

А.М. КЛИМОВА, Е.А. ТЕРЕНТЬЕВ

СРАВНЕНИЕ КАЧЕСТВА ДАННЫХ PAPI И CAPi В ЛОНГИТЮДНОМ ИССЛЕДОВАНИИ ДОМОХОЗЯЙСТВ: РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТА В РАМКАХ RLMS–HSE¹

Аннотация. В статье представлены результаты экспериментального исследования того, как переход от традиционного бумажного способа сбора социологических данных (PAPI) к сбору данных с использованием компьютерных технологий (CAPi) влияет на качество данных в лонгитюдных исследованиях домохозяйств. Эксперимент был реализован в 2018–2019 гг. в рамках проведения Российского мониторинга экономического положения и здоровья населения (RLMS–HSE). В предыдущих работах на данных экспериментального исследования в рамках 26-й волны RLMS–HSE было показано, что использование CAPi приводит к значимому уменьшению доли отказов от ответов и доли несодержательных вариантов ответа («затрудняюсь ответить»), а также к значимым различиям в отношении ответов на чувствительные вопросы. В данной работе мы проверяли консистентность полученных результатов с использованием новых данных, собранных в рамках 27-й волны RLMS–HSE. Результаты эксперимента показывают, что использование CAPi приводит к снижению доли отказов от ответа и доли выбравших несодержательные ответы, что способствует повышению качества собираемых данных. Вместе с тем было показано, что использование CAPi может приводить к повышению эффекта социальной желательности.

Климова Айгуль Маратовна — кандидат социологических наук, доцент, Департамент социологии; старший научный сотрудник, Лаборатория сравнительных социальных исследований, Национальный исследовательский университет «Высшая школа экономики».

Адрес: 101000, Москва, Кривоколенный пер., д. 3, каб. 3-334.

Телефон: +7 (495) 772-95-90 * 12244. **Электронная почта:** amavletova@hse.ru

Терентьев Евгений Андреевич — кандидат социологических наук, старший научный сотрудник, Центр социологии высшего образования Института образования, Национальный исследовательский университет «Высшая школа экономики». **Адрес:** 101000, Москва, Потаповский пер., д. 16, стр. 10, каб. 207.

Телефон: +7 (985) 386-63-49. **Электронная почта:** eterentev@hse.ru

¹ Статья подготовлена при поддержке Российского научного фонда. Грант № 17-78-20172.

Ключевые слова: PAPI; CAPI; качество данных; лонгитюдные исследования домохозяйств; RLMS–HSE; доля ответов «затрудняюсь ответить»; чувствительные вопросы; эффект социальной желательности.

Для цитирования: Климова А.М., Терентьев Е.А. Сравнение качества данных PAPI и CAPI в лонгитюдном исследовании домохозяйств: результаты эксперимента в рамках RLMS–HSE // Социологический журнал. 2020. Том 26. № 4. С. 8–30. DOI: 10.19181/socjour.2020.26.4.7640

Введение

Почти все лонгитюдные исследования домохозяйств в Европе и Америке используют одновременно несколько методов сбора социологических данных, включая такие методы, как PAPI (paper and pencil interviewing) и CAPI (computer assisted personal interviewing), когда интервьюер заполняет анкету со слов респондента на бумаге или компьютере/планшете. Стратегия комбинирования разных количественных методов сбора данных была названа в англоязычной литературе mixed-mode (не путать с mixed methods, когда исследователь одновременно применяет качественные и количественные методы сбора данных). В обзоре, посвященном влиянию метода CAPI, когда интервьюеры заполняют анкету на компьютере или планшете, на качество данных в лонгитюдных исследованиях домохозяйств, применявших до этого только метод PAPI, мы показали, что CAPI имеет ряд преимуществ перед PAPI, но также потенциально несет и риски, связанные с возможным увеличением ошибки измерения [6].

В рамках 26-й волны лонгитюдного исследования домохозяйств Российского мониторинга экономического положения и здоровья населения (RLMS–HSE) в 2017–2018 гг. наряду с традиционным методом сбора данных интервьюерами на бумаге (PAPI) был использован метод опроса с использованием компьютерных технологий (CAPI). В обоих методах опрос проводился интервьюером, однако различие состояло в том, что интервьюер заполнял анкету на бумаге или на планшете. В рамках 27-й волны RLMS–HSE в 2018–2019 гг. данный эксперимент был продолжен. Различия между PAPI и CAPI на основе результатов 26-й волны RLMS–HSE были обнаружены в таких индикаторах, как доля ответов «затрудняюсь ответить» и ответы на чувствительные вопросы².

Целью данной статьи является проверка консистентности полученных различий между PAPI и CAPI на основе результатов данных 27-й волны лонгитюдного исследования домохозяйств RLMS–HSE. Таким образом, мы сравним качество данных 27-й волны лонгитюдного исследования домохозяйств RLMS–HSE между PAPI и CAPI по таким индикаторам, как *доля несодержательных ответов* («затрудняюсь ответить» и «отказ от ответа» на определенные вопросы) и *ответы на чувствительные вопросы*.

² Klimova A., Terentev E., Cernat A. An experimental study of paper-based interviewing and computer-assisted interviewing in the longitudinal household survey, 2020 [рукопись].

Обзор литературы

Эффект перехода к использованию CAPI на качество данных изучался экспериментально в ряде срезовых [12; 22; 24; 27], а также в небольшом количестве лонгитюдных исследований [8; 27; 29; 32]. Кроме того, в некоторых работах приводится оценка влияния перехода на различные показатели качества данных, которая не основана на данных экспериментальных исследований [9; 13]. Для сравнения данных, полученных с помощью CAPI и PAPI, использовались следующие индикаторы: 1) время заполнения опросника; 2) уровень отклика и осыпание панели (в случае лонгитюдных исследований); 3) доля неотчетов и несодержательных ответов; 4) доля социально одобряемых ответов; 5) качество ответов на открытые вопросы. Подробный русскоязычный обзор результатов предыдущих лонгитюдных исследований в отношении рассмотренных индикаторов представлен в работе Е.А. Терентьева, А.М. Мавлетовой и М.С. Косолапова [6], поэтому в данной статье не будем подробно останавливаться на них, а сфокусируемся на двух пунктах, связанных с долей отказов от ответа и несодержательных ответов, а также ответов на сенситивные вопросы. В отношении этих показателей были получены значимые различия между CAPI и PAPI при проведении экспериментального исследования в рамках реализации 26-й волны RLMS–HSE³.

Доля неотчетов и несодержательных ответов. Исключение человеческого фактора, с которым связаны ошибки интервьюеров при заполнении опросников, через выстраивание системы автоматических проверок и маршрутизации респондентов, а также дополнительные возможности визуализации вопросов являются одними из важных преимуществ, обычно обсуждаемых в контексте перехода от PAPI к CAPI [6]. Это подтверждается результатами исследований, где фиксируется снижение доли неотчетов [8; 12; 23] при использовании CAPI. Результаты экспериментального исследования, выполненного в рамках 26-й волны RLMS–HSE, консистентны этим результатам и показывают, что использование CAPI приводит к уменьшению доли ответов «затрудняюсь ответить» или других форм ухода от ответа на вопросы⁴. Обнаруженные различия были связаны с тем, что данные возможности ухода от ответа не были представлены на экране планшета в открытых вопросах, и предполагалось, что интервьюеры должны были вводить определенный набор цифр в текстовое поле для обозначения ухода респондента от ответа. Отсутствие подобных возможностей уйти от ответа и необходимые дополнительные усилия интервьюеров, чтобы обозначить, что респондент затрудняется с ответом или отказывается отвечать на вопрос,

³ Klimova A., Terentev E., Cernat A. An experimental study of paper-based interviewing and computer-assisted interviewing in the longitudinal household survey, 2020 [рукопись].

⁴ Там же.

привели к уменьшению доли таких малосодержательных ответов на 17% ($t(2211) = 4,90, p < 0,001$).

Вместе с тем некоторые исследования показывают, что методы сбора данных могут по-разному влиять на долю неотвеченных и несодержательных ответов в зависимости от типа вопросов. Так, в работах Р. Бейкера и коллег [8], Э. Де Лю и коллег [17] показывается, что использование САPI может привести к снижению доли неотвеченных на чувствительные вопросы. Предложенное объяснение этого эффекта — респонденты воспринимают ситуацию интервью с использованием компьютерных технологий как более анонимную и приватную. Это, в свою очередь, приводит к тому, что респонденты более склонны раскрывать личную информацию в компьютерном опросе. В качестве одного из подтверждений такого обоснования Бейкер и коллеги приводят результаты опроса респондентов после проведения интервью, где значительная доля респондентов отметили САPI как более конфиденциальный способ сбора данных по сравнению с РАPI [8].

