



DOI: 10.22363/2313-2272-2018-18-4-627-637

РАЗРАБОТКА ПОДХОДА К КОЛИЧЕСТВЕННОЙ МУЛЬТИМОДАЛЬНОЙ ОЦЕНКЕ КОГНИТИВНОЙ НАГРУЗКИ ИНТЕРВЬЕОРОВ: РЕЗУЛЬТАТЫ ПИЛОТНОГО КВАЗИЭКСПЕРИМЕНТА*

И.Ф. Девятко

Национальный исследовательский университет «Высшая школа экономики»
ул. Мясницкая, 11, Москва, 101000, Россия

Институт социологии ФНИСЦ РАН
ул. Кржижановского, 24/35-5, Москва, 117218, Россия
(e-mail: deviatko@gmail.com)

Представление о том, что опросный инструментарий, опосредующий коммуникацию между интервьюером и респондентом, влияет на качество опросных данных, давно превратилось в общеизвестную методологическую мудрость. Однако влияние на качество данных когнитивной нагрузки, испытываемой интервьюером в процессе проведения интервью, пока не получило систематического рассмотрения в существующей литературе. В ситуации, когда анкета заполняется интервьюером, именно последний должен распределять ограниченные индивидуальные ресурсы внимания, памяти, визуального и моторного контроля, а также использовать навыки активного слушания и интерпретации, чтобы свести к минимуму недопонимание вопроса респондентом и собственные ошибки фиксации ответов. Тем не менее, среди различных методов предварительного тестирования качества опросного инструментария или оценки потенциальных эффектов способа проведения опроса методы мультимодальной количественной оценки когнитивной нагрузки интервьюера, сопряженной с использованием опросного инструментария во время интервью, все еще редки или практически отсутствуют. В статье представлены: краткий перечень субъективных, поведенческих и физиологических мер когнитивной нагрузки, используемых в таких дисциплинарных областях, как когнитивная наука, эргономика и т.д., а также описание и обсуждение предварительных результатов пилотного квазиэксперимента, участниками которого стали интервьюеры «Российского мониторинга экономического положения и здоровья населения НИУ ВШЭ», проходившие обучение в процессе перехода от интервьюирования с помощью бумажной анкеты (РАPI) к использованию компьютерных технологий в процессе проведения интервью (САPI). Результаты исследования позволили продемонстрировать некоторые преимущества и технические ограничения, связанные с параллельным использованием варианта популярной шкалы самооценки когнитивной нагрузки (шкала Пааса) и простого физиологического индикатора (частота сердечных сокращений), дополненных данными видеозаписи с экрана планшета, применявшегося в процессе интервьюирования, в целях мультимодальной количественной оценки когнитивной нагрузки интервьюера при проведении интервью.

Ключевые слова: когнитивная нагрузка; мультимодальная оценка когнитивной нагрузки, связанной с проведением интервью; РАPI и САPI; качество опросных данных; шкала Пааса; частота сердечных сокращений

* © Девятко И.Ф., 2018.

Статья подготовлена при поддержке Российского научного фонда. Проект № 17-78-20172.
Статья поступила в редакцию 22.08.2018 г.

Величина когнитивной нагрузки на интервьюера и респондента при проведении опроса — ключевой компонент оценки затрачиваемых ими когнитивных ресурсов и умственных усилий, который позволяет не только локализовать потенциальные источники ошибок измерения или учесть риски использования интервьюерами или респондентами стратегий, минимизирующих усилия и время заполнения в ущерб качеству данных, но и выбирать оптимальные форматы вопроса [11], управлять эффектами социальной желательности [17] или планировать оптимальное сочетание способов проведения опроса (например, телефонное интервью, онлайн-опрос и т.д.) и технических решений, касающихся, в частности, использования компьютерных технологий в сборе данных [см., напр.: 3; 13].

В ситуации проведения интервью «лицом-к-лицу» анкета заполняется интервьюером со слов респондента, и именно первый должен использовать ограниченные ресурсы внимания, памяти, зрительного и моторного контроля, активного слушания и интерпретации таким образом, чтобы минимизировать ошибки в понимании вопроса респондентом и ошибки фиксации ответа.

