНИИ

Правильная ссылка на статью:

Князева Е. В. Response rate — показатель качества опроса или методическая проблема? Как мы считали достижимость респондентов в комбинированном исследовании // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 4. С. 4—22. <https://doi.org/10.14515/monitoring.2021.4.1855>.

For citation:

Knyazeva E. V. (2021) Response Rate — An Indicator of the Research Quality or a Methodological Problem? How We Estimated the Response Rate in the Mixed Survey. *Monitoring of Public Opinion: Economic and Social Changes*. No. 4. P. 4–22. <https://doi.org/10.14515/monitoring.2021.4.1855>. (In Russ.)

RESPONSE RATE — ПОКАЗАТЕЛЬ КАЧЕСТВА ОПРОСА ИЛИ МЕТОДИЧЕСКАЯ ПРОБЛЕМА? КАК МЫ СЧИТАЛИ ДОСТИЖИМОСТЬ РЕСПОНДЕНТОВ В КОМБИНИРОВАННОМ ИССЛЕДОВАНИИ

КНЯЗЕВА Елена Владимировна — кандидат социологических наук, руководитель, ведущий научный сотрудник, Социологический информационно-исследовательский центр «Пульт», Одесса, Украина

*E-MAIL: knyazeva.elena.mail@gmail.com
https://orcid.org/0000-0001-5625-768X*

Аннотация. Подробная и точная информация об отказах от участия в социологических опросах, а также уровень достижимости респондентов считаются ключом к научным исследованиям и имеют особое значение для сбора высококачественных данных. В статье представлены результаты эксперимента по расчету уровня достижимости респондентов (response rate) на материалах актуального коллаборативного опросного исследования «Открытое мнение — Санкт-Петербург. Образ идеального губернатора», проведенного в течение мая-июля 2019 г. Основное внимание в статье сосредоточено на методических и технических трудностях сбора, кодирования и интерпретации параданных в комбинированном (mixed) опросе (с раздельными подвыборками телефонного и уличного опросов). Для оценки результативности опросов, а также для анализа отказов использована стандартная методика Американской ассоциации изучения общественного мнения (AAPOR). Описаны основные проблемы, возникающие при попытке рассчитать уровень достижимости респондентов и оценить качество па-

RESPONSE RATE — AN INDICATOR OF THE RESEARCH QUALITY OR A METHODOLOGICAL PROBLEM? HOW WE ESTIMATED THE RESPONSE RATE IN THE MIXED SURVEY

*Elena V. KNYAZEVA¹ — Cand. Sci. (Soc.), Head of the Center, Leading Researcher
E-MAIL: knyazeva.elena.mail@gmail.com
https://orcid.org/0000-0001-5625-768X*

¹ Sociological Information and Research Center «Pulse», Odessa, Ukraine

Abstract. Detailed and accurate information on the refusal rate in sociological surveys, as well as the response rate, are considered the key to scientific research and are of particular importance for collecting high-quality data. The article presents the results of an experiment to calculate the response rate based on the materials of the current collaborative survey «Open Opinion — St. Petersburg. The Image of the Ideal Governor», conducted in May—July 2019. The article focuses on the methodological and technical difficulties of collecting, coding and interpreting paradata in a mixed survey (with separate subsamples of telephone (CATI) and street (face-to-face) surveys). The standard methodology of the American Association for Public Opinion Research (AAPOR) was used to assess the effectiveness of the surveys, as well as to analyze their failures. The author describes the main problems that arise during the assessment of the response rate and paradata quality in a mixed survey with disposition codes recommended by the AAPOR. The study shows that paradata registered during the survey contain many omissions and inaccuracies, as well as errors in coding refusals,

радных в комбинированном опросе с применением диспозиционных кодов, рекомендованных Американской ассоциацией изучения общественного мнения, а именно: зарегистрированные в процессе опроса параданные содержат много пропусков и неточностей, а также ошибки кодирования ответов, которые приводят к неверным расчетам уровня достижимости и оценкам результативности исследования в целом. Основная проблема — качество параданных. Информация, представленная в статье, в первую очередь предназначена для того, чтобы привлечь внимание исследователей к данной проблеме и ее последствиям. Предложены ее возможные решения, которые варьируются от повышения квалификации интервьюеров, ориентированных на сбор сопутствующих исследованию данных и внедрение автоматической системы учета параданных, до разработки стандартизированной, хорошо документированной и воспроизводимой процедуры сбора параданных, адаптированной к разным методам сбора данных и российскому контексту.

Ключевые слова: открытое исследование, комбинированный опрос, параданные, уровень достижимости респондентов, AAPOR

Благодарность. Автор выражает благодарность Содружеству «Открытое мнение» за открытые массивы данных и материалы исследования «Открытое мнение — Санкт-Петербург. Образ идеального губернатора», которые позволили заглянуть в глаза «объективности» и представить результаты проделанной работы, акцентируя внимание на трудностях, с которыми при-

which lead to bias in estimations of the response rate and effectiveness of the survey. The main problem is the paradata quality. The information presented in the article is primarily intended to draw the attention of researchers to this problem and its consequences. The author proposes possible solutions, which range from improving the qualifications of interviewers, developing the process of collecting the data accompanying the survey, and introducing an automatic system for recording paradata, to developing a standardized, well-documented and reproducible procedure for collecting paradata, adapted to different methods of data collection and Russian context.

Keywords: open research, mixed methods, paradata, response rate, AAPOR

Acknowledgments. The author expresses gratitude to the Open Opinion Project for the open data sets and materials of the «Open Opinion — St. Petersburg. The Image of the Ideal Governor» study, which allowed to examine the «objectivity» and present the results of the study, focusing on the difficulties that the researchers had to face together with unresolved methodological problems. Also, many

шлось столкнуться исследователям, и нерешенных методических проблемах. Также огромная благодарность коллегам Марии Мацкевич, Игорю Задорину, Владимиру Звонковскому, Роману Кузнецову, Юлии Баскаковой, Надежде Корытниковой за советы и конструктивную критику в процессе работы над материалами исследования.

thanks to my colleagues Maria Matkevich, Igor Zadorin, Vladimir Zvonovsky, Roman Kuznetsov, Yulia Baskakova, and Nadezhda Korytnikova for the advice and constructive criticism they provided in the process of working on research materials.

