



**НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»**

НАУЧНЫЙ ДОКЛАД

**по результатам подготовленной
научно-квалификационной работы (диссертации)**

ФИО Сапонова Анастасия Владимировна

Направление подготовки 39.06.01 Социологические науки

Профиль (направленность) программы 5.4.1. Теория, методология и история социологии

Аспирантская школа по социологическим наукам

Аспирант _____ /Сапонова А.В./
подпись

Научный руководитель _____ /Девятко И.Ф./
подпись

Академический директор
Аспирантской школы _____ /Савинская О.Б./
подпись

Москва, 2024

Аннотация

Настоящий научный доклад представляет результаты диссертационного исследования по теме: «Интеграция опросных данных и цифровых следов в социологических исследованиях: возможности и ограничения». Актуальность исследования определяется растущим числом доступных для исследователей данных цифровой природы – цифровых следов, параданных, данных трекеров (шагомеров, фитнес-браслетов, трекеров цифровой активности). Вместе с тем, процедуры и стандарты оценки качества таких данных, а также способы интеграции с другими типами социологических данных, ещё находятся в разработке. Цель диссертационного исследования заключатся в систематизации теоретических оснований интеграции опросных данных и цифровых следов, а также определении возможностей и ограничений этой интеграции на основе ряда методологических квази-экспериментов. В качестве теоретической основы исследования выступает концепция «нереактивного» измерения, а также представления о социологическом измерении и критериях оценки его качества. В методологической части исследования апробируется подход «пожертвование данных» (*data donation*), в рамках которого участники исследования делятся данными, которые уже собрали их устройства. Предусмотрена установка специальных трекеров со стороны исследователя на устройства участников исследования в таком случае не требуется. В рамках квази-эксперимента была измерена конкурентная валидность ретроспективного самоотчёта об использовании смартфона. Результаты исследования показали умеренный уровень конкурентной валидности самоотчёта о мобильном использовании, что позволяет говорить о ретроспективном самоотчёте как о сравнительно валидном методе измерения практик использования смартфона в пределах недели. Однако, при прочих равных, цифровые следы представляются наиболее предпочтительными для оценки объема использования смартфона по следующим критериям: доступности данных для сбора, отсутствия существенных дополнительных издержек, оптимизации опросного инструментария, снижения когнитивной нагрузки на респондентов.

Abstract

The current report presents the findings of a dissertation study on the topic of “Integration of Survey Data and Digital Traces in Sociological Research: Opportunities and Limitations”. The significance of this study is determined by the increasing availability of digital data available to researchers, such as digital footprints, metadata, and trackers (e.g. pedometers, fitness trackers, digital activity trackers). At the same time procedures and standards for evaluating the quality of these data, as well as ways to integrate with other types of sociological data, are still under development. The aim of this study was to systematically explore the theoretical foundations of the integration of survey data and digital traces, as well as to determine the possibilities and limitations of this integration based on a number of methodological quasi-experiments. The theoretical basic of this study is the concept of “unobtrusive measures” and the sociological measurement concept, as well as and criteria for evaluating its quality. To conduct the study, we used a quasi-experimental design with the “data donation” approach. This approach involves participants sharing data that their devices have collected. This approach does not require researcher to pre-install trackers on the participants’ devices. As part of the quasi-experiment, we measured the concurrent validity of a retrospective self-report on smartphone usage. The results showed a moderate level of concurrent validity of self-reporting, indicating that retrospective self-report on mobile use can be a relatively valid method for measuring smartphone usage patterns within a week. However, other things being equal, it seems that digital footprints are be more preferable for assessing the volume of smartphone usage according to the following criteria: the availability of data for collection, the absence of significant additional costs, the optimization of survey tools, and the reduction of respondent burden.

Актуальность исследования

Цифровизация и медиатизация способствовали существенному увеличению числа данных, доступных для использования в социологическом исследовании. «Новые» цифровые данные (цифровые следы, параданные, данные трекеров различных видов активности – от физической до цифровой) представляют для социологов интерес, однако процедуры и стандарты оценки их качества, а также способы интеграции с другими социологическими данными, ещё предстоит разработать.

Интеграция данных как процедура в социологическом исследовании представляет собой обширный комплекс методологических задач, среди которых: связывание разных типов данных вероятностными или же детерминированными способами, изучение мотивации респондентов делиться персональной информацией, кросс-валидация параметров, измеренных на разных типах данных, разработка этических норм, связанных с исследованием цифровых следов на индивидуальном уровне и т.д.