Очевидно, что решение множественных параллельных задач — существенная характеристика деятельности интервьюера, которая предполагает не только значительные требования к перечисленным когнитивным ресурсам и навыкам, но и наличие метакогнитивных и саморегуляторных умений, в частности, навыка осуществления параллельного основной задаче контроля над работой техники в ходе интервью, умения поддерживать оптимальный темп интервью с учетом используемых средств фиксации (например, ручного письма в случае личного интервью с помощью бумажной версии вопросника — PAPI — или навыков использования сенсорной или физической клавиатуры устройства в случае интервью, проводимого с помощью компьютерных технологий — CAPi). Описанная ситуация решения множественных задач в процессе интервьюирования, включающих в себя поддержание коммуникации с респондентом, контроль над собственными действиями по заполнению вопросника и т.п., может приводить к исчерпанию доступных когнитивных ресурсов и достижению «красных линий» когнитивной перегрузки [6. Р. 34—39] — точек, в которых требования задачи превосходят доступные когнитивные ресурсы интервьюера, что, в конечном счете, ведет к снижению качества данных.

Если особенности дизайна опросного инструментария, влияющие на качество измерения, или потенциальное воздействие личности интервьюера на коммуникацию с респондентом привлекают устойчивый интерес и являются темой множества качественных исследований и методических экспериментов, задача прямой количественной оценки когнитивной нагрузки и умственных усилий (1), затрачиваемых интервьюером (и респондентами) в процессе интервью, до сих пор редко оказывалась в центре внимания социологов. Так, в сравнительно немногочисленных исследованиях производилась оценка времени заполнения или количества исправлений как косвенных поведенческих индикаторов, связанных с уровнем когнитивных усилий респондента, потенциально влияющих на качество данных (см., в частности, [4; 11]). Однако попытки разработки комплексного подхода

к интегральной количественной оценке разных аспектов когнитивной нагрузки, связанной с разными задачами, психическими процессами и модальностями (слух, зрительное внимание и т.п.), задействованными в процессе интервью, до сих пор не предпринимались.

Вместе с тем в последние десятилетия в таких дисциплинарных областях, как когнитивная наука, эргономика, психология образования, накоплен значительный опыт практического использования субъективных, поведенческих и физиологических индикаторов когнитивной нагрузки при выполнении множественных задач, которые используются как в исследованиях обучения, так и в оценке профессиональной деятельности (например, программистов, проверяющих компьютерный код [12]).

Систематический анализ накопленного в междисциплинарных исследованиях опыта измерения когнитивной нагрузки позволяет предварительно определить первичный набор индикаторов, потенциально обладающих свойствами надежности измерения и практической применимости для оценки когнитивной нагрузки интервьюера, связанной с заполнением отдельных блоков и анкеты в целом в полевых условиях или на стадии разработки инструментария и проведения тренингов интервьюеров. На основе такого анализа ранее [2] были очерчены ближайшие перспективы, возможности и потенциальные трудности разработки подхода к интегральной оценке когнитивной нагрузки применительно к деятельности по заполнению опросников, преимущественно с точки зрения оценки когнитивной нагрузки интервьюеров при проведении опросов с использованием разных способов заполнения (бумажная анкета, заполнение вопросников с использованием переносных компьютеров и планшетов). В данной статье, исходя из широкого определения когнитивной нагрузки как *многомерного конструкта, характеризующего ментальные усилия, связанные с решением одной или нескольких задач, структура и содержание которых задают необходимый объем привлекаемых для их успешного решения ограниченных когнитивных ресурсов — внимания, кратковременной памяти, моторного контроля и т.д.*, охарактеризованы три основные группы мер когнитивной нагрузки: 1) опросные техники, прежде всего основанные на субъективной оценке уровня когнитивной нагрузки по одномерной или многомерной шкале; 2) объективные физиологические (и нейрофизиологические) неинвазивные методы измерения когнитивной нагрузки, использующие показатели, которые в большинстве ситуаций не поддаются произвольному контролю со стороны испытуемых (чаще всего — получаемые с помощью пупиллометрии данные о динамике размера зрачка, частоте сердцебиений и ее вариабельности, реже — данные айтрекеров о динамике размера зрачка или фиксации взгляда, результаты ЭЭГ, измерения кожно-гальванической реакции и т.д.), 3) дополнительные методы измерения когнитивной нагрузки, основанные на объективных поведенческих данных об уровне исполнения, включая параданные и другие поведенческие индикаторы, часто измеряемые с использованием параллельно выполняемой второй задачи (время исполнения задачи, количество ошибок).