Введение

В научном и публичном дискурсе продолжает набирать актуальность тема отказов от участия в опросах общественного мнения. Повсеместно исследователи транслируют мнение, что в последние годы достижимость респондентов (response rate) снижается (см., например, [Романович, 1996; Рогозин, 2004, 2008; Юдин, 2008; Османов, Рогозин, 2013; Корытникова, 2013; Паниотто, Харченко, 2012]). В связи с этим ученые все чаще выражают озабоченность по поводу качества данных и занимаются поиском стратегий их сбора, которые позволили бы преодолеть данную тенденцию. Считается, что высокий показатель уровня достижимости респондентов обеспечивает большую точность результатов социологического исследования, из-за чего его величина часто используется для оценки качества данных в целом. Однако в действительности высокий уровень отказов от участия в опросе оказывается проблемой лишь тогда, когда группы респондентов и «неответивших» отличаются друг от друга по одному или нескольким параметрам.

Следует отметить, что социологи начали указывать на снижение величины откликов уже более четверти века назад. Начиная с конца 70-х годов прошлого века осуществляются попытки введения общеотраслевых стандартов, поиск смыслов, интерпретаций и методов расчета показателя достижимости респондентов в социологических опросах. При поддержке CASRO американские ученые Ф. Вайзман и П. Макдональд организовали исследование, в котором сорок директоров проектов в фирмах CASRO получили параданные трех опросов, для каждого из которых их попросили рассчитать процент откликов (response rate). В результате сорок специалистов выдали двадцать девять разных значений этого показателя, который варьировался от 12% до 90%, а наиболее часто упоминаемое значение приводилось только три раза [Wiseman, 2003]. После этого Совет директоров CASRO¹ сформировал специальную рабочую группу с целью выработки стандартизированного определения показателя достижимости (response rate), которая рекомендовала следующую формулу его расчета: response rate (коэффициент ответов)² = количество законченных интервью / количество всех единиц

¹ CASRO — Council of American Survey Research Organizations (Совет американских исследовательских организаций), основан в 1975 г. Это ассоциация организаций, занимающихся исследованиями в области опросов. В Совет входят более 275 компаний и исследовательских центров в США и за рубежом. Веб-сайт: <http://www.casro.org>.

² Response rate = Number of completed interviews with reporting units / Number of eligible reporting units in the sample. См.: CASRO on the Definition of Response Rates // CASRO. URL: <https://www.insightsassociation.org/issues-policies/best-practice/casro-definition-response-rates> (дата обращения: 29.07.2021).

наблюдения, удовлетворяющих критериям выборки. Однако проведя несколько экспериментов по расчету показателей достижимости в разных опросных компаниях, авторы формулы пришли к выводу, что во многих случаях невозможно ее обосновать и определить правильность выбранных для расчетов отчетных единиц.

В конце 1990-х годов в разработку стандартов расчета показателя достижимости включилась Американская ассоциация изучения общественного мнения (AAPOR). В 2016 г. она выпустила рекомендации³, в которых указывалось, как следует определять и рассчитывать этот показатель в разных типах опросов. AAPOR предложила «Стандартные определения», которые включают конечные коды, сгруппированные в четыре основные категории:

- 1) Лица, соответствующие критериям отбора и опрошенные (интервью у них взяты);
- 2) Лица, соответствующие критериям отбора, но не были опрошены;
- 3) Лица с неизвестным статусом, не опрошенные (контакт не состоялся);
- 4) Лица, не соответствующие критериям отбора, не опрошенные.

Помимо этого, AAPOR предлагает шесть альтернативных формул расчета коэффициента ответов с помощью калькулятора RR⁴ — электронной таблицы Excel, вычисляющей показатели результативности опроса после того, как исследователь вводит диспозиционные коды⁵. Разработчики надеялись, что с внедрением рекомендаций CASRO и AAPOR в профессиональном сообществе будет достигнута стандартизация отчетности по показателям результативности опросов, однако этого не случилось. На практике в большинстве исследований, сообщающих о коэффициенте достижимости, не указывается, как он был рассчитан. Это затрудняет или даже делает невозможным сравнение таких коэффициентов в разных типах опросов.

Таким образом, проблема интерпретации терминов и методов расчетов коэффициента по-прежнему существует. В качестве основных причин можно назвать использование разных определений и стратегий регистрации конечных кодов, применение более сложных методов сбора данных и регистрации параданных. Поэтому особенно важно, чтобы исследователи не только информировали об уровне достижимости, но и подробно описывали характеристики диспозиционных кодов, а также то, как был рассчитан показатель результативности исследования. В противном случае сообщаемые показатели не дают практически никакой полезной информации и могут приводить либо к завышению, либо к занижению уровня достижимости. Несмотря на то, что нормативный характер стандартов AAPOR представляет собой «передовой опыт», социологи-исследователи заинтересованы в их эмпирическом применении и выработке общих, понятных кодов, указывающих на результат интервью. Поэтому мы провели анализ отказов респондентов от участия в опросе на базе исследования «Открытое мнение — Санкт-Петербург.

³ Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys // AAPOR. URL: https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf (дата обращения: 20.07.2021).

⁴ Подробнее см.: Response Rates — An Overview: AAPOR Response Rate Calculator 4.1 // AAPOR. URL: <https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx> (дата обращения: 20.07.2021).

⁵ Количественные значения параметров исследования.

Образ идеального губернатора», задаваясь вопросами, как фиксируются отказы в социологических исследованиях и насколько стандарт AAPOR распространен среди исследовательского сообщества? С какими трудностям сталкиваются исследователи в процессе сбора и кодирования параданных? Каковы риски искажения качества результатов и к чему может привести неправильная интерпретация конечных кодов? Какие имеются ограничения при сборе и анализе параданных в комбинированном исследовании?