Как правило, выделяют три типа данных, пригодных для обогащения и интеграции с данными опросов – параданные, административные данные и так называемые контекстные данные, под последними, как правило, понимают обширный спектр условно неактивных (малореактивных) данных¹ от показателей физической активности и цифровых следов до более точных данных о характеристиках района жительства респондента, или же месте его работы.

С одной стороны интеграция и связывание данных может использоваться в целях валидации и повышения уровня надежности измерения, с другой – она позволяет оптимизировать инструментарий исследования, например, методику опросов (снижать нагрузку на респондента, заполняя некоторые ответы исходя из неактивных данных).

¹ Девятко И.Ф. Новые данные, новая статистика: от кризиса воспроизводимости к новым требованиям к анализу и представлению данных в социальных науках // Социологические исследования. 2018. № 12. С. 30–38.

Исследовательская проблема

Проблема исследования может быть рассмотрена с теоретической и методологической позиций. Во-первых, несмотря на обилие методологических и эмпирических попыток осуществить операцию интеграции на концептуальном уровне остаётся неясным, в какой общей теоретической рамке реактивные и неактивные данные могут быть объединены и проанализированы, в чем заключается их концептуальное сходство и различие. Во-вторых, в исследовательской практике ещё не сформировалось устойчивых методологических программ по интеграции (и в частности кросс-валидации) реактивных и неактивных данных.

Настоящее диссертационное исследование фокусируется на изучении перспектив интеграции реактивных и неактивных данных на примере сравнения опросных данных, основанных на ретроспективных самоотчётах респондентов, и их цифровых следах.

Объект исследования – процесс интеграции данных опросов и цифровых следов.

Предмет – методологические проблемы и возможности интеграции данных опросов и цифровых следов.

Цель исследования заключаются в систематизации теоретических оснований интеграции опросных данных и цифровых следов, а также определении возможностей и ограничений этой интеграции на основе ряда методологических квази-экспериментов.

Задачи исследования:

- 1) Концептуализировать процесс интеграции реактивных и неактивных данных как методологическую проблему;
- 2) Систематизировать корпус работ в области интеграции указанных типов данных;
- 3) Расширить существующую типологию подходов к связыванию (интеграции) реактивных и неактивных данных;

- 4) Разработать инструментарий для оценки валидности ретроспективного самоотчёта;
- 5) Провести апробацию квази-эксперимента с использованием методического подхода «пожертвование данных»;
- 6) Оценить валидность самоотчётов об использовании смартфона с использованием цифровых следов, полученных в результате «пожертвования данных» респондентами.

Степень разработанности темы исследования

Идея совмещения разных исследовательских методов и техник с целью проверки надёжности, валидности или же их взаимной интеграции и обогащения в самом общем виде уже ставилась и ранее, например, в классических работах Кэмпбелла Д.Т., Фиске Д.В.², Бушара Т.Д. (младший)³, Зеллера Р.А., Кармина Е.Г.⁴, и возможно является одной из основных в области методологии социальных наук. Несмотря на некоторую концептуальную разработанность этой идеи, её частные методологические особенности представляют интерес для изучения. Вместе с тем, оценка валидности применимых инструментов измерения (в частности конструктивной и критериальной) не является широко распространённой практикой в социологических исследованиях⁵, однако при интеграции данных потребность в оценке валидности имеет особое значение.

Несмотря на относительную новизну направления интеграции данных, в литературе уже предпринимаются попытки систематизировать возможные методологические подходы в рамках интеграции реактивных и нереактивных данных. Так, Себастьян Штир⁶ и соавторы предлагают двумерную классификацию, где связывание может являться либо частью дизайна

² Campbell D. T., Fiske D. W. Convergent and discriminant validation by the multitrait-multimethod matrix // Psychological bulletin. 1959. Vol. 56. №. 2. С. 81–105.

³ Bouchard Jr T. J. Field research methods: Interviewing, questionnaires, participant observation, systematic observation, unobtrusive measures // Handbook of industrial and organizational psychology / ed. M. D. Dunnette. Chicago: Rand McNally. 1976. Vol. 1. P. 363–413.

⁴ Zeller R.A., Carmines E.G. Measurement in the social sciences: the link between theory and data. Cambridge; New York: Cambridge University Press, 1980. 198 p.

⁵ Девятко И. Ф. Социологическая теория: старые трудности, новые вызовы // Социологические исследования. 2021. № 10. С. 3–11.

⁶ Stier S. et al. Integrating Survey Data and Digital Trace Data: Key Issues in Developing an Emerging Field // Social Science Computer Review. 2020. Vol. 38, № 5. P. 503–516.