В статье представлены результаты пилотного полевого исследования, направленного на оценку уровня когнитивной нагрузки на первой стадии перехода интервьюеров Российского мониторинга экономического положения и здоровья населения НИУ ВШЭ (далее — РМЭЗ НИУ ВШЭ) (2) от бумажных вопросников (РАPI) к их заполнению на планшете (САPI). В ходе исследования оценивались также возможности и технические особенности интегральной количественной оценки когнитивной нагрузки интервьюеров с помощью элементарных субъективного и объективного индикаторов (3). При проведении пилотного квазиэксперимента интервьюеры, проходившие офлайн тренинг в Москве ($N = 24$, женщины, средний возраст — $60,9 \pm$ стандартное отклонение — $10,6$), сначала заполняли анкету «до тренинга», содержащую необходимую информацию о социально-демографических характеристиках, наличии релевантного опыта работы в РМЭЗ и других опросах, в том числе с применением планшетов или смартфонов, об общем уровне «цифровой грамотности», а также о воспринимаемой сложности заполнения девяти основных разделов анкеты РМЭЗ НИУ ВШЭ. Полное описание процедур тренинга и анкетирования до и после обучения заполнению вопросника РМЭЗ НИУ ВШЭ на планшете, а также основные результаты анкетирования интервьюеров приводятся в статье [1].

После практико-ориентированного тренинга, в каждой сессии которого участвовало от четырех до шести интервьюеров, участникам предлагалось самостоятельно провести тестовое интервью с «модельным» респондентом (членом исследовательской команды или техническим сотрудником) по полной версии содержавшего ряд дополнительных блоков «Вопросника для взрослых» 26-й волны РМЭЗ НИУ ВШЭ с использованием планшета (САPI). Участие в таком тестовом интервью было добровольным и обычно не вызывало отказов (4). После тестового интервью участники заполняли бланк оценки когнитивной нагрузки для всех блоков (разделов) «Вопросника» 26-й волны при интервьюировании в режиме САPI и, позднее, анкету «после тренинга» [1].

Шкала воспринимаемой когнитивной нагрузки. До начала и после завершения тестового интервью на планшете интервьюеры заполняли краткую анкету, содержащую модифицированную (пятибалльную) версию одномерной шкалы воспринимаемой когнитивной нагрузки Пааса (5). Для девяти основных разделов бумажной версии вопросника (т.е. для режима РАPI) когнитивная нагрузка как «сложность заполнения данного раздела при проведении интервью» оценивалась до, а для шестнадцати разделов «Вопросника для взрослых» 26-й волны — после проведения интервью в режиме САPI (если некоторый раздел, например, «Пенсия» или «Только для женщин», в ходе проведения интервью с «модельным» респондентом не заполнялся, оценивание когнитивной нагрузки для него не проводилось).

Шкала Пааса, которая основана на ретроспективном самоотчете индивидов о воспринимаемом уровне когнитивной нагрузки в ходе выполнения разных заданий — от решения задач по статистике до нахождения ошибок в тексте — это очень популярный метод оценки когнитивной нагрузки, который использо-

вался сначала в исследованиях интеллекта [5], но основное развитие получил в педагогической психологии и прикладных исследованиях образования [14; 15]. Надежность-согласованность и валидность шкалы Пааса подтверждены в значительном количестве исследований [см., напр.: 2. С. 8—10].