Однако противоположные результаты были получены в рамках экспериментального исследования на базе Немецкого социально-экономического панельного исследования домохозяйств: использование САPI привело к значимому увеличению доли неотвеченных на чувствительные вопросы о валовом доходе семьи [29]. Для объяснения этого эффекта авторы обращаются к противоположному объяснению Бейкера и коллег и отмечают, что, возможно, использование компьютеров приводит к возникновению дополнительных опасений респондентов в отношении конфиденциальности и приватности процесса сбора данных.

Доля социально одобряемых ответов. Интерес представляют также возможные влияния использования САPI на содержательные ответы на чувствительные вопросы, в первую очередь на долю социально одобряемых ответов. Эти эффекты изучались в ряде предыдущих исследований, которые показали противоречивые результаты [8; 22; 24; 27; 32].

В ряде исследований в качестве ключевого индикатора использовалась доля выбравших крайние варианты ответа [24; 27; 32]. Отмечалось, что большее значение этого показателя может свидетельствовать о меньшем эффекте социальной желательности и, как следствие, о большей надежности данных. Мартин и коллеги обнаружили большую долю респондентов, выбравших крайние варианты ответа на чувствительные вопросы при использовании САPI [27]. Схожие результаты были получены в исследовании Н. Уотсона и Р. Уилкинса в отношении некоторых вопросов (о потреблении овощей и о курении) [32]. Однако противоположные результаты были получены в отношении установочных вопросов об удовлетворенности разными аспектами своей жизни: использование САPI привело к уменьшению доли выбравших крайние варианты ответа [32]. П. Линн также получил противоречивые результаты в отношении эффекта САPI на долю выбравших крайние варианты ответа: для 7 из 13 вопросов, по которым были обнаружены значимые различия,

фиксирувалась большая доля выбравших крайние варианты ответа при использовании CAPI, для оставшихся шести вопросов был получен противоположный результат [24]. В исследовании, выполненном в рамках 26-й волны RLMS–HSE, мы не обнаружили значимых различий в доле респондентов, использующих крайние оценки в шкальных вопросах, и в доле пропущенных данных (когда был исключен небольшой блок вопросов, в которых допущена ошибка программирования CAPI)⁵.

Однако мы обнаружили значимые различия при сравнении средних значений ответов на некоторые чувствительные вопросы с использованием CAPI и PAPI. В одних вопросах эффект социальной желательности был выше в данных CAPI (меньшее значение алкогольного индекса и меньшее количество потребляемых сигарет по сравнению с данными PAPI), а в других — он оказался выше в данных PAPI (количество форм физической активности за последние 12 месяцев)⁶. Вместе с тем не было найдено различий в доле респондентов, отметивших, что курят или что пили алкогольные напитки за последние 30 дней. Не было также обнаружено различий между данными PAPI и CAPI в отмеченном респондентами возрасте, в котором они начали курить или пить алкогольные напитки. В целом такая противоречивость характерна и для других исследований, измерявших влияние перехода от PAPI к CAPI на качество данных. Так, Р. Бейкер и коллеги выявили снижение эффекта социальной желательности при ответе на вопросы о потреблении алкоголя и об использовании контрацептивов, однако не обнаружили этот эффект для ряда других вопросов [8]. П. Линн выявил значимые различия между CAPI и PAPI в отношении 15 из 90 вопросов на установки. Однако направление эффекта варьировалось для разных вопросов, что не позволило прийти к консистентным выводам [24]. Н. Уотсон и Р. Уилкинс отметили снижение эффекта социальной желательности при использовании CAPI для вопросов о курении и потреблении овощей, однако усиление этого эффекта было обнаружено для вопроса о валовом доходе семьи, что также не позволило авторам прийти к консистентным выводам в отношении полученных результатов [32].

Исходя из противоречивости выводов, полученных в предыдущих исследованиях, крайне актуальным представляется дальнейшее изучение влияния способа сбора данных на качество данных с учетом различий в типах вопросов. В рамках данной статьи мы ставим цель проверить консистентность полученных различий между PAPI и CAPI в отношении доли ответов «затрудняюсь ответить» или отказов от ответа на определенные вопросы, а также эффекта социальной желательности при ответах на сенси-

⁵ Klimova A., Terentev E., Cernat A. An experimental study of paper-based interviewing and computer-assisted interviewing in the longitudinal household survey, 2020 [рукопись].

⁶ Эти результаты будут опубликованы в: Klimova A., Terentev E., Cernat A. An experimental study of paper-based interviewing and computer-assisted interviewing in the longitudinal household survey, 2020 [рукопись].

тивные вопросы. Основываясь на результатах предыдущих исследований, мы тестируем *две ключевые гипотезы*. При этом из-за противоречивости результатов, имеющей отношение к эффекту социальной желательности во второй гипотезе, мы выдвигаем два альтернативных предположения.

H1: Использование технологии CAPI приведет к *уменьшению* доли ответов «затрудняюсь ответить» или других форм ухода от ответа на вопросы по сравнению с методом PAPI.

H2: CAPI приведет к *ослаблению* эффекта социальной желательности при ответах на чувствительные вопросы по сравнению с PAPI.

H2* (альтернативная гипотеза): CAPI приведет к *усилению* эффекта социальной желательности при ответах на чувствительные вопросы по сравнению с PAPI.

Эмпирическая база и экспериментальный дизайн

Экспериментальное исследование было реализовано в рамках проведения 27-й волны обследования RLMS–HSE. Этот мониторинг реализуется с 1994 г. и является единственным общероссийским мониторинговым исследованием социально-экономического положения и здоровья населения в России. Ключевые темы, рассматриваемые в исследовании, касаются миграционной истории, образовательных и профессиональных траекторий, оценки социального благополучия, здоровья и экономического положения. Исследование базируется на использовании вероятностной стратифицированной многоступенчатой территориальной выборки⁷ и предполагает сбор данных на индивидуальном уровне и на уровне домохозяйств. Результаты используются как для академических целей, так и для проектирования, поддержки и оценки различного рода реформ.

Сбор данных 27-й волны обследования осуществлялся в период с октября 2018 г. по январь 2019 г. Для проведения экспериментального исследования использовались индивидуальные опросники для взрослого населения. В исследовании приняли участие 58 интервьюеров и 3065 респондентов в возрасте 14 лет и старше. 631 респондент (21%) был опрошен с использованием CAPI, 2434 (79%) — с использованием PAPI.

Экспериментальное исследование было реализовано в восьми населенных пунктах: в Москве, Казани, Нижнем Новгороде, Саратове, Кургане, Соликамске, Вольске и Бердске. Распределение интервьюеров и респондентов по населенным пунктам, а также ключевые характеристики этих населенных пунктов представлены в таблице 1. Отбор участников исследования проходил в несколько этапов. На первом этапе были отобраны опросные точки (первичные единицы отбора); на втором — интервьюеры; на третьем этапе отбирались домохозяйства и респонденты для участия в эксперименте.

⁷ Подробнее о выборке и принципах ее построения см. страницу на сайте RLMS–HSE: URL: <https://www.hse.ru/rlms/sample>

Таблица 1

Основные характеристики населенных пунктов, где проводилось экспериментальное исследование, и данные о количестве интервью с использованием PAPI и CAPI в каждом из пунктов

Название населенного пункта	Численность населения, чел.	Административно-территориальный статус	Количество:		
			интервью-еров, чел.	интервью в PAPI	интервью в CAPI
Москва (включая Новую Москву и Московскую область)	12 000 000	столица РФ мегаполис	18	976	218
Казань	1 250 000	столица Республики Татарстан	7	227	73
Нижний Новгород	1 259 000	областной центр Нижегородская область	5	303	47
Саратов	845 000	областной центр Саратовская область	6	203	60
Курган	325 000	областной центр Курганская область	7	206	73
Соликамск	95 000	райцентр Пермский край	4	107	46
Вольск	64 000	райцентр Саратовская область	5	181	52
Бердск	100 000	Новосибирская агломерация	6	231	62
Всего			58	2434	631

Отбор населенных пунктов. Отбор первичных единиц осуществлялся в два этапа: в рамках реализации экспериментального исследования в ходе проведения 26-й и 27-й волн исследования RLMS–HSE. На первом этапе (26-я волна) отбор осуществлялся из списка поселений, участвующих в RLMS–HSE. Были отобраны только городские поселения, которые стратифицированы по численности населения и административно-территориальному статусу. Далее в каждой страте были отобраны городские поселения, где работали наиболее квалифицированные супервайзеры, которые могли провести обучение интервьюеров и помочь с использованием CAPI в поле. В результате было выделено шесть первичных единиц отбора: Москва и Московская область, Казань, Курган, Соликамск, Вольск и Бердск. На втором этапе (27-я волна) к городским поселениям, отобранным на первом этапе, добавились еще два городских поселения — Нижний Новгород и Саратов. Принцип отбора на

втором этапе повторял процедуру первого этапа: были отобраны городские поселения со статусом областных центров, где работают наиболее компетентные супервайзеры и интервьюеры, которые могли справиться с проведением опроса с использованием CAPI.