исследования с самого начала и данные для интеграции собираются исходя из установленных задач («Ex Ante» – «До»), либо связывание происходит постфактум (и сбор происходил независимо друг от друга) на основе собранных данных («Ex Post» – «После»). С другой стороны, связывание этими двумя способами может происходить на уровне агрегированном, индивидуальном, а также – уровней организаций (публичных акторов)⁷.

Среди работ по связыванию данных на *индивидуальном уровне* можно выделить несколько направлений. Первое (довольно узкое) продолжает долгую традицию эмпирико-функционалистской школы, связанную с интеграцией данных опросов и медиапотребления с целью измерения эффекта медиавоздействия, можно отметить работы Яновицкого И. и Капелла Й.⁸, Де Вриза К. и Семетко Х.⁹, Ван Спанье Й. и Де Вриза К.¹⁰. Фактически предпринимаются попытки фиксации на микро-уровне воздействия на респондентов экзогенных факторов в виде источников информации.

Другое направление представлено в работах Врага Е.К. и Талли М.¹¹, Хеншена Х.¹², Юргенса П. и соавторов¹³, Хоппа Т.¹⁴ и связано с изучением эффекта памяти и/или сравнения данных о припоминании респондентов своих действий с их цифровыми следами для измерения точности самооценок респондентов (результаты предыдущих исследований говорят преимущественно о низкой степени точности такой самофиксации).

Стоит отметить и корпус работ с *агрегированными данными* разной природы

⁷ Там же.

⁸ Yanovitzky I. Effect of Call-In Political Talk Radio Shows on Their Audiences: Evidence from a Multi-Wave Panel Analysis // International Journal of Public Opinion Research. 2001. Vol. 13, № 4. P. 377–397.

⁹ de Vreese C., Semetko H.A. News matters: Influences on the vote in the Danish 2000 euro referendum campaign // European Journal of Political Research. 2004. Vol. 43, № 5. P. 699–722.

¹⁰ van Spanje J., de Vreese C. Europhile Media and Eurosceptic Voting: Effects of News Media Coverage on Eurosceptic Voting in the 2009 European Parliamentary Elections // Political Communication. 2014. Vol. 31, № 2. P. 325–354.

¹¹ Vraga, E. K., Tully, M. Who is exposed to news? It depends on how you measure: Examining self-reported versus behavioral news exposure measures // Social Science Computer Review. 2020. Vol. 38, №5. P. 550–566.

¹² Haenschen K. Self-Reported Versus Digitally Recorded: Measuring Political Activity on Facebook // Social Science Computer Review. 2020. Vol. 38, № 5. P. 567–583.

¹³ Jürgens P., Stark B., Magin M. Two Half-Truths Make a Whole? On Bias in Self-Reports and Tracking Data // Social Science Computer Review. 2020. Vol. 38, № 5. P. 600–615.

¹⁴ Hopp, T., Vargo, C. J., Dixon, L., Thain, N. Correlating Self-Report and Trace Data Measures of Incivility: A Proof of Concept // Social Science Computer Review. 2020. Vol. 38, № 5. P. 584–599.

в области электоральных исследований, так, например, в работе Пасека Й.¹⁵ и соавторов сравниваются результаты сентимент-анализа твитов и опросов общественного мнения, в работе Хофстра Б.¹⁶ и соавторов данные опросов подростков об их социальном окружении интегрируются с информацией об их друзьях в социальных сетях. В первом случае интеграция используется с целью измерения разных латентных конструкторов, объединенных единым объектом исследования, во втором – анализ нереактивных данных используется в качестве предположительно более валидного метода, т.е. с целью более точного измерения одного и того же латентного конструктора.

Помимо ряда традиционных методологических вопросов (насколько и по каким признакам смещена группа, которая готова добровольно делиться персональной информацией с исследователями), которые возникают в связи с интеграцией данных, отмечают ещё законодательные и этические. Так, отдельно изучается процесс получения согласия респондентов на использование их нереактивных данных, их мотивы согласия и/или несогласия делиться персональными данными – работы Койша Ф. и соавторов,¹⁷ Бенинджера К. и соавторов¹⁸.

В России тема настоящего диссертационного исследования изучена сравнительно слабо, и если попытки отрефлексировать природу цифровых следов и их потенциальное место в социальных науках неоднократно предпринимались, здесь можно отметить работы Богданова М.Б., Смирнова И.Б.¹⁹, Губы К.С.²⁰,

¹⁵ Pasek, J., McClain, C. A., Newport, F., Marken, S. Who's Tweeting About the President? What Big Survey Data Can Tell Us About Digital Traces? // *Social Science Computer Review*. 2020. Vol. 38, № 5. P. 633–650.