При использовании той или иной версии одномерной шкалы Пааса перед участниками исследования ставится задача прямой субъективной оценки умственных усилий, затраченных ими при решении отдельной задачи или выполнении блока заданий, по девятибалльной (иногда семибалльной) шкале ликертовского (лайкертовского) типа (в диапазоне от «очень, очень/крайне небольшие» до «очень, очень/крайне большие»).

В нашем исследовании испытуемые заполняли в бумажной форме сокращенную и инвертированную версию шкалы, которая своим форматом напоминала другие оценочные шкалы, использовавшиеся в анкетах «до тренинга» и «после тренинга». Задача участника исследования заключалась в том, чтобы ответить на вопрос, насколько были сложны для нее/него разные разделы вопросника при проведении интервью «с точки зрения затраченных ею/им умственных усилий (в частности, внимания, памяти, самоконтроля)» (6). Для ответа во всех случаях использовалась пятибалльная шкала, где оценка «1» означала ответ «блок вопросов очень сложный для заполнения», а «5» — «блок вопросов очень легкий для заполнения». Исходя из данных похожих по замыслу предыдущих исследований установок интервьюеров при переходе с PAPI на CAPI [см., напр.: 7; 9]), мы ожидали, что общая оценка когнитивной нагрузки при заполнении «Вопросника» с помощью планшета будет не выше или даже ниже оценки когнитивной нагрузки при заполнении бумажной версии.

Частота сердечных сокращений. В ходе исследования изучалась возможность фонового и малозаметного применения надежного (хотя и косвенного) коррелята *динамики общей активации* при когнитивной нагрузке — частоты сердечных сокращений. Увеличение частоты сердечных сокращений — простейший с точки зрения доступности измерения индикатор активности симпатической нервной системы (наряду с ростом давления, потоотделением и изменениями электрической активности кожи). Умеренный, т.е. не слишком низкий или не слишком высокий, уровень общей активации оказывает положительное воздействие на когнитивные процессы, в частности, на функции памяти, способность поддерживать концентрацию внимания и т.п. [8; 10; 16].

В нашем исследовании были использованы результаты пульсометрии — средние значения частоты сердечных сокращений за минуту (ЧСС), полученные с применением наручного браслета Mi Band 2 и приложения для выгрузки данных Mi Heart Rate (кастомизированного для Android или iOS). Браслет работал в фоновом режиме непрерывного измерения пульса, передавая в приложение данные о среднем значении ЧСС для заданного временного интервала с частотой около 0,083 Гц (поскольку для выгрузки данных использовались различные смартфоны, этот параметр зависел от устройства, в частности, для одной из моделей с iOS не удалось получить точные временные метки для выгружаемых данных в секундном диапазоне).

Перед началом каждого тестового интервью интервьюеры заполняли «Форму согласия» и получали разъяснения относительно функций браслета. Интервал адаптации к ношению браслета (устройство выглядит как небольшие пластиковые смарт-часы на силиконовом ремешке) составлял в среднем 7—9 минут. Данные за последние 1,5—2 минуты этого интервала использовались при анализе динамики ЧСС по разделам анкеты в качестве индивидуального базового значения ЧСС, характерного для испытуемого. Каждого интервьюера спрашивали о наличии заболеваний сердечно-сосудистой системы и систематически принимаемых в этой связи препаратах (поскольку для нашей небольшой выборки эффекты этих двух переменных не были значимы, мы не будем здесь рассматривать эти данные).

Видеозапись с экрана планшета. С целью синхронизации временных меток для рядов данных пульсометрии и фонового мониторинга прохождения отдельных разделов вопросника использовалось приложение для записи видео с экрана DU Recorder в версии, оптимизированной для планшетов, работающих на Android (1.5.7), предустановленное на используемых в ходе тренинга планшетах. Полученная видеозапись с указанием временных меток прохождения вопросника была использована для разбиения последовательности данных пульсометрии по разделам (блокам) «Вопросника», осуществлявшегося вручную двумя кодировщиками. В будущем полученные таким образом видеозаписи могут быть использованы и для анализа ошибок и исправлений при заполнении анкеты (в данном случае СAPI-интервью проводилось каждым испытуемым однократно, поэтому количество ошибок было невелико и соответствующие данные не анализировались).