Отбор и обучение интервьюеров. Отбор интервьюеров также проводился в два этапа: в рамках реализации экспериментального исследования в ходе 26-й и 27-й волн RLMS–HSE. На обоих этапах отбор проводился среди всех интервьюеров, работающих в первичных единицах отбора (городских поселениях), по трем основным критериям, которые были выбраны руководителями RLMS–HSE: 1) ответственность кандидатов (оценивалась супервайзером), 2) способность освоить процедуру работы с планшетом (оценивалась супервайзером) и 3) согласие принять участие в эксперименте. На первом этапе (26-я волна) были отобраны 38 интервьюеров. Еще 20 интервьюеров были дополнительно отобраны на втором этапе (27-я волна). Отбор дополнительных интервьюеров проводился как в новых опросных точках, так и в тех, которые участвовали в экспериментальном исследовании в рамках 26-й волны RLMS–HSE. В результате общее количество интервьюеров, участвовавших в экспериментальном исследовании, составило 58 человек.

Отбор домохозяйств и респондентов. На третьем этапе проводился случайный отбор домохозяйств для каждого интервьюера, участвующего в эксперименте, с помощью процедуры механического отбора с заданным шагом. Отбор производился до выхода в поле руководителя исследования RLMS–HSE, с использованием полного списка домохозяйств, участвующих в опросе в отобранных городских поселениях. Изначально предполагалось, что в среднем в эксперименте каждый интервьюер должен опросить взрослых респондентов с использованием CAPI не менее чем в пяти домохозяйствах. Поэтому шаг отбора для каждого интервьюера рассчитывался как общее количество домохозяйств, опрашиваемых данным интервьюером, деленное на 5. Поскольку в среднем на каждое домохозяйство приходится по два респондента⁸, ожидалось, что каждый интервьюер опросит по 10 респондентов с использованием CAPI. Если 10 взрослых членов домохозяйств уже опрошены на планшете, но в последнем из опрашиваемых домохозяйств еще оставались взрослые члены семьи, которых нужно было опросить, то интервьюер должен был опросить на планшете всех оставшихся. Из-за разницы в составе домохозяйств, приписанных к процедуре CAPI, количество опрошенных с использованием CAPI для разных интервьюеров варьируется между 8 и 16 респондентами.

Программное обеспечение и планшеты

Для разработки и программирования электронных опросных форм применялось программное обеспечение Survey Solutions, разработанное Всемирным банком и свободно распространяемое для проведения ис-

⁸ В соответствии с данными по предыдущим волнам RLMS–HSE.

следований с использованием компьютерных технологий⁹. Электронная версия опросника была запрограммирована таким образом, чтобы минимизировать возможные различия между его бумажной и электронной версиями. В электронной версии использовалась система маршрутизации респондентов внутри опросника (фильтрация показа вопросов в зависимости от ответов на предыдущие), которая повторяла инструкции для интервьюеров в бумажной версии. Также в электронной версии опросника использовались валидационные процедуры: 1) проверки ответов на вопросы, предполагающие ввод числовых данных, на соответствие ожидаемому разбросу допустимых значений (например, год рождения, количество рабочих часов в неделю, потребляемые объем алкоголя и количество сигарет и др.); 2) проверки на консистентность ответов на разные вопросы внутри опросника (например, год рождения и год начала проживания по текущему месту жительства). В случае нарушения допустимых условий проверок интервьюер должен был скорректировать ответ либо в специальном поле дать обоснование, почему отмеченный вариант является корректным. Некоторые вопросы сопровождалась специальными инструкциями для интервьюеров по особенностям заполнения ответов, которые полностью повторяли инструкции, представленные в бумажной версии опросника. При организации сбора данных не использовалась возможность предварительной загрузки в систему данных о респондентах из предыдущих волн обследования.

Для сбора данных использовались планшетные компьютеры Samsung Galaxy Tab A 16,0 SM-T355. Большая часть планшетов (40 штук) были закуплены для проведения экспериментального исследования в рамках 26-й волны RLMS–HSE. Для проведения экспериментального исследования в рамках 27-й волны обследования были закуплены дополнительно 20 планшетных компьютеров. Выбор модели планшетного компьютера обусловлен тремя группами характеристик: 1) требованием к устройствам для использования программного обеспечения Survey Solutions¹⁰, 2) техническими характеристиками устройств (вес, размер, объем встроенной памяти и др.), 3) ценовыми характеристиками.

Данные

В эксперименте в рамках 27-й волны лонгитюдного исследования домохозяйств RLMS–HSE приняли участие 3065 респондентов в возрасте 14 лет и старше: 2434 респондента (79%) были опрошены методом PAPI, 631 респондент (21%) — методом CAPI.

Обнаружены различия в нескольких базовых социально-демографических характеристиках между выборками PAPI и CAPI (табл. 2). Так, немного меньше средний возраст респондентов, опрошенных методом

⁹ Подробную информацию о системе см. на сайте Survey Solutions 20.10: URL: <https://mysurvey.solutions/>

¹⁰ См. сайт Survey Solutions: URL: <http://support.mysurvey.solutions/customer/en/portal/articles/2505822-what-tablets-should-i-buy->

РАPI, по сравнению с возрастом респондентов САPI ($M = 45,5$, $SD = 18,82$ и $M = 49,6$, $S = 19,05$ соответственно, $t(3065) = 4,82$, $p < 0,001$). Были также обнаружены некоторые различия в распределении по населенным пунктам ($\chi^2(7) = 32,51$, $p < 0,001$). Кроме того, в данных РАPI была зафиксирована большая доля работающих респондентов по сравнению с данными САPI (57% и 45% соответственно, $\chi^2(1) = 28,91$, $p < 0,001$). Вместе с тем не было обнаружено различий по полу (41% мужчин, $\chi^2(1) = 0,17$, $p = 0,68$), по семейному статусу (59% живут в гражданском или зарегистрированном браке, $\chi^2(1) = 3,21$, $p = 0,07$) и в доле респондентов с высшим образованием (31%, $\chi^2(1) = 0,19$, $p = 0,67$).

Таблица 2

Таблица ковариат

Ковариаты	РАPI	САPI	Статистический критерий
доля мужчин, %	41,6	40,6	$\chi^2(1) = 0,17$, $p = 0,68$
средний возраст, лет	45,5 (SD = 18,82)	49,6 (SD = 19,05)	$t(3063) = 4,82^{***}$
распределение по населенным пунктам, %	Соликамск = 4,4 Казань = 9,3 Курган = 8,5 Саратов = 8,3 Вольск = 7,4 Нижний Новгород = 12,4 Москва, Новая Москва, Московская область = 40,1 Бердск = 9,5	Соликамск = 7,3 Казань = 11,6 Курган = 11,6 Саратов = 9,5 Вольск = 8,2 Нижний Новгород = 7,4 Москва, Новая Москва, Московская область = 34,5 Бердск = 9,8	$\chi^2(7) = 32,51^{***}$
имеют работу, %	57,0	45,0	$\chi^2(1) = 28,91^{***}$
живут в гражданском или зарегистрированном браке, %	58,2	62,2	$\chi^2(1) = 3,21$, $p = 0,07$
имеют высшее образование, %	31,1	30,1	$\chi^2(1) = 0,19$, $p = 0,67$
N	2434	631	

* $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$; SD — стандартное отклонение.