¹⁶ Hofstra B. et al. Sources of Segregation in Social Networks: A Novel Approach Using Facebook // *American Sociological Review*. 2017. Vol. 82, № 3. P. 625–656.

¹⁷ Beninger K. et al. Understanding Society: How people decide whether to give consent to link their administrative and survey data / *Kantar Public UK*, 2017. № 2017 – 13. 65 p.

¹⁸ Keusch, F., Struminskaya, B., Antoun, C., Couper, M. P., & Kreuter, F. Willingness to Participate in Passive Mobile Data Collection // *Public Opinion Quarterly*. 2019. Vol. 83. No. S1. P. 210–235.

¹⁹ Богданов М.Б., Смирнов И.Б. Возможности и ограничения цифровых следов и методов машинного обучения в социологии // *Мониторинг общественного мнения: экономические и социальные перемены*. 2021. № 1. С. 304–328.

²⁰ Губа К. С. Большие данные в социологии: Новые данные, новая социология? *Социологическое обозрение*, 2018. № 17(1). С. 213–234.

Девятко И.Ф.²¹, Дудиной В. И.²², Одинцова А.В.²³, Федоровой А.А., Николаенко Г.А.²⁴, то литературы с результатами методических экспериментов по интеграции/сравнения данных разной природы можно отметить сравнительно немного – Давыдов С.Г., Логунова О.С.²⁵, Климова А. М., Куликов С. П., Чмель К. Ш.²⁶

Теоретическая и методологическая основы исследования

Прежде всего, теоретическую основу исследования составляет *концепция «нереактивного измерения»*. Представления о реактивности и нереактивности впервые обстоятельно были изложены Юджином Уэббом и коллегами в книге «Незаметные меры: Нереактивное исследование в социальных науках»²⁷, где они формулируют термин «нереактивные меры» (*unobtrusive measures*). Нереактивное измерение определяется как измерение, которое не требует от испытуемого ни участия в исследовании, ни, что возможно более важно, осознания факта участия в исследовании²⁸. Такое измерение противопоставляется «реактивному» измерению – с использованием классического методологического арсенала социальных и психологических наук (интервью, опросы, эксперименты, фокус-группы и др.), которые основывались на непосредственном и осознанном вовлечении исследуемого в исследовательские процедуры.

Так, в качестве «нереактивных» можно обозначить целый класс данных, полученных в результате такого измерения. Уэбб и коллеги исходно выделили три

²¹ Девятко И. Ф. Инструментарий онлайн-исследований: попытка каталогизации // В кн.: Онлайн-исследования в России 3.0 / Отв. ред.: И. Ф. Девятко, А. В. Шашкин, С. Г. Давыдов; науч. ред.: И. Ф. Девятко. М.: OMI RUSSIA, 2012. С. 17–30.

²² Дудина В. И. Цифровые данные — потенциал развития социологического знания // Социологические исследования. 2016. № 9. С. 21–30.

²³ Одинцов А.В. Социология общественного мнения и вызов Big Data //Мониторинг общественного мнения: экономические и социальные перемены. 2017. № 3. С. 30—43.

²⁴ Федорова А.А., Николаенко Г.А. Нереактивная стратегия: применимость незаметных методов сбора социологической информации в условиях web 2.0 на примере цифровой этнографии и big data // Социология власти, 2017. Т. 29, № 4. С. 36–54.

²⁵ Давыдов С.Г., Логунова О.С. Проект «Индекс цифровой грамотности»: методические эксперименты // Социология: методология, методы, математическое моделирование (Социология: 4М). 2015. Том. 0. № 41. С. 120–141.

²⁶ Климова А. М., Куликов С. П., Чмель К. Ш. Роль социальных медиа в формировании регионального экологического протеста в России // Мониторинг общественного мнения: Экономические и социальные перемены. 2021. № 6. С. 28–52.

²⁷ Webb E.J., Campbell D.T., Schwartz R.D., Sechrest L. Unobtrusive measures: nonreactive research in the social sciences. Chicago: Rand McNally, 1966. 220 p.

²⁸ Там же.

типа неактивных данных – наблюдение (*observation*), физические следы (*physical traces*) и архивы (*archives*)²⁹, в последнюю группу входят также документы, дневниковые записи, фото- и видео-записи. Однако широкого распространения концепция не получила, интерес к ней появился только в начале 2000-х. Так, одним из первых на возрастающий объем онлайн-информации обратил внимание Реймонд Ли³⁰, который описал цифровые данные как новый тип неактивных данных. Он же предложил другую классификацию данных, основанную не на разной физической природе, а на роли исследователя в сборе и обработке неактивных данных, а именно различать данные на: «найденные» (*found*), «собранные/зарегистрированные» (*captured*), и «извлеченные» (*retrieved*)³¹. По сравнению с дихотомией «реактивный/неактивный» в основе этой классификации артикулируется характер активности исследователя (и возможное влияние методов сбора данных на ошибку измерения).