Результаты эксперимента позволили оценить качество получаемых с применением описанных мер оценки когнитивной нагрузки данных (в том числе с точки зрения масштаба технических потерь данных в ходе фоновой записи, точности временной разметки при экспорте данных пульсометрии на разные мобильные устройства — в частности, были обнаружены неточности при экспорте на некоторые мобильные устройства на iOS); были оценены возможные потери в результате технических ошибок и полной или частичной потери записи с экрана (около 25%).

Говоря о возможности использования приложения DU Recorder для *видеозаписи действий* испытуемых на экране планшета при проведении интервью, следует отметить его преимущества: бесплатность, фоновый характер записи, не отвлекающий испытуемого от выполнения основной задачи на планшете, возможность использования видеозаписей как источника множественных поведенческих индикаторов. Впрочем, упомянутый малозаметный и фоновый характер работы приложения имеет и обратную сторону: прекращение записи в случае сбоя в работе приложения нередко оставалось незамеченным и вело к полной или частичной потере данных для наблюдения (по крайней мере, в трех случаях потеря данных предположительно могла быть связана с интерференцией с работой других устройств, поскольку происходила в конкретной локации использовавшегося для проведения тренингов помещения).

В результате анализа данных *пульсометрии* для зависимой переменной ЧСС ($N = 18$ испытуемых) с использованием линейной модели со смешанными эффектами и повторными измерениями (оценивание проводилось методом ограни-

ченного максимального правдоподобия), позволившей анализировать данные с пропусками (7) и оценивать эффект испытуемого/интервьюера, были выявлены значимые основные эффекты и эффект взаимодействия внутрисубъектного фиксированного фактора «Раздел вопросника» и межсубъектного фактора «Испытуемый», которые планируется уточнить в дополнительном эксперименте: для основного эффекта фактора ««Раздел вопросника» $F(17; 305,512) = 6,985, p < 0,0001$; для основного эффекта испытуемого $F(17; 889,241) = 180,637, p < 0,0001$; для взаимодействия этих факторов $F(263; 509,583) = 4,657, p < 0,0001$. Полученные данные могут рассматриваться как предварительная демонстрация конструктивной валидности показателей пульсометрии при косвенной оценке общей активации как компонента когнитивной нагрузки, возникающей при заполнении интервьюерами отдельных разделов анкеты. Были выявлены, соответственно, и значимые попарные различия для оцененных средних значений ЧСС (например, разность оцененных маргинальных средних между базовым уровнем и уровнем при заполнении блока «Образование», составила 2,704 при $p \leq 0,005$), однако содержательная интерпретация этих различий в условиях квазиэксперимента затруднена тем, что при фиксированном естественном порядке следования блоков анкеты (т.е. без рандомизации, которая была невозможна в реальных условиях обучения интервьюеров работе с «Вопросником») и в отсутствие интервалов покоя при переходе от одного блока к другому динамика ЧСС не может быть однозначно атрибутирована конкретному блоку (значительные изменения могут наблюдаться в течение некоторого времени после завершения опроса по конкретному разделу, т.е. при заполнении следующего раздела). Кроме того, нелинейный характер связи между общим уровнем активации и выполнением задачи требует учета независимо измеренной сложности задачи, а также межиндивидуальных различий в компетенции [см., напр.: 10]. Вместе с тем можно сделать вывод, что данные о динамике ЧСС, наряду с другими индикаторами уровня активации симпатической нервной системы, прежде всего вариабельностью сердечного ритма и изменениями электрической активности кожи, в дальнейшем могут быть использованы для комплексной оценки различных компонентов когнитивной нагрузки.