Эмпирическая база исследования

Статья подготовлена на основе результатов опроса «Открытое мнение — Санкт-Петербург. Образ идеального губернатора»⁶, проведенного с мая по июль 2019 г. содружеством профессиональных социологов «Открытое мнение»⁷ в Санкт-Петербурге. В проекте участвовали специалисты из разных исследовательских центров России и ближнего зарубежья, включая автора данной публикации. Основное внимание исследователей фокусировалось на представлениях петербуржцев о предпочтительных для них качествах градоначальника.

Полевые работы выполнялись четырьмя разными исследовательскими компаниями⁸. Опрос проводился методом формализованного интервью. Использовался комбинированный (mixed) тип выборки с отдельными подвыборками для телефонного (CATI)⁹ и уличного (face-to-face) опросов. Общий объем выборочной совокупности составил 1012 респондентов: 512 приняли участие в телефонном опросе (B1T1 и B2T2) и 500 — в уличном, где для 250 интервью использовались бумажные анкеты (B3S1) и для еще 250 — планшеты (B4S2). Телефонный и уличный опросы проводились по единой анкете. По замыслу разработчиков ее вопросы должны были одинаково восприниматься как при личном контакте, так и при телефонном интервью, хотя на практике имелись некоторые различия в восприятии, связанные с методом.

При подготовке данной статьи мы использовали данные, предоставленные рабочей группой проекта «Открытое мнение», — финальный массив с результатами опроса и массив «отказов». Мы сосредоточили внимание на трех исследовательских вопросах:

⁶ Подробнее об исследовании, его целях, методике, инструментарии и результатах см.: Открытое мнение: Санкт-Петербург: Результаты проекта // Открытое мнение. Независимый социологический проект. 2019. URL: <https://openopinion.ru/sankt-peterburg> (дата обращения: 20.07.2021).

⁷ Команда содружества «Открытое мнение» на протяжении ряда лет занимается исследованиями, направленными на улучшение существующих методов опроса, а также ищет альтернативные способы социологического измерения. Применяя передовые методы исследований, команда «Открытое мнение» старается быть максимально прозрачной в отношении качества полученных данных и того, как они собираются. В этом содружество следует возникшей в начале XXI века концепции открытой науки, фундаментальные принципы которой — открытые данные, открытая методология, открытая экспертная оценка. Проекты, подобные «Открытому мнению», стремятся внести вклад в изменение процедуры проведения социологических исследований, сделав открытость данных их нормой. Полное раскрытие информации об опросе «Образ идеального губернатора» — методологии, описания выборки и дизайна исследования, а также другой специфической информации можно рассматривать как вклад Содружества в развитие концепции открытой науки.

⁸ Из этических соображений компании-исполнители в массиве закодированы как B1T1, B2T2, B3S1 и B4S1. Для различения этих компаний в статье мы будем использовать маркеры: B1T1 и B2T2 — компании, проводившие телефонный опрос; B3S1 и B4S1 — уличный.

⁹ CATI — Computer Assisted Telephone Interview.

- 1) Насколько информативны собранные в процессе опроса параданные?
- 2) Различаются ли те, кто согласился участвовать в опросе, и те, кто отказался принять в нем участие, по своим социально-демографическим характеристикам?
- 3) Зависят ли уровень достижимости респондентов и качество сбора параданных от интервьюеров, участвующих в исследовании?

Насколько информативны исходные данные об отказах от участия в опросе?

Значительная по объему и важная часть работы по решению методологических и организационных проблем комбинированного исследования «Образ идеального губернатора» освещена в «Аналитическом отчете о результатах исследования»¹⁰. Вопросам расчетов важных показателей качества выборки, анализу достижимости респондентов, их отказов и неответов посвящен «Методический отчет о проведении исследования»¹¹. Он также содержит анализ систематических ошибок, связанных с отказами как на уровне всего опроса в целом, так и в случае конкретных методов сбора данных. Как правило, данные об отказах (параданные), фиксируемые в процессе исследования, зависят от его типа и должны быть четко изложены организаторами — указаны в описании методологии и отмечены в полевом этапе [Лебедев, 2020]. Учитывая, что сбор параданных в исследовании «Образ идеального губернатора» не ставился как необходимая задача, специальные контактные формы для регистрации отказов не разрабатывались, чтобы не увеличивать нагрузку на интервьюеров и стоимость опроса. Поэтому наше внимание было сосредоточено на анализе ключевых параметров исследования, указанных в методике формирования выборочной совокупности и в инструментарии уличного и телефонного опросов. Поскольку эти данные в силу определенных причин фиксировались различными способами, мы будем рассматривать их только как «минимальное раскрытие».

Так, в «Методике формирования выборочной совокупности респондентов уличного опроса»¹² было четко указано, кто должен быть исключен из исследования, а также какую информацию необходимо фиксировать интервьюеру, а именно:

- число людей, приглашенных к опросу, которые не являлись жителями Петербурга,
- число людей, для которых квоты были уже превышены (возраст, пол, образование),
- количество запланированных интервью,
- количество проведенных интервью,
- количество отказов,

¹⁰ Подробнее см.: «Образ идеального губернатора» (ОМ—Санкт-Петербург). Аналитический отчет о результатах исследования // Открытое мнение. Независимый социологический проект. 2019. URL: <https://openopinion.ru/uploads/s/f/k/O/fk0srzydghw/file/DZ30NmCk.pdf?preview=1> (дата обращения: 26.07.2021).

¹¹ Подробнее см.: Проект «Открытое мнение — Санкт-Петербург». Методический отчет о проведении исследования // Открытое мнение. Независимый социологический проект. 2019. URL: <https://openopinion.ru/uploads/s/f/k/O/fk0srzydghw/file/D8eMTaw7.pdf?preview=1> (дата обращения: 26.07.2021).

¹² Подробнее см.: Открытое мнение: Санкт-Петербург. Инструментарий: Методика (уличный опрос) // Открытое мнение. Независимый социологический проект. 2019. URL: <https://openopinion.ru/sankt-peterburg> (дата обращения: 26.07.2021).

- количество случаев, когда респондент не подходил по скринингу (вопрос S0),
- количество случаев, когда респондент не подходил по квотам (вопрос S2 и Q32),
- количество прерванных интервью,
- характеристики интервьюеров (пол, возраст, образование, в каком опросе участвует).