Можно отметить низкую степень изученности концепции «неактивности» российскими социологами. Федорова А.А. и Николаенко Г.А. отмечают, что к «неактивности» как к исследовательской стратегии российские исследователи не обратились ни в период первых волн интереса (в 1970-е и 1990-е), ни в настоящее время, когда очевидно растет внимание к изучению цифровых данных³². Вместе с тем «неактивность» представляется артикулировать при обсуждении методологических вопросов интеграции как важное концептуальное различие, где одним из важных вопросов являются различия между типами связываемых данных, а также способами их получения.

Вторым важным теоретическим концептом в настоящей диссертации являются *представления о социологическом измерении и критериях оценки его качества*. Общие представления об измерении в социальных науках были

²⁹ Там же.

³⁰ Lee R. M. *Unobtrusive Measures in Social Research*. Philadelphia, PA: Open University Press. 2000. 175 p.

³¹ Девятко И. Ф. Инструментарий онлайн-исследований: попытка каталогизации // В кн.: *Онлайн-исследования в России 3.0* / Отв. ред.: И. Ф. Девятко, А. В. Шашкин, С. Г. Давыдов; науч. ред.: И. Ф. Девятко. М.: OMI RUSSIA, 2012. С. 17–30.

³² Федорова А.А., Николаенко Г.А. Неактивная стратегия: применимость незаметных методов сбора социологической информации в условиях web 2.0 на примере цифровой этнографии и big data // *Социология власти*, 2017. Т. 29, № 4. С. 36–54.

заложены Огюстом Конттом. В его методологии в качестве возможных методов познания социальной реальности обозначаются наблюдение, эксперимент, сравнение³³. Измерение в социологии можно рассматривать как способ дать количественную оценку социальным феноменам/явлениям, или же, согласно определению П. Лазарсфельда, как «поиск упорядоченной классификации»³⁴. Прежде всего в качестве индикаторов таких явлений выступают субъективные переменные (установки, самоотчёты). Пик академического интереса к вопросам социологического измерения (в частности, измерению установок, шкалированию, процедурам факторного анализа и латентно-структурного анализа) пришёлся на 1940 – 1950-е гг. После, достигнув некоторого плато, исследования в этой области приостановились, а в повседневном социологическом обороте остались сравнительно простые шкалы и индексы³⁵. Однако стремительное развитие технологических возможностей для проведения исследований (например, сбор и анализ больших данных), как видится, вновь проблематизирует обозначенные ранее вопросы к качеству социологического измерения. Прежде всего, вопросы качества адресованы новым типам неактивных данных, которые и находятся в фокусе внимания настоящего диссертационного исследования.

Измерение – процесс привязки абстрактных концепций к эмпирическим показателям³⁶. Так как эти концепты нельзя напрямую наблюдать для их измерения разрабатываются специальные индикаторы, при чем из-за априорной открытости этих абстрактных концептов возможна множественность наблюдаемых эмпирических индикаторов. Вместе с этим у исследователей возникает потребность определить, в какой степени конкретный эмпирический показатель (или набор эмпирических показателей) отражает данную теоретическую концепцию³⁷, т.е. произвести оценку качества используемой модели. Для оценки

³³ Гофман А. Б. Семь лекций по истории социологии. Изд. 10-е. М.: Книжный Дом «Университет», 2015. 240 с.

³⁴ Лазарсфельд П. Измерение в социологии // Американская социология. Перспективы, проблемы, методы / пер.: В.В. Воронина, Е.В. Зиньковского; под ред. Г.В. Осипова. М.: Прогресс, 1972. С. 134–149.

³⁵ Blalock, H. M. (Ed.). *Measurements in the Social Sciences: Theories and Strategies*. N.J.: AldineTransaction, 1974. 464 p.

³⁶ Zeller R.A., Carmines E.G. *Measurement in the Social Sciences: The Link between Theory and Data*. Cambridge: Cambridge University Press, 1980. 198 p.

³⁷ Carmines, E. G., & Zeller, R. A. *Reliability and Validity Assessment*. New York: Sage Publications, 1979. 72 p.