Анализ данных оценки *воспринимаемой когнитивной нагрузки* по шкале Пааса по итогам тестового заполнения анкеты на планшете отчасти подтвердил ожидаемый паттерн различий между этими оценками и ретроспективными оценками сложности заполнения бумажной анкеты в пользу САРІ (при сравнении с использованием теста Уилкоксона) для трех из девяти основных разделов «Вопросника» («Интернет», «Пенсия», «Только для женщин», $p < 0,05$) и тенденцию к значимости для еще двух разделов («Миграция», «Работа», $p = 0,07$). Для всех этих разделов заполнение на планшете оценивалось как более легкое с точки зрения затраченных умственных усилий в сравнении с ретроспективно оцениваемой сложностью заполнения бумажной анкеты, что может свидетельствовать в пользу конструктивной валидности данной версии шкалы Пааса при оценке когнитивной нагрузки интервьюеров.

Однако более определенные выводы могут быть сделаны лишь при большем количестве наблюдений и для более гетерогенной выборки (абсолютное большинство наших испытуемых относились к старшим возрастным группам и имели

многолетний опыт работы интервьюером). С точки зрения содержательной интерпретации данные о том, что для большинства разделов «Вопросника» воспринимаемая когнитивная нагрузка при заполнении в бумажной форме и на планшете не различаются, а для нескольких разделов когнитивные усилия при заполнении на планшете субъективно оцениваются как меньшие, хорошо согласуются и с данными более ранних исследований сравнительных установок интервьюеров по отношению к работе в режимах PAPI и CAPi [7], и с данными анкетирования «после тренинга» относительно общей оценки опыта работы на планшете [1].

Также в исследовании была продемонстрирована высокая надежность-согласованность оценки общей когнитивной нагрузки интервьюеров по шкале Пааса (альфа Кронбаха для оценки когнитивной нагрузки при работе на планшете была равна 0,945, для ретроспективной оценки когнитивной нагрузки при работе с бумажной анкетой — 0,867), что в целом согласуется с ранее полученными оценками [2]. Однако требуется большее количество наблюдений для статистически обоснованной оценки возможного влияния наличия опыта интервьюирования с помощью планшета или смартфона и общей «цифровой грамотности» интервьюеров на воспринимаемую когнитивную нагрузку в процессе проведения опроса в формате CAPi.

В целом пилотное исследование подтвердило возможности применения таких субъективных и объективных индикаторов когнитивной нагрузки, как шкала Пааса, данные пульсометрии и видеозаписи с экрана мобильного устройства, используемого для заполнения анкеты, для разработки подхода к мультимодальной оценке когнитивной нагрузки интервьюеров, однако уточнение полученных выводов с точки зрения как методологических преимуществ, так и технических удобств и ограничений, требует дальнейших исследований с использованием больших и более разнородных с точки зрения возраста и опыта участников выборок, большей степени контроля над условиями эксперимента, а также оценки конвергентной и дискриминантной валидности использованных в пилотном квазиэксперименте индикаторов в сравнении с другими показателями, широко используемыми в практике междисциплинарных исследований когнитивной нагрузки (прежде всего, с данными пупиллометрии и измерениями динамики показателей электрической активности кожи).

ПРИМЕЧАНИЯ

- (1) Умственные, или ментальные, усилия понимаются обычно как «аспект когнитивной нагрузки, отражающий когнитивные способности, выделенные для удовлетворения запросов, связанных с задачей» [15. Р. 64]. При этом другие аспекты когнитивной нагрузки могут быть связаны с иными аспектами ситуации решения задачи — например, с необходимостью поддерживать концентрацию на задаче в условиях шума или при наличии отвлекающих от задачи нерелевантных стимулов.
- (2) Российский мониторинг экономического положения и здоровья населения (РМЭЗ НИУ ВШЭ, RLMS-HSE) — единственное в России лонгитюдное обследование домохозяйств (см.: <https://www.hse.ru/rlms>).
- (3) Автор выражает признательность коллегам, участвовавшим в обсуждении, проведении исследования и кодировании данных, — А.М. Мавлетовой, Д.В. Лебедеву, М.Б. Богданову.