Индикаторы сравнения данных

Были использованы следующие *индикаторы качества данных*:

— *доля несодержательных ответов*. Рассчитывалась доля ответов «затрудняюсь ответить» или отказов от ответа от общего количества вопросов, на которые должен был ответить респондент;

— *ответы на сенситивные вопросы*. Мы сравнили качество данных в следующих вопросах:

- возраст, в котором респондент впервые попробовал курить (wm72);
- возраст, в котором респондент впервые попробовал алкогольные напитки (wm80.2);
- доля курящих (wm71);
- доля пьющих алкогольные напитки в течение последних 30 дней (wm80.0);
- среднесуточный объем потребления «чистого алкоголя» (на основе блока вопросов wm8411a–wm84.7d).

Среднесуточный объем потребления алкоголя был рассчитан на основе вопросов о частоте и количестве потребления алкогольных напитков (более подробно о методике см. в работе [5]). На основе вопросов о частоте и количестве потребления разных типов алкогольных напитков (пиво промышленного производства, домашнее пиво, домашнее вино, вино промышленного производства, крепленое вино, самогон, водка, коньяк, виски, ликер, алкогольные коктейли) был рассчитан суммарный среднесуточный объем потребленного «чистого алкоголя»;

- среднесуточное количество потребления сигарет/папирос/трубок (wm75);

- количество форм физической активности за последние 12 месяцев (блок вопросов wm113.1a–wm113.6d): занимался ли респондент такими формами физической активности, как бег трусцой, упражнения на тренажерах, езда на велосипеде, баскетбол, волейбол, футбол, хоккей и т. д. Максимальное количество форм физической активности — 11;

- общее количество занятий разными формами физической активности за последние 12 месяцев (блок вопросов wm113.1a–wm113.6d). Рассчитано общее количество занятий на основе вопросов о количестве месяцев занятий той или иной формой физической активности и о частоте занятий в месяц. Данное значение было суммировано по всем формам активности;

- общее время занятий физической активностью за последние 12 месяцев (в часах, блок вопросов wm113.1a–wm113.6d). Рассчитано общее количество часов занятий на основе вопросов о количестве месяцев занятий той или иной формой физической активности, о частоте занятий в месяц и о длительности занятия в часах. Данное значение было суммировано по всем формам активности.

По всем индикаторам сравним качество данных между интервьюированием методами РАPI и САPI, используя статистические критерии, а также измеряя величину эффекта. В качестве основной величины эффекта будем использовать стандартизованную разницу средних d Коэна [15; 16]. Принятая интерпретация полученных результатов состоит в том, что d Коэна, равная 0,2, означает минимальную величину эффекта, 0,5 — среднюю величину эффекта, 0,8 и выше — большую величину эффекта. Однако следует иметь в виду, что в методологических

экспериментах, когда речь идет о сравнении разных методов сбора данных между собой, величина эффекта в классическом понимании является минимальной (см., например: [21]). Эта величина эффекта для исследователей является существенной с учетом того, что она обозначает различие в методе сбора одних и тех же данных.

Результаты

Доля несодержательных ответов. В соответствии с ожиданиями мы обнаружили *меньшую долю ответов «затрудняюсь ответить» или «отказ от ответа» в опросе методом CAPI по сравнению с опросом методом PAPI*: 0,029 (SD = 0,016) и 0,037 (SD = 0,019) соответственно, $t(3063) = 9,12$, $p < 0,001$, критерий Уилкоксона $W = 953830$, $p < 0,001$; d Коэна = 0,42 (см. табл. 3). Величина эффекта приближается к средней, что говорит о достаточно *большом снижении доли несодержательных ответов в опросе CAPI по сравнению с опросом PAPI*. Линейная регрессионная модель с включением контрольных переменных, по которым наблюдались различия между выборками, также показала статистически значимые различия между данными CAPI и PAPI ($\beta = -0,007$, $p < 0,001$).

Ответы на чувствительные вопросы. В ряде чувствительных вопросов *не было обнаружено различий между данными опросов методами PAPI и CAPI*. Так, не найдено различий при сопоставлении отмеченного респондентами возраста, в котором они впервые попробовали курить ($t(875) = 0,35$, $p = 0,72$, табл. 3) или пить алкогольные напитки ($t(1943) = 0,84$, $p = 0,40$). Не было обнаружено и различий в доле тех, кто отметил, что курит ($M = 29,2\%$, $\chi^2(1) = 1,36$, $p = 0,24$, табл. 3) или пьет алкогольные напитки ($M = 66,9\%$, $\chi^2(1) = 3,15$, $p = 0,08$). Вместе с тем вопреки результатам 26-й волны RLMS–HSE в рамках 27-й волны не было обнаружено статистически значимых различий в значении среднесуточного объема потребленного «чистого алкоголя» ($M = 11,83$, $SD = 21,37$, $t(1413) = 0,20$, $p = 0,84$, табл. 3).

Однако *большой эффект социального одобрения был обнаружен при опросе методом CAPI в вопросе о количестве потребляемых сигарет и в блоке вопросов о занятиях разными формами физической активности*. Так, меньшее количество потребляемых сигарет было обнаружено методом CAPI по сравнению с методом PAPI: 13,94 (SD = 6,60) и 15,19 (SD = 8,06) соответственно, $t(889) = 1,88$ (табл. 3). Хотя статистическая значимость $p = 0,06$ не достигает классического значения 0,05, отметим, что величина эффекта показывает наличие существенных различий между методами для эффекта социального одобрения (d Коэна = 0,17, см. табл. 3). Различия между CAPI и PAPI подтверждают и линейные регрессионные модели, в которых были включены контрольные переменные. Так, при интервьюировании методом CAPI респонденты отмечают меньшее потребление сигарет в среднем на 1,5 штуки в день ($\beta = -1,5$, $p < 0,05$).

Эффект социального одобрения также обнаружен выше при CAPI-опросе в блоке вопросов о занятиях разными формами физической активности. Хотя нет статистически значимых различий в ответах на вопросы о количестве форм физической активности, которыми занимались респонденты

за последние 12 месяцев ($M = 0,48$, $SD = 0,87$ в PAPI и $M = 0,45$, $SD = 0,97$ в CAPI, $t(3063) = 0,90$, $p = 0,37$, табл. 3), мы обнаружили различия в значениях суммарного количества занятий и общего времени занятий физической активностью. Оба значения были выше в опросах методом PAPI по сравнению с опросами методом CAPI, а d Коэна варьировалась от 0,18 до 0,25 (см. табл. 3). При этом линейные регрессионные модели, контролирующие ряд социально-демографических переменных, показывают значимость различий на уровне $p = 0,09$ для индикатора количества занятий физической активностью и $p = 0,001$ для общего времени занятий физической активностью.

Таблица 3

Значения индикаторов качества данных при опросах методами PAPI и CAPI

Индикатор качества данных	PAPI	CAPI	Статистический критерий	d Коэна
доля ответов «затрудняюсь ответить» или «отказ от ответа»	0,037 (SD = 0,019)	0,029 (SD = 0,016)	$t(3063) = 9,12^{***}$	0,42
возраст, в котором респондент впервые попробовал курить (wm72)	17,24 (SD = 3,32)	17,39 (SD = 4,37)	$t(875) = 0,35$, $p = 0,72$	0,03
возраст, в котором респондент впервые попробовал алкогольные напитки (wm80.2)	17,29 (SD = 3,32)	17,13 (SD = 3,21)	$t(1943) = 0,84$, $p = 0,40$	0,05
доля курящих (wm71), %	29,8	27,3	$\chi^2(1) = 1,36$, $p = 0,24$	0,05
доля пьющих алкогольные напитки (wm80.0), %	67,7	63,9	$\chi^2(1) = 3,15$, $p = 0,08$	0,08
среднесуточный объем потребленного «чистого алкоголя»	11,88 (SD = 22,30)	11,59 (SD = 16,78)	$t(1413) = 0,20$, $p = 0,84$	0,01
среднесуточное количество потребления сигарет/папирос/трубок (wm75)	15,19 (SD = 8,06)	13,94 (SD = 6,60)	$t(889) = 1,88x$	0,17
количество форм физической активности за последние 12 месяцев	0,48 (SD = 0,87)	0,45 (SD = 0,97)	$t(3063) = 0,90$, $p = 0,37$	0,04
количество занятий физической активностью за последние 12 месяцев	191,08 (SD = 154,76)	222,48 (SD = 187,86)	$t(933) = 2,26^*$	0,18
общее время занятий физической активностью за последние 12 месяцев, час.	185,12 (SD = 188,00)	247,99 (SD = 300,62)	$t(933) = 3,44^{***}$	0,25

* $p < 0,06$; * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

Дискуссия

Мы показали, что способ программирования вопросов в CAPI оказывает влияние на качество данных. Так, отсутствие эксплицированной возможности уйти от содержательного ответа привело к значительному

уменьшению доли ответов «затрудняюсь ответить» или других форм ухода от ответа при опросе методом САPI. В интервьюировании методами РАPI и САPI есть четыре основные возможности изображения и озвучивания интервьюером ответов «затрудняюсь ответить» или «отказ от ответа». Первая возможность — эксплицитно представить для респондента ответы на бумаге или экране и интервьюеру озвучить эти ответы. Вторая возможность — представить ответы на бумаге или экране, однако интервьюер не озвучивает эти ответы респонденту. Третья возможность — не представлять эксплицитно данные ответы на бумаге или экране, однако разрешить интервьюеру вносить определенные коды для обозначения таких ответов. И наконец, четвертая возможность — не представлять данные ответы на бумаге или экране. Подобная опция может усложниться при опросе методом САPI, когда в случае пропуска вопроса на экране показывается повторная просьба ответить на вопрос, однако на экране уже появляются ответы «затрудняюсь ответить» и «отказ от ответа».