качества измерения были сформированы представления о валидности и надёжности исследования. *Валидность измерения (validity)* принято определять как соответствие измерения его цели, иначе говоря, в какой мере мы измеряет то, что намереваемся измерить. *Надёжность (reliability)* исследования принято характеризовать как показатель согласованности и устойчивости (воспроизводимости) измерения. Он характеризует, в какой мере с помощью одного и того же инструмента возможно получить одни и те же результаты при повторных замерах³⁸. В настоящем диссертационном исследовании мы фокусируемся на оценке *конкурентной валидности (concurrent validity)*³⁹ ретроспективных самоотчетов об использовании смартфона.

Проведение оценки качества измерения представляется трудоемким занятием, которое требует разработки методологических программ, стандартов и критериев оценки. Однако эти процедуры представляются необходимым условием для построения теорий среднего уровня и развития социологической теории в целом.

Методологической основой исследования является методический подход «пожертвование данных» (*data donation*)⁴⁰. Он позволяет исследователям собирать цифровые поведенческие следы непосредственно от участников исследования, т.е. получать доступ к тем данным, которые уже собрали о пользователях цифровые платформы. В рамках диссертационного исследования был проведен квази-эксперимент по измерению валидности ретроспективного самоотчёта с использованием реактивных и нереактивных данных. Цель эксперимента состояла в оценке конкурентной валидности самоотчетов о мобильном использовании. Кросс-валидация данных самоотчёта с цифровыми поведенческими следами представляется как первый шаг к их дальнейшей интеграции. Второй эксперимент в настоящее время находится на этапе реализации.

³⁸ Zeller R.A., Carmines E.G. Measurement in the Social Sciences: The Link between Theory and Data. Cambridge: Cambridge University Press, 1980. 198 p.

³⁹ Конкурентная валидность – подвид критериальной валидности.

⁴⁰ Ohme, J., Araujo, T., De Vreese, C. H., Piotrowski, J. T. Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function // Mobile Media & Communication. 2021. Vol. 9. No. 2. P. 293–313.

Научная новизна исследования

Научная новизна исследования заключается в следующих результатах диссертационного исследования:

1. Рассмотрены возможности использования нереактивных данных в социологическом исследовании.
2. Определены основные тематические направления интеграции данных, которые охарактеризованы по следующим параметрам: объекту исследования, уровню интеграции, типу сбора данных, решаемым методологическим проблемам и перспективам развития.
3. Проблематизированы общие вопросы качества данных ретроспективных самоотчётов, систематизирована информация об альтернативных самоотчёту источниках данных для измерения поведения и установок.
4. Разработан план эксперимента для измерения конкурентной валидности ретроспективного самоотчёта об использовании смартфона, в рамках которого был пробирован новый подход к сбору нереактивных данных – «пожертвование данных» (*data donation*). Результаты эксперимента с использованием такого подхода представляются в русскоязычной научной литературе впервые.
5. Методом корреляционного анализа оценена конкурентная валидность самоотчёта о мобильном медиапотреблении, что позволяет охарактеризовать уровень конкурентной валидности самоотчёта этого вида как удовлетворительный.

Основные результаты исследования и положения, выносимые на защиту

1. Цифровые следы представляют собой новый класс «малореактивных» данных и могут выступать в качестве объективных поведенческих индикаторов, предоставляя возможность получения более валидных данных с более высоким уровнем детализации. Потенциально использование цифровых следов может снизить нагрузку на респондентов и уменьшить

ошибку измерения, однако их использование в социологическом исследовании должно быть теоретико-методологически обосновано.

2. В работе определяются три вида альтернативных ретроспективному самоотчёту источников данных для измерения поведения и установок: малореактивные данные («достоверные» внешние источники), наблюдение, дневники.
3. Результаты эксперимента позволяют оценить методический подход «пожертвование данных» как перспективный для сбора цифровых следов с мобильных устройств. Рассматриваемый подход может использоваться для изучения практики мобильного использования, что позволяет исследователям получать объективные поведенческие данные об активности смартфона без установки сторонних приложений (тайм-трекеров). Функции трекинга, предусмотренные разработчикам в цифровых устройствах, исходно имеют целью предоставить пользователю возможность осуществления контроля за своим использованием платформ, однако могут использоваться и в исследовательских целях. В целом, знания о валидности данных, полученные в рамках рассматриваемого методического подхода ещё достаточно ограничены, а практика применения сопряжена со рядом сложностей в процессе сбора и обработки данных, однако можно предположить, что подход займёт свою нишу как один из возможных способов сбора малореактивных данных.
4. Уровень конкурентной валидности самоотчёта может быть охарактеризован как умеренный с точки зрения тесноты связи. Такие результаты позволяют говорить о ретроспективном самоотчёте о мобильном использовании как о сравнительно валидном методе измерения практик использования смартфона для временных интервалов в пределах недели.
5. При прочих равных цифровые следы представляются наиболее предпочтительными для оценки объема использования смартфона с точки зрения доступности данных для сбора, отсутствия существенных дополнительных издержек, а также, предположительно, сокращения времени

заполнения анкеты и снижения нагрузки на респондентов. Однако в рамках полномасштабных опросов сбор цифровых поведенческих следов потребует разработки новых методических подходов, решения ряда технологических и этических вопросов.