Помимо этого, в процессе опроса интервьюеры/операторы должны были фиксировать параметры, перечисленные в разделе «Параданные»¹³ анкеты для уличного/телефонного опроса¹⁴, такие как:

- дата опроса,
- время опроса (B1 — начало интервью, B2 — конец интервью),
- название колл-центра, который проводил опрос (для телефонного опроса) / бригада (для уличного опроса),
- интервьюер (фамилия, имя или номер), который проводил опрос,
- точка опроса (для уличного опроса),
- характеристики набора телефонного номера: ошибка (нет гудка); неверный номер; номер не обслуживается и т. п.; нет ответа; вне зоны действия сети; занято; взяли трубку (для телефонного опроса).

Первое затруднение, с которым мы столкнулись в процессе работы, — значимые различия в параданных комбинированного опроса. Данные об отказах, предоставленные колл-центрами и организациями, проводившими уличные интервью, оказались несопоставимы, поэтому не имело смысла анализировать их на уровне исследования в целом. Некоторые типы параданных отсутствовали в массивах отказов по результатам уличных опросов либо фиксировались разными способами в случае телефонных интервью. Все это приводило к ошибкам измерения (см. табл. 1).

Таблица 1. Сведения о предоставленных параданных (массив «Отказы»)

База «Отказы»	B1T1	B2T2	B3S1	B4S2
Интервьюер/оператор	+	+	+ точка опроса	+ точка опроса
Дата опроса	+	+	+	+
Время начала и окончания разговора	+	+	-	-
Длительность интервью	+	+	-	-
Имя отказавшегося	+	-	-	-

¹³ Речь идет об отказах от участия в опросе. Отказы — это информация, которая традиционно собирается в рамках опроса, при этом никакой дополнительной формы для фиксации не требуется, вопросы об отказах встроены в опросник.

¹⁴ Подробнее см.: Открытое мнение: Санкт-Петербург. Инструментарий: Анкета опроса // Открытое мнение. Независимый социологический проект. 2019. URL: <https://openopinion.ru/sankt-peterburg> (дата обращения: 26.07.2021).

База «Отказы»	B1T1	B2T2	B3S1	B4S2
Статус разговора	+	–	результат обращения	–
Телефон отказника	+	+	–	–
Возраст отказника	+	+	+	+
Пол отказника	+	+	+	+
Основное занятие	+	–	+	
Образование	+	–	–	+
Прописка	–	–	–	+
Район проживания	–	–	–	+

Отсутствие единого классификатора параданных для всех типов опроса сделало невозможным расчет коэффициентов результативности для массива целиком. Кроме того, в массиве отказов невозможно различить «недостигнутых респондентов» и «отказавшихся», чтобы оценить их влияние на возможные смещения. Попытка закодировать параданные так, как рекомендовано в методике AAPOR [Стандартные определения..., 2016], то есть с использованием диспозиционных кодов, не увенчалась успехом. Например, такой важный для телефонных опросов показатель, как «результаты дозвона», в массивах отказов отсутствовал или был представлен в другой модификации, как «статус разговора» (см. табл. 2).

Таблица 2. Категории отказов в массиве телефонного опроса B1T1

	Абсол. значение	%
Бросили трубку	382	10,5
Категорический отказ	3018	82,96
Неверный номер	3	0,08
Нет ответа	13	0,36
Получен ответ	150	4,12
Результативная анкета	70	1,92
Всего	3637	100

Время начала и окончания разговора фиксировалось только в телефонных опросах, но не в уличных, хотя если при проведении интервью с использованием бумажной анкеты это действительно затруднительно, то при наличии планшета такие данные должны отмечаться автоматически. Также было установлено, что на качество и полноту данных влияют интервьюеры/операторы: в массиве отказов

оказались как отчеты, в которых отсутствовала большая часть параданных, так и такие, где все требования методики их фиксации были соблюдены. К сожалению, первых было значительно больше. Из тринадцати заданных параметров только по четырем имелись полные сведения, а именно:

- ФИО интервьюера/оператора,
- дата опроса,
- пол респондента,
- возраст респондента.

Всего мы проанализировали четыре массива параданных — для двух телефонных опросов и двух уличных. Остановимся на каждом из них подробнее.

Анализ параданных в телефонных опросах

Не имея единой инструкции, колл-центры сами настраивали запись разговоров/контактов в процессе опроса. Для проведения телефонных интервью оба центра использовали случайную выборку стационарных и мобильных номеров. Такая выборка всегда содержит какую-то часть несуществующих телефонных номеров. Судя по массивам «отказов», в ходе телефонных опросов идентификация таких случаев не производилась. Следовательно, проблематично было отличить номера респондентов, не отвечающих критериям отбора, от подходящих.

Отдельная проблема связана с тем, как следовало бы интерпретировать различные сигналы и к какой категории их отнести. Еще одна трудность касалась того, как интерпретируется окончательный результат разговора, который связан с процедурой отбора, структурой выборки и характером обращения оператора к респондентам.

В телефонном опросе B2T2 все случаи, в которых никто не ответил после заранее определенного количества звонков, были закодированы как «неподходящие случаи» — таких оказалось 1 628. В телефонном опросе B1T1 были выделены такие категории, как «категорический отказ» (3 018), «бросили трубку» (382), «неверный номер» (3), «нет ответа» (13), «получен ответ» (150) и «результативная анкета» (70) — всего 3 637 случаев (см. табл. 3).

Таблица 3. **Количество зафиксированных отказов по партнерам**

Код из выборки	B1T1 (CAT1 1)	B2T2 (CAT1 2)	B3S1 (F2F 1)**	B4S2 (F2F 2)**
Завершенные интервью	261	251	250	250
Количество зафиксированных отказов	3 637	1 628	173	412
Коэффициент ответов (RR3)*	7 %	13 %	59 %	38 %

* Коэффициент ответов RR3 рассчитывается как отношение полностью взятых интервью ко всем телефонным номерам (ко всем респондентам), включенным в выборку, и показывает в ней долю опрошенных респондентов.