В рамках анкеты RLMS—HSE возможности уйти от ответа в открытых вопросах были реализованы по-разному. Если в опросе методом РАPI реализована вторая опция, то есть интервьюеры видели на бумаге ответы «затрудняюсь ответить» и «отказ от ответа», однако не озвучивали их, то при использовании метода САPI была реализована третья опция, когда интервьюеры не видели на экране ответов «затрудняюсь ответить» и «отказ от ответа», однако могли внести определенные коды в текстовое поле для ответа на открытый вопрос, чтобы обозначить формы ухода от ответа. Данная особенность была связана с ограничениями программного обеспечения Survey Solutions, когда для открытых вопросов невозможно добавить дополнительные ответы.

Ряд исследований показали значимое увеличение доли несодержательных ответов в случае эксплицитного представления форм ухода от ответа или в случае наличия вопроса-фильтра о том, могут ли респонденты ответить на вопросы по этой теме или нет, по сравнению с имплицитным представлением, когда интервьюеры не зачитывают ответы «затрудняюсь ответить» или «отказ от ответа» [10; 30]. В онлайн-опросах на самозаполнение показано, что отсутствие эксплицитно представленных форм ухода от ответа на экране снижает долю неотчетов и даже может приводить к повышению надежности полученных данных по сравнению со случаями, когда формы ухода от ответа эксплицитированы на экране [18; 19]. Есть также ряд экспериментов в онлайн-опросах, показывающих, что расположение ответов «затрудняюсь ответить» или «отказ от ответа» на экране, а также цвет и размер шрифта могут оказывать влияние на долю несодержательных ответов [4; 31]. В рамках нашего исследования мы обнаружили, что отсутствие на экране ответов «затрудняюсь ответить» и «отказ от ответа» при использовании метода САPI привело к значительному снижению доли несодержательных ответов. Хотя при использовании обоих способов сбора опросных данных эта доля была очень маленькой,

при использовании CAPI она была меньше на 22%. Стандартизованная разница средних d Коэна = 0,42 указывает на среднюю величину эффекта, что свидетельствует о весьма существенном различии между двумя типами опросов. Такая стандартизованная разница говорит о том, что среднее значение доли несодержательных ответов при опросе по технологии PAPI соответствует 66-му перцентилю доли несодержательных ответов, полученных при опросе методом CAPI.

Мы также обнаружили *увеличение эффекта социального одобрения при сборе данных методом CAPI по сравнению с использованием метода PAPI*. Метаанализ, измеряющий различия в эффекте социального одобрения между двумя методами, когда респонденты самостоятельно заполняют анкеты без интервьюера либо на бумаге, либо на компьютере, показывает снижение эффекта социального одобрения в чувствительных вопросах при заполнении анкеты на компьютере [21]. Среднее метааналитическое значение отношения шансов — величины эффекта, измеряемой исследователями на основе 460 экспериментов, оказалось равным 1,24 [21]. Это небольшая величина эффекта, которая примерно равна d Коэна 0,07–0,09 [14]. При этом величина эффекта была выше в более чувствительных вопросах. Объясняя данные различия между методами, исследователи предполагают два основных механизма. Первый механизм связан с восприятием информации, заполненной на компьютере, как имеющей большую степень конфиденциальности по сравнению с тем, когда анкета заполняется на бумаге и есть риск, что третьи лица увидят ответы на чувствительные вопросы. Вторым механизмом связан со спецификой когнитивной обработки информации, когда она представлена на экране или когда она представлена на бумаге. Ряд исследований показали, что студенты лучше считывают и понимают информацию на бумаге, чем на экране [7; 26]. Вместе с тем есть также различия в оценке своих способностей в зависимости от того, выполняется тест на экране или на бумаге. Так, если студенты выполняли тест на экране, то они были более уверены в успешной сдаче теста, чем студенты, выполнявшие его на бумаге [7]. Подобным же образом онлайн-гемблеры более высоко оценивают свои навыки, если играют онлайн, по сравнению с теми, кто играет офлайн [25]. Другой пример когнитивных искажений — различия в восприятии рисков. В случае когнитивной обработки информации с экрана риски оцениваются как менее высокие и человек готов к более рискованному поведению по сравнению с тем, когда информация считывается с бумаги [28]. Таким образом, предполагается, что при заполнении анкеты на компьютере эффект социального одобрения будет ниже по сравнению с тем, когда анкета заполняется на бумаге, однако это относится к самостоятельному заполнению анкеты.

Что происходит, когда с компьютером или бумагой *взаимодействует интервьюер, а не сам респондент*? В нашем эксперименте мы обнаружили почти нулевую величину эффекта в 5 индикаторах из 9. Однако в 4 индикаторах были обнаружены значения величины эффекта (d Коэна), варьирующиеся от 0,08 до 0,25. При этом эффект социального одобрения был выше при использовании метода CAPI, чем метода PAPI. Так, хотя в вопросе

о потреблении алкогольных напитков статистически значимых различий не обнаружено, величина эффекта d оказалась равной 0,08, что демонстрирует небольшое уменьшение в данных CAPI доли тех, кто отметил, что потреблял алкогольные напитки в течение последних 30 дней. Более высокие величины эффекта обнаружены в вопросах с интервальной шкалой. Так, среднесуточное потребление сигарет/папирос/трубок было на 1,5 меньше по данным CAPI, чем по данным PAPI. Величина эффекта $d = 0,17$ указывает на то, что среднее значение в опросе методом PAPI соответствует 57-му процентилю значения в опросе CAPI. Два других индикатора — количество занятий физической активностью за последние 12 месяцев и общее время занятий в часах — показывают больший эффект социального одобрения в данных, полученных методом CAPI, по сравнению с данными PAPI. При этом отметим, что различий в количествах форм физической активности не было выявлено. Значит, респонденты указывали большее количество занятий и большее количество часов. При этом d достигла значения 0,25 в индикаторе общего времени занятий в часах, что указывает на то, что среднее значение в опросе по методике PAPI соответствует 60-му процентилю значения в опросе по технологии CAPI.

Важно отметить, что, хотя в 26-й волне RLMS—HSE также был обнаружен более высокий эффект социального одобрения в опросе методом CAPI по сравнению с опросом PAPI, различия были в других индикаторах, а именно в объеме среднесуточного потребления чистого алкоголя и в количестве форм физической активности. Таким образом, мы не видим систематических различий в одних и тех же вопросах между волнами, однако тенденция к более высокому эффекту социального одобрения в опросах методом CAPI остается.

Как можно объяснить подобные различия? Предыдущие эксперименты показали, что в случае опроса, проводимого интервьюером на компьютере, может быть как меньший эффект социального одобрения [8], так и больший [29] по сравнению с опросом, проводимым интервьюером на бумаге. Однако вместе с тем различий может и не быть [23], или в разных вопросах эффект социального одобрения может быть как выше, так и ниже в компьютерном опросе по сравнению с «бумажным» [32]. Кроме того, экспериментальные исследования показывают наличие различий между опросами CAPI и PAPI по ряду включенных в них вопросов, что не всегда возможно связать именно с эффектом социального одобрения [8; 12]. Так, около 14–17% вопросов анкеты показывают статистически значимые различия между данными CAPI и PAPI [8; 24]. Таким образом, эксперименты приводят к разным результатам, в связи с чем однозначного вывода о различии в эффекте социального одобрения между двумя типами опросов пока не может быть сделано, тем более в условиях отсутствия метааналитических исследований. В нашем эксперименте мы обнаружили, что эффект социального одобрения был выше в опросе на компьютере, нежели на бумаге. Как можно объяснить подобные результаты? На наш взгляд, есть три возможных объяснения.