- (4) Немногочисленные отказы были вызваны преимущественно необходимостью успеть вернуться домой до наступления темноты или временного интервала, когда нужно долго ожидать электрички в Подмоскowie, и т.п. (тренинг проводился в вечернее время, в ноябре 2017 года).
- (5) С точки зрения экспериментального плана, предполагавшего сравнение ретроспективной оценки опыта работы с «Вопросником для взрослых» РМЭЗ НИУ ВШЭ в режиме PAPI с вновь полученным опытом CAPI с тем же вопросником, это был квазиэксперимент, а не истинный эксперимент, поскольку мы не могли рандомизировать приписывание интервьюеров к условиям.
- (6) Следует отметить, что при оценке интервьюерами сложности заполнения основных разделов бумажной версии «Вопросника» до тренинга использовалась более короткая формулировка задачи оценки когнитивной нагрузки: «Насколько, по Вашему мнению, сложны разные разделы анкеты „Проект «Россия»“ лично для Вас при проведении интервью?», т.е. без уточнения «Пожалуйста, оцените сложность заполнения каждого блока вопросов с точки зрения затраченных Вами умственных усилий (в частности, внимания, памяти, самоконтроля)». Однако форматы ответа в обоих случаях не отличались, что обеспечивает базовую сопоставимость оценок.
- (7) Отдельные блоки «Вопросника» могли не заполняться в конкретном случае в силу нерелевантности для «модельного» респондента или, в нескольких случаях, быть частично утрачены из-за некорректной работы приложения для видеозаписи.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК / REFERENCES

- [1] Богданов М.Б., Лебедев Д.В. «Главное не бояться!» — воздействие тренингов на ожидания интервьюеров при переходе от PAPI к CAPI // Социология: методология, методы, математическое моделирование. 2017. № 45. С. 102—132 / Bogdanov M.B., Lebedev D.V. “Glavnoe — ne boyatsya” — vozdejstvie treningov na ozhidaniya intervvyuerov pri perekhode ot PAPI k CAPI [“Do not worry!” — training’s impact on interviewers’ expectations towards the change from PAPI to CAPI]. *Sociologiya: Metodologiya, Metody, Matematicheskoe Modelirovanie*. 2017; 45: 102—132 (In Russ.).
- [2] Девиатко И.Ф., Лебедев Д.В. Глазами интервьюера, глазами респондента: контуры нового подхода к оценке когнитивной нагрузки при проведении опроса // Мониторинг общественного мнения: Экономические и социальные перемены. 2017. № 5. С. 1—19 / Deviatko I.F., Lebedev D.V. Glazami intervvyuera, glazami respondenta: kontury novogo podkhoda k otsenke kognitivnoj nagruzki pri provedenii oprosa [Through the eyes of the interviewer, through the eyes of the respondent: A new approach to the assessment of cognitive load during the interview. *Monitoring Obschestvennogo Mneniya: Ekonomicheskie i Socialnye Peremeny*. 2017; 5: 1—19 (In Russ.).
- [3] 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 / Terentev E.A., Mavletova A.M., Kosolapov M.S. Intervvuirovaniye s pomoschyu kompyuternyh tekhnologij v longityudnyh obsledovaniyah domohozyajstv [Computer-assisted personal interviewing for longitudinal household studies]. *Monitoring Obschestvennogo Mneniya: Ekonomicheskie i Socialnye Peremeny*. 2018; 3: 47—64 (In Russ.).
- [4] Bassili J.N., Scott B.S. Response latency as a signal to question problems in survey research. *Public Opinion Quarterly*. 1996; 60 (3): 390—399.
- [5] Bratfisch O., Borg G., Dornic S. *Perceived Item Difficulty in Three Tests of Intellectual Performance Capacity*. Stockholm: Institute of Applied Psychology; 1972. Report No. 29.
- [6] Chen F. et al. *Robust Multimodal Cognitive Load Measurement*. Springer International Publishing; 2016.