Апробация результатов исследования

Результаты диссертационного исследования были (будут) представлены на следующих российских и международных конференциях:

1. 15 сентября 2023 г. – Inhabiting Uncertainty. A Forum On Qualitative, Quantitative, And Visual Methods (Рим, Италия). Доклад: “Integrating Self-Reports and Mobile Tracker Data: Results of a Methodological Experiment”
2. 22 сентября 2023 г. – WAPOR 76th Annual Conference (Зальцбург, Австрия). Доклад: “Integrating Self-Reports and Mobile Tracker Data: Results of a Methodological Experiment”
3. *Ожидается (поданы тезисы):* 22 октября 2024 г., научная конференция «Социология: вчера, сегодня, завтра» (Москва, Россия). Организаторы: Федеральный научно-исследовательский социологический центр РАН, редакция журнала «Социологические исследования» и др. Доклад: «Оценка валидности данных самоотчёта об использовании смартфона с использованием цифровых поведенческих следов: результаты методического эксперимента»

Научные публикации по теме диссертационного исследования:

1. Сапонова А. В., Куликов С. П. Интеграция опросных данных и цифровых следов: обзор основных методологических подходов // Социология: методология, методы, математическое моделирование. 2021. № 53. С. 117–164.
2. Сапонова А. В. Оценка валидности данных самоотчёта об использовании смартфона с использованием цифровых поведенческих следов // Социологический журнал. 2024. (в печати)

Третья статья находится в настоящее время в разработке. Также автор диссертации дважды выступал в качестве куратора проектных групп в рамках Зимней школы факультета социальных наук НИУ ВШЭ в 2023 и 2022 гг., и осуществлял руководство разработкой студенческих проектов по теме своего диссертационного исследования.

Список использованных источников и литературы

1. Богданов М.Б., Смирнов И.Б. Возможности и ограничения цифровых следов и методов машинного обучения в социологии // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 304–328.
2. Гофман А. Б. Семь лекций по истории социологии. Изд. 10-е. М.: Книжный Дом «Университет», 2015. 240 с.
3. Губа К. С. Большие данные в социологии: Новые данные, новая социология? Социологическое обозрение, 2018. № 17(1). С. 213–234.
4. Давыдов С.Г., Логунова О.С. Проект «Индекс цифровой грамотности»: методические эксперименты // Социология: методология, методы, математическое моделирование (Социология: 4М). 2015. Том. 0. № 41. С. 120-141.
5. Девятко И. Ф. Инструментарий онлайн-исследований: попытка каталогизации // В кн.: Онлайн-исследования в России 3.0 / Отв. ред.: И. Ф. Девятко, А. В. Шашкин, С. Г. Давыдов; науч. ред.: И. Ф. Девятко. М.: OMI RUSSIA, 2012. С. 17–30.
6. Девятко И. Ф. Социологическая теория: старые трудности, новые вызовы // Социологические исследования. 2021. № 10. С. 3–11.
7. Девятко И.Ф. Новые данные, новая статистика: от кризиса воспроизводимости к новым требованиям к анализу и представлению данных в социальных науках // Социологические исследования. 2018. № 12. С. 30–38.
8. Дудина В. И. Цифровые данные — потенциал развития социологического знания // Социологические исследования. 2016. № 9. С. 21–30.
9. Климова А. М., Куликов С. П., Чмель К. Ш. Роль социальных медиа в формировании регионального экологического протеста в России // Мониторинг общественного мнения: Экономические и социальные перемены. 2021. № 6. С. 28–52.
10. Лазарсфельд П. Измерение в социологии // Американская социология. Перспективы, проблемы, методы / пер.: В.В. Воронина, Е.В. Зиньковского;