Первое — различие в оценке конфиденциальности данных, когда респонденты отвечают интервьюеру, фиксирующему ответы на бумаге или фиксирующему ответы на планшете. Так, все больше россиян испытывают беспокойство по поводу конфиденциальности персональных данных в Интернете: по данным ФОМ, если в 2013 г. таких было 24%, то в 2017 г. — 35%¹¹. Таким образом, при наличии у интервьюера планшета опасения респондента, что данные могут утечь или быть переданы третьим лицам, могут усилиться, это приводит к увеличению эффекта социальной желательности в опросе на планшете по сравнению с опросом на бумаге.

Второе возможное объяснение — различие в когнитивной сложности для интервьюера, заполняющего анкету на планшете или на бумаге. Сюда включены различия в затрачиваемых им умственных усилиях и когнитивных ресурсах, включая моторный контроль, зрительное внимание и т. д. [2]. Так, например, заполнение анкеты на планшете занимало в среднем на 10% больше времени, чем на бумаге. На основе собранных данных среди интервьюеров, работающих в экспериментальном исследовании RLMS—HSE, Д.В. Лебедев и М.Б. Богданов [3] делают выводы, что процесс работы методом СAPI по сравнению с опросом на бумаге был для них более сложным по нескольким причинам. Во-первых, в связи с увеличением времени заполнения анкеты на планшете у интервьюеров снижаются внимание и концентрация. Во-вторых, повышение нагрузки было вызвано увеличением количества непривычных для интервьюера действий, необходимых для заполнения анкеты на планшете, особенно в открытых вопросах, где ему необходимо было самому вводить данные в текстовое поле, а не выбирать ответ из уже заданного списка. В-третьих, переживания интервьюеров были связаны с их неуверенностью в умении пользоваться техникой. Поскольку средний возраст интервьюеров 55 лет и многие из них оценили себя как недостаточно уверенных пользователей компьютеров, они испытывали дополнительный стресс по поводу того, что могут совершить не те действия, стереть полностью собранную информацию, куда-то перескочить и т. д. В-четвертых, экран планшета достаточно небольшой (диагональ экрана 8 дюймов) по сравнению с более привычным для интервьюеров форматом бумаги А4, что усложняло процесс фиксации ответов. В-пятых, в случае бумажной анкеты респонденты и интервьюеры обладают большей свободой при ее заполнении по сравнению с заполнением анкеты на планшете. Например, какие-то блоки вопросов могли заполняться респондентами самостоятельно на бумаге, или в случае затруднений интервьюер мог его пропустить и вернуться к нему позже, что невозможно в случае СAPI [1; 3]. Кроме того, опрос интервьюеров после эксперимента показал, что большинство из них (88%) столкнулись с техническими проблемами при заполнении анкеты, когда не работал тачпад на планшете, приходилось долго ждать переход на другую страницу или не срабатывала система контроля GPS-координат. Учитывая перечисленные сложности,

¹¹ См. на сайте ФОМ: Безопасность в Интернете. Какие меры безопасности принимают в интернете пользователи? URL: <https://fom.ru/Bezopasnost-i-pravo/13465> (дата обращения 05.08.2020).

мы предполагаем, что при работе методом CAPI могла увеличиться ошибка измерения по сравнению с опросом на бумаге, тем более в открытых вопросах, когда интервьюеры должны были вводить, например, количество занятий физической активностью или объем выпитого алкоголя. Отметим, что различия были обнаружены именно в индикаторах, основанных на открытых вопросах, что, по мнению интервьюеров, было сложнее заполнять на планшете, чем на бумаге.

Третье возможное объяснение — увеличение эффекта социального одобрения в случае заполнения анкеты на планшете. Так, исследования в развивающихся странах показали, что низкостатусные группы завышают свои доходы в опросах методом CAPI по сравнению с «бумажными» [11]. Можно предположить, что в случае заполнения анкеты интервьюером на планшете в нашем случае эффект социального одобрения был найден в вопросе о среднесуточном потреблении сигарет/папирос/трубок, а также о количестве и часах занятий физической активностью.

Заключение

На основе результатов экспериментального исследования в рамках 27-й волны лонгитюдного обследования домохозяйств RLMS—HSE в 2018—2019 гг. мы проверили две гипотезы о сравнении качества данных, полученных методами CAPI и PAPI. Первая гипотеза касалась доли несодержательных ответов («затрудняюсь ответить» и «отказ от ответа»). Вторая гипотеза — различий в ответах на чувствительные вопросы. Наши результаты показали, что в опросе на планшете была ниже ошибка неотчетов на уровне определенных вопросов (меньшая доля несодержательных ответов), но более высокая ошибка измерения в чувствительных вопросах. Таким образом, *подтвердились H1 и H2** (альтернативная H2 гипотеза).

Необходимо отметить, что результаты исследования имеют ряд ограничений. Во-первых, мы проверили только ограниченное количество индикаторов, что не всегда дает полное представление о различии в качестве данных, полученных методами CAPI и PAPI. Во-вторых, в экспериментальном исследовании было неравное распределение в объемах выборки между двумя типами опросов: так, лишь только пятая часть респондентов была опрошена методом CAPI, а остальные — методом PAPI. Неравное распределение между контрольной и экспериментальной группами может увеличить ошибку первого рода, когда мы ложно отвергаем нулевую гипотезу [20]. В-третьих, наблюдались различия в некоторых социально-демографических характеристиках между выборками в опросах CAPI и PAPI. Возможно, это связано со спецификой экспериментального дизайна или со спецификой выполнения инструкций интервьюерами. В-четвертых, в экспериментальное исследование были включены определенные населенные пункты, и мы не можем пока обобщить эти результаты на все локации, в которых проводится обследование домохозяйств RLMS—HSE. В-пятых, в экспериментальное исследование была включена только индивидуальная анкета для взрослых. В последующих экспериментах необходимо сравнить качество данных, полученных в двух типах опросов, по анкете для домохозяйств.

Несмотря на ограничения, данная статья, насколько нам известно, является первой в отечественной литературе попыткой сравнения качества данных, полученных методами CAPI и PAPI в лонгитюдных опросах, где анкету заполняет интервьюер. В статье мы используем уникальные данные экспериментального исследования, проведенного в рамках единственного российского лонгитюдного исследования домохозяйств. Дальнейшими задачами RLMS–HSE могут быть перевод опросов на метод CAPI в других населенных пунктах, включенных в исследование; программирование анкеты для домохозяйств и поиск способов построения индивидуализированной версии анкеты для взрослых членов домохозяйств, когда используются ответы респондентов из предыдущей волны опроса.

ЛИТЕРАТУРА

1. *Богданов М.Б., Лебедев Д.В.* «Главное не бояться!» — воздействие тренингов на ожидания интервьюеров при переходе от PAPI к CAPI // Социология: 4М. 2018. № 45. С. 103–132.
2. *Девятко И.Ф.* Разработка подхода к количественной мультимодальной оценке когнитивной нагрузки интервьюеров: результаты пилотного квазиэксперимента // Вестник Российского университета дружбы народов. Серия: Социология. 2018. Т. 18. № 4. С. 627–637. DOI: 10.22363/2313-2272-2018-18-4-627-637
3. *Лебедев Д.В., Богданов М.Б.* Переход с PAPI на CAPI: опыт интервьюеров и характеристики, влияющие на их ожидания // Мониторинг общественного мнения: Экономические и социальные перемены. 2019. № 4. С. 43–67. DOI: 10.14515/monitoring.2019.4.03
4. *Малошюнок Н.Г., Терентьев Е.А.* Эффект визуализации и формата оценочных шкал в онлайн-опросах студентов // Universitas. Журнал о жизни университетов. 2013. № 1 (2). С. 20–44.
5. *Мягков А.Ю., Журавлева С.Л.* Экспериментальная оценка качества данных в телефонном интервью // Социология: 4М. 2011. № 32. С. 26–52.
6. *Терентьев Е.А., Мавлетова А.М., Косолапов М.С.* Интервьюирование с помощью компьютерных технологий в лонгитюдных обследованиях домохозяйств // Мониторинг общественного мнения: Экономические и социальные перемены. 2018. № 3. С. 47–64. DOI: 10.14515/monitoring.2018.3.03
7. *Ackerman R., Goldsmith M.* Metacognitive Regulation of Text Learning: On Screen Versus on Paper // Journal of Experimental Psychology: Applied. 2011. Vol. 17. P. 18–32. DOI: 10.1037/a0022086
8. *Baker R.P., Bradburn N.M., Johnson R.A.* Computer-assisted personal interviewing: An experimental evaluation of data quality and costs // Journal of Official Statistics. 1995. Vol. 11. No. 4. P. 413–431.
9. *Banks R., Laurie H.* From PAPI to CAPI: The Case of the British Household Panel Survey // Social Science Computer Review. 2000. Vol. 18. No. 4. P. 397–406. DOI: 10.1177/089443930001800403
10. *Bishop G., Oldendick R., Tuchfarber A.* Effects of Filter Questions in Public Opinion Surveys // Public Opinion Quarterly. 1980. Vol. 47. P. 528–546. DOI: 10.1086/268810