** F2F — face-to-face.

Отметим, что выборка была распределена поровну между двумя колл-центрами. Каждый получил 37 500 телефонных номеров, по которым необходимо было провести не менее 250 полных интервью (не все номера из выборки были задействованы). В результате исследовательский центр В1Т1 опросил по телефону 261 респондента, а центр В2Т2—251¹⁵. В первом случае процент ответивших составил 13%, во втором — 7% (см. табл. 3). Возникает вопрос, насколько достоверны такие расчеты и правомочно ли делать вывод о том, что уровень достижимости в опросе, проведенном В2Т2, почти вдвое выше, чем у В1Т1, если случаи, не отвечающие критериям отбора, могли быть просто исключены из расчета? Ведь именно это могло привести к ошибкам измерения и завышению показателя достижимости. Интересно также, в каких случаях в опросе В1Т1 контакту присваивался статус «категорический отказ». AAPOR рекомендует «категорическими» считать «отказы, в которых потенциальный респондент просит, чтобы его „исключили из списка“, прямо говорит „не звоните снова“ или угрожает судебным иском, если будут предприняты усилия по повторному контакту»¹⁶. В процессе консультаций с руководителем опросного центра В1Т1 выяснилось, что при формировании массива отказов неверно были объединены параметры «категорический отказ» и «больше не звонить». А вариант «больше не звонить» оператор/интервьюер мог отмечать по субъективным причинам — например, из-за проблем на линии, плохой слышимости, включения автоответчика и т. д.

Анализ параданных в уличных опросах

В случае с уличными опросами применение методики AAPOR в принципе невозможно, так как, во-первых, данная техника сбора данных не пользуется популярностью на Западе в целом и в США в частности. Во-вторых, применять рекомендации AAPOR, сформулированные для интервью face-to-face, к уличным опросам неправомерно, так как в этих случаях используются разные принципы построения выборки, а также образцы регистрации параданных, которые на практике наших опросов не применяются.

Отметим, что выборка для уличного опроса (500 человек) была распределена между восемнадцатью районами Санкт-Петербурга, при этом 250 анкет делалось на планшетах (B3S 1), и еще 250 — на бумажных носителях (B4S 1). Каждая исследовательская компания составила список точек для проведения уличного опроса. Работа была распределена так, что на каждую точку приходилось от десяти до пятнадцати респондентов, а на каждого интервьюера — не более двадцати опрошенных. Полевая активность интервьюеров визуально контролировалась, а после завершения работ был проведен выборочный контроль интервью¹⁷. Такой подход дает интервьюерам больше свободы действий при выборе респондентов,

¹⁵ «Образ идеального губернатора» (ОМ—Санкт-Петербург). Аналитический отчет о результатах исследования // Открытое мнение. Независимый социологический проект. 2019. С. 8. URL: <https://openopinion.ru/uploads/s/f/k/O/fkOsrzydghw/file/DZ30NmCk.pdf?preview=1> (дата обращения: 26.07.2021).

¹⁶ Исходная цитата: «Researchers also often distinguish between „hard“ and „soft“ refusals, though the definitions of these terms vary by organization and within organizations. Certainly, refusals where a potential respondent asks to be „taken off the list“, explicitly says „do not call me again“, or threatens legal action if efforts are made at re-contact should be considered hard refusals and not re-contacted». Источник: Survey Refusals. AAPOR Report: Current Knowledge and Considerations Regarding Survey Refusals. AAPOR Task Force on Survey Refusals, September 8, 2014. URL: <https://www.aapor.org/Education-Resources/Reports/Survey-Refusals.aspx> (дата обращения: 14.08.2021).

¹⁷ Там же. С. 9—10.

однако может привести к высокому, но предвзятому количеству ответов из-за чрезмерной представленности в выборке более склонных к сотрудничеству лиц. При равном числе опрошенных число отказов у компании B3S1, использовавшей планшеты, составило 412 случаев, а у компании B4S1, использовавшей бумажную анкету, отказов оказалось вдвое меньше — 173 (см. табл. 3).

Не имея возможности рассчитать все коэффициенты результативности в силу отсутствия необходимых исходных данных, мы ограничились вычислением коэффициента достижимости RR3¹⁸ для телефонных (без разбивки по стационарным и мобильным телефонам) и уличных опросов (см. табл. 3).

В целом показатели достижимости находятся в пределах значений, типичных для техники опроса, использованной той или иной компанией-исполнителем. Согласно расчетам, телефонные интервью показали самый низкий коэффициент достижимости — 7 % (B1T1) и 13 % (B2T2) соответственно. Самый высокий показатель — 59 % — продемонстрировал уличный опрос с использованием бумажной анкеты, и 38 % оказалось у планшетного опроса. Мы объясняем эти различия следующими причинами.

Во-первых, применение окончательных кодов AAPOR зависит от интерпретации отказов и качества данных. Ошибки кодирования могут возникать из-за отсутствия понимания того, как осуществлялись полевые работы. Неправильные спецификации могут напрямую влиять на расчеты показателя достижимости. Этим объясняется разный уровень отказов в телефонных опросах.

Во-вторых, низкий уровень достижимости в телефонных опросах может быть связан как с поведением респондентов, так и с техническими факторами, специфичными для данного метода опроса.

В-третьих, большее разнообразие результатов в опросах CATI по сравнению с уличными может объясняться тем, что уличный опрос начинается с отбора конкретного человека, а телефонный — со списка номеров, которые не содержат никакой личной информации. По данным в массиве невозможно отследить последовательность, конверсию, число и результативность попыток дозвона.

В-четвертых, если предположить, что около 50 % всех сгенерированных номеров не связаны с фактическим доступом к телефону (проверить эту гипотезу на данном конкретном примере не представляется возможным, но она имеет право на существование), то соответственно доля лиц с неизвестным статусом и не опрошенных и лиц, не соответствующих критериям отбора и не опрошенных, в телефонных интервью может быть больше, чем в уличных.