- [7] Couper M.P., Burt G. Interviewer attitudes toward computer-assisted personal interviewing (CAPI). *Social Science Computer Review*. 1994; 12 (1): 38—54.
- [8] Critchley H.D. Electrodermal responses: What happens in the brain. *The Neuroscientists*. 2002; 8 (2): 132—142.
- [9] De Leeuw E.D., Hox J.J., Snijders G. The effect of computer-assisted interviewing on data quality. *International Journal of Market Research*. 1995; 37 (4): 325—344.
- [10] Hoogerheide V., Renkl A., Fiorella L., Paas F., van Gog T. Enhancing example-based learning: Teaching on video increases arousal and improves problem-solving performance. *Journal of Educational Psychology*. 2018. <http://dx.doi.org/10.1037/edu0000272>.
- [11] Höhne J.C., Schlosser S., Krebs D. Investigating cognitive effort and response quality of question formats in web surveys using paradata. *Field Methods*. 2017; 29 (4): 365—382.
- [12] Jbara A., Feitelson D.G. How programmers read regular code: A controlled experiment using eye tracking. *Empirical Software Engineering*. 2017; 22 (3): 1440—1477.
- [13] Kaminska O., Foulsham T. Real-world eye-tracking in face-to-face and web modes. *Journal of Survey Statistics and Methodology*. 2014; 2 (3): 343—359.
- [14] Paas F.G. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*. 1992; 84 (4): 429—434.
- [15] Paas F., Tuovinen J.E., Tabbers H., Van Gerven P.W.M. Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist*. 2003; 38 (1): 63—71.
- [16] Sharot T., Phelps E.A. How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective, & Behavioral Neuroscience*. 2004; 4 (3): 294—306.
- [17] Stodel M. But what will people think? Getting beyond social desirability bias by increasing cognitive load. *International Journal of Market Research*. 2015; 57 (2): 313—321.

DOI: 10.22363/2313-2272-2018-18-4-627-637

DEVELOPING AN APPROACH TO MULTIMODAL QUANTITATIVE ASSESSMENT OF INTERVIEWERS' COGNITIVE LOAD: FIRST RESULTS OF A FIELD QUASI EXPERIMENT*

I.F. Deviatko

National Research University Higher School of Economics
Myasnickaya St., 11, Moscow, 101000, Russia

Institute of Sociology of FCTAS RAS
Krzhizhanovskogo St., 24 / 35-5, Moscow, 117218, Russia
(e-mail: deviatko@gmail.com)

Abstract. The belief that survey research instruments mediating communication between interviewer and respondent influence the quality of data obtained in the interview has become a conventional methodological wisdom long ago. However, the impact of the cognitive load experienced by the interviewer has not been systematically examined. When a questionnaire is filled up by an interviewer it is the latter

* © I.F. Deviatko, 2018.

The research was supported by the Russian Science Foundation. Project No. 17-78-20172.
The article was submitted on 22.08.2018.

who has to allocate limited individual resources of attention, memory, visual and motor control, active listening and interpretation in order to minimize the respondent's misunderstanding of the questions and one's own errors of the answers' fixation. However, among various methods of pre-testing or evaluating survey mode effects and assessing instruments' quality, the methods of multimodal quantitative estimation of instrument-related cognitive load experienced by interviewers during the interview are still rare or lacking. Thus, the article presents a brief review of subjective, behavioral and physiological measures of the cognitive load, which are used in such disciplinary fields as cognitive science, ergonomics, etc., and a discussion of preliminary findings of the field quasi-experiment conducted among the interviewers of the Russian Longitudinal Monitoring Survey on the first stage of transition to the CAPI mode. The quasi-experiment findings prove some possibilities and limitations of the parallel use of a version of the cognitive load rating scale developed by F. Paas and a simple physiological measure (heart rate) supplemented with a background screen video capture from Android-based tablets used for CAPI interviews for the multimodal quantitative evaluation and optimization of the interviewer's cognitive load.

Key words: cognitive load; multimodal quantitative assessment of interview-related cognitive load; PAPI and CAPI; survey data quality; the Paas scale; heart rate