- под ред. Г.В. Осипова. М.: Прогресс, 1972. С. 134–149.
- 11.Одинцов А.В. Социология общественного мнения и вызов Big Data // Мониторинг общественного мнения: экономические и социальные перемены. 2017. № 3. С. 30–43.
 - 12.Федорова А.А., Николаенко Г.А. Нереактивная стратегия: применимость незаметных методов сбора социологической информации в условиях web 2.0 на примере цифровой этнографии и big data // Социология власти, 2017. Т. 29, № 4. С. 36–54.
 - 13.Araujo T., Neijens P. Unobtrusive Measures for Media Research // The International Encyclopedia of Media Psychology. 1st ed. / ed. Bulck J. Wiley, 2020. P. 1–7.
 - 14.Beninger K. et al. Understanding Society: How people decide whether to give consent to link their administrative and survey data / Kantar Public UK, 2017. № 2017 – 13. 65 p.
 - 15.Blalock, H. M. (Ed.). Measurements in the Social Sciences: Theories and Strategies. N.J.: AldineTransaction, 1974. 464 p.
 - 16.Bouchard Jr T. J. Field research methods: Interviewing, questionnaires, participant observation, systematic observation, unobtrusive measures // Handbook of industrial and organizational psychology / ed. M. D. Dunnette. Chicago: Rand McNally. 1976. Vol. 1. P. 363–413.
 - 17.Campbell D. T., Fiske D. W. Convergent and discriminant validation by the multitrait-multimethod matrix // Psychological bulletin. 1959. Vol. 56. №. 2. P. 81–105.
 - 18.Carmines, E. G., & Zeller, R. A. Reliability and Validity Assessment. New York: Sage Publications, 1979. 72 p.
 - 19.de Vreese C., Semetko H.A. News matters: Influences on the vote in the Danish 2000 euro referendum campaign // European Journal of Political Research. 2004. Vol. 43, № 5. P. 699–722.
 - 20.Haenschen K. Self-Reported Versus Digitally Recorded: Measuring Political Activity on Facebook // Social Science Computer Review. 2020. Vol. 38, № 5. P.

567–583.

21. Hofstra B. et al. Sources of Segregation in Social Networks: A Novel Approach Using Facebook // *American Sociological Review*. 2017. Vol. 82, № 3. P. 625–656.
22. Hopp, T., Vargo, C. J., Dixon, L., Thain, N. Correlating Self-Report and Trace Data Measures of Incivility: A Proof of Concept // *Social Science Computer Review*. 2020. Vol. 38, № 5. P. 584–599.
23. Jürgens P., Stark B., Magin M. Two Half-Truths Make a Whole? On Bias in Self-Reports and Tracking Data // *Social Science Computer Review*. 2020. Vol. 38, № 5. P. 600–615.
24. Keusch, F., Struminskaya, B., Antoun, C., Couper, M. P., & Kreuter, F. Willingness to Participate in Passive Mobile Data Collection // *Public Opinion Quarterly*. 2019. Vol. 83. No. S1. P. 210–235.
25. Lee R. M. *Unobtrusive Measures in Social Research*. Philadelphia, PA: Open University Press, 2000. 175 p.
26. Ohme, J., Araujo, T., De Vreese, C. H., Piotrowski, J. T. Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function // *Mobile Media & Communication*. 2021. Vol. 9. No. 2. P. 293–313.
27. Pasek, J., McClain, C. A., Newport, F., Marken, S. Who’s Tweeting About the President? What Big Survey Data Can Tell Us About Digital Traces? // *Social Science Computer Review*. 2020. Vol. 38, № 5. P. 633–650.
28. Stier S. et al. Integrating Survey Data and Digital Trace Data: Key Issues in Developing an Emerging Field // *Social Science Computer Review*. 2020. Vol. 38, № 5. P. 503–516.
29. van Spanje J., de Vreese C. Europhile Media and Eurosceptic Voting: Effects of News Media Coverage on Eurosceptic Voting in the 2009 European Parliamentary Elections // *Political Communication*. 2014. Vol. 31, № 2. P. 325–354.
30. Vraga, E. K., Tully, M. Who is exposed to news? It depends on how you measure: Examining self-reported versus behavioral news exposure measures // *Social Science Computer Review*. 2020. Vol. 38, №5. P. 550–566.

31. Webb E.J., Campbell D.T., Schwartz R.D., Sechrest L. Unobtrusive measures: nonreactive research in the social sciences. Chicago: Rand McNally, 1966. 220 p.
32. Yanovitzky I. Effect of Call-In Political Talk Radio Shows on Their Audiences: Evidence from a Multi-Wave Panel Analysis // International Journal of Public Opinion Research. 2001. Vol. 13, № 4. P. 377–397.
33. Zeller R.A., Carmines E.G. Measurement in the Social Sciences: The Link between Theory and Data. New York: Cambridge University Press, 1980. 198 p.