Наконец, высокий показатель достижимости в уличных опросах может объясняться тем, что случаи, не отвечающие критериям отбора, исключены из расчета доли ответивших. Однако меньшее количество ответов означает, что исследователи тратят больше ресурсов на то, чтобы связаться и опросить людей по телефону. В целом, чем больше взаимодействия между потенциальным респондентом и людьми, собирающими данные, тем выше уровень достижимости.

Таким образом, даже на основании анализа ограниченной информации наглядно видно, что на уровень достижимости и число отказов существенно влияют:

¹⁸ Интерпретацию коэффициента см. под таблицей 3.

метод сбора данных, качество параданных и ошибки измерения. Все эти проблемы указывают на то, что уровень достижимости не может считаться надежным показателем качества исследования. Во всяком случае, пока...

Характеристика основных демографических параметров «отказников»

Одной из методических задач нашего исследования было выявление характеристик людей, которые отказываются принимать участие в опросе, а также понимание того, насколько они отличаются от опрошенных.

Анализ показал, что возрастные параметры индивидов, отказавшихся от интервью, и респондентов, вошедших в выборочную совокупность, совпадают либо незначительно различаются (от 1 % до 7 %) во всех типах опроса (см. табл. 4).

Таблица 4. Основные демографические параметры «отказников»

Значения шкалы	В1Т1 (телефонный опрос 1)		В2Т2 (телефонный опрос 2)		В3S1 (уличный опрос 1)		В4S2 (уличный опрос 2)		Все опрошенные (N = 1 012)
	частота	%	частота	%	частота	%	частота	%	
<i>Возраст отказавшихся</i>									
18—29 лет	12	18	17	18	34	20	56	18	17
30—44 лет	18	27	30	32	54	31	74	24	31
45—59 лет	14	21	24	25	40	24	83	27	25
60 лет и старше	22	33	23	25	43	25	94	31	28
Всего	66	100	94	100	171	100	307	100	100
<i>Пол отказавшихся</i>									
Мужской	59	57	499	41	81	47	323	81	44
Женский	44	43	733	59	91	53	75	19	56
Всего	103	100	1232	100	172	100	398	100	100

На основании этих сравнений можно сделать вывод, который тем не менее требует дальнейших наблюдений и проверки, что по признаку «возраст» те, кто отказывается от участия в опросе, практически не отличаются от опрошенных респондентов. Возможно, этот параметр достаточно устойчив к разным методам сбора данных и может служить критерием оценки внешней валидности всего исследования в целом.

С переменной «пол» ситуация иная. Только в двух опросах — телефонном В2Т2 и уличном В3S1 (с использованием бумажных анкет) — распределения по полу у отказавшихся участвовать в исследовании и опрошенных респондентов отклоняются незначительно — в пределах 3 %. В телефонном опросе В1Т1 и уличном опросе В4S2 (с использованием планшетов) по признаку «пол» зафиксированы существенные различия: мужчины в несколько раз чаще отказывались от ин-

тервью (см. табл. 4). Несоответствие гендерных половозрастных характеристик «отказавшихся» соответствующей структуре выборочной совокупности, наблюдаемое при использовании разных техник опроса, может свидетельствовать как о предвзятости при отборе участников исследования, так и о значительном различии групп тех, кто согласился и не согласился в нем участвовать, а также о том, что методика исследования сама по себе влияет на достижимость респондентов.

Влияние интервьюеров на уровень отказов

Можно предположить, что, если не использовать стандартизированные инструменты сбора параданных, интервьюеры/операторы могут повлиять на их полноту и качество и стать источником систематических ошибок измерения в опросах.

Наш анализ показал, что у разных интервьюеров уровень отказов значительно колеблется. Возможно, это связано с их опытом или отношением к работе, однако проверить это предположение нам не представилось возможным.

Рассмотрим на примере телефонных опросов, влияют ли интервьюеры на уровень достижимости респондентов. В. Б. Звоновский и Ю. В. Соловьева утверждают, что «использование CATI при проведении телефонных опросов теоретически позволяет собирать значительный объем параданных, необходимый для расчета характеристик достижимости, отказов и кооперации <...>. Программы, управляющие системами CATI, традиционно собирают в автоматическом режиме такие данные, как длительность интервью, число звонков, число состоявшихся интервью у каждого оператора-интервьюера» [Звоновский, Соловьева, 2017: 60]. В анализируемых нами телефонных опросах параданные собирались автоматически, поэтому у нас была возможность изучить число отказов и длительность разговора у каждого оператора-интервьюера (см. табл. 5).

Таблица 5. Отказы у интервьюеров

Интервьюер	В1Т1				В2Т2			
	Число случаев (n)	Доля отказов, %	Среднее время разговора, сек	Станд. отклонение	Число случаев (n)	Доля отказов, %	Среднее время разговора, сек	Станд. отклонение
1	1048	98	18,92	25,632	54	83	40,74	29,536
2	1	**	38	***	9	47	56,44	17,501
3	69	86	24,43	19,967	5	46	30,8	22,421
4	1243	94	18,23	25,068	1	67	46	***
5	7	*	19,86	19,987	729	99	27,76	8,935
6	13	**	19,15	5,728	1	25	54	***
7	17	*	18,76	16,123	9	81	42,56	11,458
8	8	**	30,63	24,715	99	90	27,8	9,91
9	595	97	13,78	31,836	4	24	21,5	27
10	15	**	30,4	34,278	2	60	45,5	64,347

Интер- вьюер	В1Т1				В2Т2			
	Число случаев (n)	Доля отказов, %	Среднее время разгово- ра, сек	Станд. отклоне- ние	Число случаев (n)	Доля отказов, %	Среднее время разгово- ра, сек	Станд. откло- нение
11	216	93	25,74	37,44	18	72	44,94	12,609
12	4	*	29,25	11,701	9	81	46,89	19,199
13	37	*	39,49	101,659	1	14	0	***
14	166	83	25,83	29,644	4	71	31	36,028
15	105	91	24,78	17,683	20	20	43,9	18,784
16	93	96	16,57	12,243	1	**	39	***
17					9	81	53,67	28,262
18					2	67	55	29,698
19					2	**	33,5	7,778
20					2	40	9	0
21					2	15	4	5,657
22					1	20	0	***
23					179	90	41,57	22,853
24					56	85	30,66	29,243
25					366	99	19,3	8,732
26					2	13	18,5	10,607
27					1	4	0	***
28					24	89	31,83	11,293
29					1	17	0	***
30					3	60	26	1
31					10	77	48,7	23,372
32					2	67	36	14,142
Всего	3637	93	19,08	28,964	1628	87	28,95	16,367

*Одно результативное интервью.