11. *Bush S.S., Prather L.* Do Electronic Devices in Face-To-Face Interviews Change Survey Behavior? Evidence from a Developing Country // *Research and Politics*. 2019. Vol. 6. No. 2. P. 1–7. DOI: 10.1177/2053168019844645
12. *Caeyers B., Chalmers N., Weerdt J.D.* Improving consumption measurement and other survey data through CAPI: Evidence from a randomized experiment // *Journal of Development Economics*. 2012. Vol. 98. No. 1. P. 19–33. DOI: 10.1016/j.jdeveco.2011.12.001
13. *Caviglia-Harris J., Hall S., Mullan K., Macintyre C., Bauch S.C., Harris D., Cha H.* Improving household surveys through computer-assisted data collection: Use of touch-screen laptops in challenging environments // *Field Methods*. 2012. Vol. 24. No. 1. P. 74–94. DOI: 10.1177/1525822x11399704
14. *Chen H., Cohen P., Chen S.* How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies // *Communications in Statistics — Simulation and Computation*. 2010. Vol. 39. No. 4. P. 860–864. DOI: 10.1080/03610911003650383
15. *Cohen J.* *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers, 1988. — 567 p.
16. *Cohen J.* A Power Primer // *Psychological Bulletin*. 1992. Vol. 112. P. 155–159.
17. *De Leeuw E.D., Hox J.J., Snijders G.* The effect of computer-assisted interviewing on data quality. A review // *Market Research Society Journal*. 1995. Vol. 37. No. 4. P. 1–19. DOI: 10.1177/147078539503700401
18. *De Leeuw E., Hox J., Boevé A.* Handling Do-Not-Know Answers Exploring New Approaches in Online and Mixed-Mode Surveys // *Social Science Computer Review*. 2015. Vol. 34. No. 1. P. 116–132. DOI: 10.1177/0894439315573744
19. *DeRouvray C., Couper M.* Designing a Strategy for Reducing “No Opinion” Responses in Web-based Surveys // *Social Science Computer Review*. 2002. Vol. 20. P. 3–9. DOI: 10.1177/089443930202000101
20. *Dumville J.C., Hahn S., Miles J.N.V., Torgerson D.J.* The use of unequal randomisation ratios in clinical trials: A review // *Contemporary Clinical Trials*. 2006. Vol. 27. No. 1. P. 1–12. DOI: 10.1016/j.cct.2005.08.003
21. *Gnambs T., Kaspar K.* Disclosure of sensitive behaviors across self-administered survey modes: A meta-analysis // *Behavior Research Methods*. 2015. Vol. 47. P. 1237–1259. DOI: 10.3758/s13428-014-0533-4
22. *Heiskanen M., Ahlqvist K.* Experiences of Two CAPI-PAPI Comparisons in Finland // *IBUC 97*. 1997. P. 95–146.
23. *Laurie H.* From PAPI to CAPI: consequences for data quality on the British Household Panel Study // *Working Papers of the Institute for Social and Economic Research*. 2003. No. 14. P. 1–14. [online]. Accessed 17.11.2020. URL: <https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2003-14.pdf>
24. *Lynn P.* Data collection mode effects on responses to attitudinal questions // *Journal of Official Statistics*. 1998. Vol. 14. No. 2. P. 1–14.
25. *MacKay T.L., Bard N., Bowling M., Hodgins D.C.* Do Pokers Players Know How Good They Are? Accuracy of Poker Skill Estimation in Online and Offline Players // *Computers in Human Behavior*. 2014. Vol. 31. P. 419–424. DOI: 10.1016/j.chb.2013.11.006
26. *Mangen A., Walgermo B.R., Brønnick K.* Reading Linear Texts on Paper versus Computer Screen: Effects on Reading Comprehension // *International Journal of Educational Research*. 2013. Vol. 58. P. 61–68. DOI: 10.1016/j.ijer.2012.12.002

27. *Martin C.L., O’Muircheartaigh C., Curtice J.* The use of CAPI for attitudes surveys: an experimental comparison with traditional methods // *Journal of Official Statistics*. 1993. Vol. 9. P. 356–369.
28. *Oeberst A., Haberstro S., Gnambs T.* Not Really the Same: Computerized and Real Lotteries in Decision Making Research // *Computers in Human Behavior*. 2015. Vol. 44. P. 250–257. DOI: 10.1016/j.chb.2014.10.060
29. *Schräpler J.-P., Schupp J., Wagner G.G.* Changing from PAPI to CAPI: Introducing CAPI in a longitudinal Study // *Journal of Official Statistics*. 2010. No. 2 (26). P. 233–269.
30. *Schuman H., Scott J.* Response Effects Over Time: Two Experiments // *Sociological Methods and Research*. 1989. No. 17. P. 398–408. DOI: 10.1177/0049124189017004005
31. *Tourangeau R., Couper M., Conrad F.* Spacing, Position, and Order Interpretive Heuristics for Visual Features Of Survey Questions // *Public Opinion Quarterly*. 2004. Vol. 68. No. 3. P. 368–393. DOI: 10.1093/poq/nfh035
32. *Watson N., Wilkins R.* Experimental change from paper-based interviewing to computer-assisted interviewing in the HILDA Survey // *Melbourne Institute Working Paper Series*. 2012. No. 6 (12) [online]. Accessed 17.11.2020. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2018624 DOI: 10.2139/ssrn.2018624

Дата поступления: 27.05.2020.

SOTSIOLOGICHESKIY ZHURNAL = SOCIOLOGICAL JOURNAL. 2020.
VOL. 26. NO. 4. P. 8–30. DOI: 10.19181/socjour.2020.26.4.7640

A.M. KLIMOVA, E.A. TERENCEV

National Research University Higher School of Economics,
Moscow, Russian Federation.

Aigul M. Klimova — Candidate of Sociological Sciences, Associate Professor, Department of Sociology; Senior Research Fellow, Laboratory for Comparative Social Research, National Research University Higher School of Economics.

Address: room 3-334, bl. 3, Krivokolenny Lane, 101000, Moscow, Russian Federation.
Phone: +7 (495) 772-95-90 * 12244. **Email:** amavletova@hse.ru

Evgeniy A. Terentev — Candidate of Sociological Sciences, Senior Research Fellow, Centre of Sociology of Higher Education, Institute of Education, National Research University Higher School of Economics. **Address:** room 20716, bl. 10, 16, Potapovsky Lane, 101000, Moscow, Russian Federation.

Phone: +7 (985) 386-63-49. **Email:** eterentev@hse.ru

**COMPARISON OF DATA QUALITY BETWEEN PAPI AND CAPI MODES
OF DATA COLLECTION IN LONGITUDINAL HOUSEHOLD SURVEYS:
RLMS–HSE EXPERIMENT RESULTS**

Abstract. This article presents the results of an experimental study on how the transition from PAPI to CAPI modes affected data quality in longitudinal household surveys. The study was conducted in 2018–2019 within the Russian Longitudinal Monitoring Survey (RLMS–HSE). In the previous paper, which was based on data from the 26th wave of the RLMS HSE, it was shown that the use of CAPI leads to a significant decrease in the rate

of non-substantive responses (“Don’t know”), as well as significant differences in sensitive questions. This paper was aimed at verifying these findings using new data collected during the 27th wave of the RLMS–HSE. The results show that the use of CAPI leads to a decrease in the rate of non-substantive responses, which helps to improve data quality. However, it was shown that the use of CAPI could lead to an increase in social desirability bias.

Keywords: PAPI; CAPI; data quality; panel household survey; RLMS–HSE; rate of nonsubstantive responses; sensitive questions; social desirability effect.

For citation: Klimova A.M., Terentev E.A. Comparison of Data Quality between PAPI and CAPI Modes of Data Collection in Longitudinal Household Surveys: RLMS–HSE Experiment Results. *Sotsiologicheskii Zhurnal = Sociological Journal*. 2020. Vol. 26. No. 4. P. 8–30. DOI: 10.19181/socjour.2020.26.4.7640

Acknowledgment. This work is supported by the Russian Science Foundation under grant No. 17-78-20172.

REFERENCES

1. Bogdanov M.B., Lebedev D.V. “Do not worry!” — training’s impact on interviewers’ expectations towards the change from PAPI to CAPI. *Sotsiologiya: 4M*. 2018. No. 45. P. 103–132. (In Russ.)
2. Devyatko I.F. Developing an approach to multimodal quantitative assessment of interviewers’ cognitive load: first results of a field quasi experiment. *Vestnik Rossiiskogo universiteta družby narodov. Seriya: Sotsiologiya*. 2018. No. 18 (4). P. 627–637. DOI: 10.22363/2313-2272-2018-18-4-627-637 (In Russ.)
3. Lebedev D.V., Bogdanov M.B. Transition from PAPI to CAPI: interviewers’ experience and factors influencing their expectations. *Monitoring of Public Opinion: Economic and Social Changes*. 2019. No. 4. P. 43–67. DOI: 10.14515/monitoring.2019.4.03 (In Russ.)
4. Maloshonok N.G., Terent’ev E.A. Effects of visualization and assessment scales’ format in online student surveys. *Universitas. Zhurnal o zhizni universitetov*. 2013. No. 1 (2). P. 20–44. (In Russ.)
5. Myagkov A.U., Zhuravleva S.L. Experimental evaluation of data quality in telephone interview. *Sotsiologiya: 4M*. 2011. No. 32. P. 26–52. (In Russ.)
6. Terentev E.A., Mavletova A.M., Kosolapov M.S. Computer-assisted personal interviewing for longitudinal household studies. *Monitoring of Public Opinion: Economic and Social Changes*. 2019. No. 3. P. 47–64. DOI: 10.14515/monitoring.2018.3.03 (In Russ.)
7. Ackerman R., Goldsmith M. Metacognitive Regulation of Text Learning: On Screen Versus on Paper. *Journal of Experimental Psychology: Applied*. 2011. Vol. 17. P. 18–32. DOI: 10.1037/a0022086
8. Baker R.P., Bradburn N.M., Johnson R.A. Computer-assisted personal interviewing: An experimental evaluation of data quality and costs. *Journal of Official Statistics*. 1995. Vol. 11. No. 4. P. 413–431.
9. Banks R., Laurie H. From PAPI to CAPI: The Case of the British Household Panel Survey. *Social Science Computer Review*. 2000. Vol. 18. No. 4. P. 397–406. DOI: 10.1177/089443930001800403
10. Bishop G., Oldendick R., Tuchfarber A. Effects of Filter Questions in Public Opinion Surveys. *Public Opinion Quarterly*. 1980. Vol. 47. P. 528–546. DOI: 10.1086/268810
11. Bush S.S., Prather L. Do Electronic Devices in Face-To-Face Interviews Change Survey Behavior? Evidence from a Developing Country. *Research and Politics*. 2019. Vol. 6. No. 2. P. 1–7. DOI: 10.1177/2053168019844645
12. Caeyers B., Chalmers N., Weerdt J.D. Improving consumption measurement and other survey data through CAPI: Evidence from a randomized experiment. *Journal of Development Economics*. 2012. Vol. 98. No. 1. P. 19–33. DOI: 10.1016/j.jdeveco.2011.12.001
13. Caviglia-Harris J., Hall S., Mulllan K., Macintyre C., Bauch S.C., Harris D., Cha H. Improving household surveys through computer-assisted data collection: Use of

- touch-screen laptops in challenging environments. *Field Methods*. 2012. Vol. 24. No. 1. P. 74–94. DOI: 10.1177/1525822x11399704
14. Chen H., Cohen P., Chen S. How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Communications in Statistics — Simulation and Computation*. 2010. Vol. 39. No. 4. P. 860–864. DOI: 10.1080/03610911003650383
 15. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers, 1988. 567 p.
 16. Cohen J. A Power Primer. *Psychological Bulletin*. 1992. Vol. 112. P. 155–159.
 17. De Leeuw E.D., Hox J.J., Snijders G. The effect of computer-assisted interviewing on data quality. A review. *Market Research Society Journal*. 1995. Vol. 37. No. 4. P. 1–19. DOI: 10.1177/147078539503700401
 18. De Leeuw E., Hox J., Boevé A. Handling Do-Not-Know Answers Exploring New Approaches in Online and Mixed-Mode Surveys. *Social Science Computer Review*. 2015. Vol. 34. No. 1. P. 116–132. DOI: 10.1177/0894439315573744
 19. DeRouvray C., Couper M. Designing a Strategy for Reducing “No Opinion” Responses in Web-based Surveys. *Social Science Computer Review*. 2002. Vol. 20. P. 3–9. DOI: 10.1177/089443930202000101
 20. Dumville J.C., Hahn S., Miles J.N.V., Torgerson D.J. The use of unequal randomisation ratios in clinical trials: A review. *Contemporary Clinical Trials*. 2006. Vol. 27. No. 1. P. 1–12. DOI: 10.1016/j.cct.2005.08.003
 21. Gnamb T., Kaspar K. Disclosure of sensitive behaviors across self-administered survey modes: A meta-analysis. *Behavior Research Methods*. 2015. Vol. 47. P. 1237–1259. DOI: 10.3758/s13428-014-0533-4
 22. Heiskanen M., Ahlqvist K. Experiences of Two CAPI-PAPI Comparisons in Finland. *IBUC 97*. 1997. P. 95–146.
 23. Laurie H. From PAPI to CAPI: consequences for data quality on the British Household Panel Study. *Working Papers of the Institute for Social and Economic Research*. 2003. No. 14. P. 1–14. Accessed 17.11.2020. URL: <https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2003-14.pdf>
 24. Lynn P. Data collection mode effects on responses to attitudinal questions. *Journal of Official Statistics*. 1998. Vol. 14. No. 2. P. 1–14.
 25. MacKay T.L., Bard N., Bowling M., Hodgins D.C. Do Pokers Players Know How Good They Are? Accuracy of Poker Skill Estimation in Online and Offline Players. *Computers in Human Behavior*. 2014. Vol. 31. P. 419–424. DOI: 10.1016/j.chb.2013.11.006
 26. Mangen A., Walgermo B.R., Brønnick K. Reading Linear Texts on Paper versus Computer Screen: Effects on Reading Comprehension. *International Journal of Educational Research*. 2013. Vol. 58. P. 61–68. DOI: 10.1016/j.ijer.2012.12.002
 27. Martin C.L., O’Muircheartaigh C., Curtice J. The use of CAPI for attitudes surveys: an experimental comparison with traditional methods. *Journal of Official Statistics*. 1993. Vol. 9. P. 356–369.
 28. Oeberst A., Haberstro S., Gnamb T. Not Really the Same: Computerized and Real Lotteries in Decision Making Research. *Computers in Human Behavior*. 2015. Vol. 44. P. 250–257. DOI: 10.1016/j.chb.2014.10.060
 29. Schräpler J.-P., Schupp J., Wagner G.G. Changing from PAPI to CAPI: Introducing CAPI in a longitudinal Study. *Journal of Official Statistics*. 2010. No. 2 (26). P. 233–269.
 30. Schuman H., Scott J. Response Effects Over Time: Two Experiments. *Sociological Methods and Research*. 1989. No. 17. P. 398–408. DOI: 10.1177/0049124189017004005
 31. Tourangeau R., Couper M., Conrad F. Spacing, Position, and Order Interpretive Heuristics for Visual Features Of Survey Questions. *Public Opinion Quarterly*. 2004. Vol. 68. No. 3. P. 368–393. DOI: 10.1093/poq/nfh035
 32. Watson N., Wilkins R. Experimental change from paper-based interviewing to computer-assisted interviewing in the HILDA Survey. *Melbourne Institute Working Paper Series*. 2012. No. 6 (12). Accessed 17.11.2020. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2018624 DOI: 10.2139/ssrn.2018624