** Данных нет или интервьюер не опознан в финальном массиве.

*** Стандартное отклонение не рассчитано.

Зафиксированные различия были вызваны в первую очередь организационными моментами. В телефонном опросе В1Т1 согласно финальному массиву и массиву отказов было задействовано в два раза меньше интервьюеров, чем в опросе В2Т2. На одного интервьюера-оператора в опросе В1Т1 пришлось максимум 81 результативное интервью, минимум — 4, а в опросе В2Т2—25 и 2 соответственно. Вопрос о том, почему респонденты оказались так неравномерно распределены между интервьюерами-операторами, остался открытым. Также необходимо отме-

титель существенный разброс в длительности разговоров. Интервьюеры-операторы В1Т1 тратили на разговор в среднем на десять секунд больше, чем интервьюеры-операторы В2Т2. Возможно, это одна из причин, почему уровень достижимости в опросе В1Т1 существенно ниже, чем в опросе В2Т2. Практически у всех интервьюеров В1Т1 (за исключением случаев, которые не удалось идентифицировать) зафиксирован высокий уровень отказов, который колеблется в интервале 83%-98%. Напротив, у интервьюеров компании В2Т2 разброс в уровне отказов оказался значительным — от 4% до 99%.

По техническим причинам мы не смогли рассчитать доли отказов для интервьюеров компании В4S2, проводивших опрос на планшетах: и в массиве отказов, и в финальном массиве они были закодированы. Можно только сказать, что в опросе, проводимом компанией В4S2, принимало участие 11 интервьюеров, и на каждого приходилось от 16 до 37 результативных интервью. Мы проанализировали доли отказов для уличного опроса с использованием бумажных анкет (компания В3S1): на одного интервьюера здесь приходилось от 15 до 60 результативных интервью, а средняя доля отказов составила 41% (разброс отказов между интервьюерами зафиксирован в диапазоне от 27% до 54%).

Мы полагаем, что на высокую дисперсию отказов у интервьюеров/операторов влияют следующие факторы:

1) Отсутствие понимания важности и полезности сбора параданных. Интервьюер должен понимать важность сбора параданных для компании и для себя лично (оплата труда).

2) Плохо составленные инструменты и инструкции по сбору параданных приводят к низкой мотивации для выполнения этого задания и, как следствие, существенному влиянию интервьюера на качество и полноту собранных данных об отказах.

Заключение

На основе материалов открытого исследования мы попытались рассчитать уровень достижимости респондентов в комбинированном опросе, сравнить и оценить качество параданных, полученных во время телефонных и уличных интервью. В результате мы выделили проблемы и ограничения, возникающие при сборе параданных в различных типах опросов: параданные содержат много пропусков и неточностей; они могут включать ошибки измерения, приводящие к неверным расчетам уровня достижимости и оценкам результативности исследования в целом. Мы пришли к выводу, что проблемы качества параданных связаны, прежде всего, с отсутствием единой методики, адаптированной к разным методам сбора данных. Однако эти проблемы усугубляются, если плохо проинструктировать интервьюеров по поводу важности сбора информации, связанной с отказами, в результате чего они относятся к этому разделу своей работы недостаточно ответственно.

Ни одна из исследовательских организаций, задействованных в изучаемом нами комбинированном опросе, не отслеживала уровень достижимости респондентов в соответствии со стандартами AAPOR. Это обусловлено тем, что диспозиционные коды, предложенные AAPOR, основаны на практике сбора данных в США и, не смотря на отдельные попытки [Звоновский, Соловьева, 2017], на се-

годняшний день нет четкого представления о том, как адаптировать их к российскому контексту. Восполнение этого пробела ставит перед отечественными исследователями актуальную методическую задачу: адаптировать расчет уровня достижимости респондентов к реальным условиям сбора данных, сделать эту процедуру стандартизированной, хорошо документированной и воспроизводимой.

Показатель уровня достижимости представляет большую ценность для методологии и практики опросных компаний. Потребность в сопоставимости растущего числа исследований, возрастающая значимость открытой науки — все это требует стандартизации процедур отчетности. Пока не сформированы общие критерии и стандарты сбора параданных и представления результатов, у исследователей будут возникать методологические проблемы и сложности [Ипатов, 2014]. Однако сбор сопутствующих основному исследованию данных требует дополнительно времени, финансовых затрат и более высокой квалификации интервьюеров. Следовательно, одной из проблем оказывается то, что не все интервьюеры понимают и принимают необходимость фиксации дополнительной информации в ходе опроса, что приводит к неточностям, пропускам и ошибкам [там же].

Делают ли низкие показатели отклика опрос ненадежным? Является ли количество ответов хорошим показателем качества опроса? Наш короткий ответ здесь — «нет». Как мы смогли убедиться, сами по себе показатели достижимости ничего не говорят о том, увеличивается ли систематическая ошибка при их низких значениях. Наше исследование показало, что систематическая ошибка, связанная с низким уровнем достижимости, не влияет на результаты опросов и оказывается управляемой проблемой.

Исходя из того, как общество продолжает меняться, а технологии развиваются, в будущем социологические исследования, вероятно, будут представлять собой комбинацию опросов и других форм сбора данных. Уменьшение количества ошибок при фиксации параданных должно стать целью каждой исследовательской организации. Для этого есть как минимум два пути. Первый — более тщательно обучать и инструктировать интервьюеров, контролировать их в полевых условиях. Второй — активно использовать методы автоматического сбора параданных. Для обоснования этих предложений необходимы дальнейшие исследования, направленные на понимание природы ошибок измерения и их последствий. Качественная и количественная оценка этих ошибок — сложная задача. Такие исследования дороги и, следовательно, редки. Но наш анализ показал (во всяком случае, мы на это надеемся), что дальнейшие подобные исследования оправданы.

Список литературы (References)

- Звоновский В. Б., Соловьева Ю. В. Адаптация системы конечных кодов AAPOR к российской практике CATI-опросов // Мониторинг общественного мнения: экономические и социальные перемены. 2017. № 3. С. 59—72. <https://doi.org/10.14515/monitoring.2017.3.06>.
- Zvonovski V. B., Solovieva Yu. V. (2017) Adaptation of the AAPOR Final Code System for CATI Technique in Russia. *Monitoring of Public Opinion: Economic and Social Changes*. No. 3. P. 59—72. (In Russ.) <https://doi.org/10.14515/monitoring.2017.3.06>. (In Russ.)

Ипатова А. А. Использование параданных в анализе телефонных опросов // Телескоп. 2014. № 6. С. 34—31.

Ipatova A. A. (2014) The Use of Paradata in the Analysis of Telephone Surveys. *Telescope: A Journal of Sociological and Marketing Research*. No. 6. P. 34—41. (In Russ.)

Корытничкова Н. В. Стадии повышения уровня достижимости респондентов при проведении интернет-опросов // Социология: теория, методы и маркетинг. 2013. № 1. С. 80—94.

Korytnikova N. V. (2013) Stages of Increasing the Level of Reachability of Respondents When Conducting Internet Surveys. *Sociology: Theory, Methods and Marketing*. No. 1. P. 80—94. (In Russ.)

Лебедев Д. В. Параданные: определения, типы, сбор и возможное применение // Мониторинг общественного мнения: экономические и социальные перемены. 2020. № 2. С. 4—32. <https://doi.org/10.14515/monitoring.2020.2.915>.

Lebedev D. V. (2020) Paradata: Definition, Types, Collection, and Possible Uses. *Monitoring of Public Opinion: Economic and Social Changes*. No. 2. P. 4—32 <https://doi.org/10.14515/monitoring.2020.2.915>. (In Russ.)

Османов Т. Э., Рогозин Д. М. Методическое представление общероссийского опроса по мобильным телефонам, или процедуры оценки качества выборочного исследования на примере опроса трудоспособного населения России // Мониторинг общественного мнения: экономические и социальные перемены. 2013. № 2. С. 40—54.

Osmanov T. E., Rogozin D. M. (2013) Methodological Presentation of the All-Russian Survey on Mobile Phones, or the Procedure for Assessing the Quality of a Sample Study on the Example of a Survey of the Working-Age Population of Russia. *Monitoring of Public Opinion: Economic and Social Changes*. No. 2. P. 40—54. (In Russ.)

Паниотто В., Харченко Н. Кризис в методах опроса и пути его преодоления // Вестник общественного мнения. 2012. № 1. С. 100—108. URL: http://www.levada.ru/sites/default/files/vom1_111.pdf (дата обращения: 26.07.2021).

Paniotto V., Kharchenko N. (2012) Crisis in Survey Methods and Ways to Overcome It. *Bulletin of Public Opinion*. No. 1. P. 100—108. URL: http://www.levada.ru/sites/default/files/vom1_111.pdf (accessed: 26.07.2021). (In Russ.)

Рогозин Д. М. Влияние интервьюера на доступность респондентов в телефонном опросе // Социологический журнал. 2004. № 1/2. С. 75—105. URL: <https://jour.fnisc.ru/index.php/socjour/article/view/825/779> (дата обращения: 26.07.2021).

Rogozin D. M. (2004) Influence of the Interviewer on the Availability of Respondents in a Telephone Survey. *Sociological Journal*. No. 1/2. P. 75—105. URL: <https://jour.fnisc.ru/index.php/socjour/article/view/825/779> (accessed: 26.07.2021). (In Russ.)

Рогозин Д. М. Конформная выборка в торговых центрах // Социологический журнал. 2008. № 1. С. 22—49. URL: <https://jour.fnisc.ru/index.php/socjour/article/view/982/936> (дата обращения: 26.07.2021).

Rogozin D. M. (2008) Conformal Sampling in Shopping Centers. *Sociological Journal*. No. 1. P. 22—49. URL: <https://jour.fnisc.ru/index.php/socjour/article/view/982/936> (accessed: 26.07.2021). (In Russ.)

Романович Н. А. Отказы респондентов в зависимости от восприятия ими опроса и его методов // Мониторинг общественного мнения: экономические и социальные перемены. 1996. № 4. С. 42—45.

Romanovich N. A. (1996) Respondents' Refusals Depending on Their Perception of the Survey and Its Methods. *Monitoring of Public Opinion: Economic and Social Changes*. No. 4. P. 42—45. (In Russ.)

Стандартные определения: Систематическое описание диспозиционных кодов и коэффициентов результативности для массовых опросов / Американская ассоциация исследователей общественного мнения; пер. с англ. А. А. Ипатовой, Д. М. Рогозина. 8-е изд. М.: АО «ВЦИОМ», 2016.

American Association of Opinion Researchers. (2016) Standard Definitions: Systematic Description of Disposition Codes and Performance Rates for Mass Surveys. Transl. from English by A. A. Ipatova, D. M. Rogozin. 8th ed. Moscow: VCIOM. (In Russ.)

Юдин Г. Б. Территориальная локализация и уровень неответов в массовом опросе // Социологический журнал. № 1. С. 49—72. URL: <https://jour.fnisc.ru/index.php/socjour/article/view/983/937> (дата обращения: 26.07.2021).

Yudin G. B. (2008) Territorial Localization and the Level of Non-Responses in a Mass Survey. *Sociological Journal*. No. 1. P. 49—72. URL: <https://jour.fnisc.ru/index.php/socjour/article/view/983/937> (accessed: 26.07.2021). (In Russ.)

Wiseman F. (2003) On the Reporting of Response Rates in Extension Research. *Journal of Extension*. Vol. 41. No. 3. URL: <https://archives.joe.org/joe/2003june/comm1.php> (accessed: 26.07.2021).