

---

---

## МЕТОДИЧЕСКИЙ ЭКСПЕРИМЕНТ

А.Ю. Мягков, С.Л. Журавлева, Е.Н. Прокофьев  
(Иваново)

### МОДЕЛЬ «ВЫНУЖДЕННОГО ОТВЕТА»: ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА ЭФФЕКТИВНОСТИ

Статья посвящена экспериментальному изучению модели «вынужденного ответа», предназначенной для повышения анонимности респондентов в чувствительных исследованиях. В работе представлены результаты сравнительного анализа качества данных, полученных в опросах с ее использованием и без нее. Эффективность модели оценивается по валидности ответов респондентов, уровню «отвечаемости» и субъективным оценкам опрашиваемых. Обсуждаются также возможности дальнейшего совершенствования модели.

*Ключевые слова:* метод рандомизированного ответа, модель «вынужденного ответа», чувствительные вопросы, валидационный эксперимент, качество данных, субъективная анонимность, индивидуальная валидизация.

#### *Описание проблемной ситуации*

В современной психологии доказано, что неискреннее поведение представляет собой универсальный «эволюционный

---

**Александр Юрьевич Мягков** – доктор социологических наук, профессор, заведующий кафедрой социологии Ивановского государственного энергетического университета. E-mail: socio@sociol.ispu.ru.

**Светлана Львовна Журавлева** – старший преподаватель кафедры социологии Ивановского государственного энергетического университета. E-mail: socio@sociol.ispu.ru.

**Евгений Николаевич Прокофьев** – преподаватель кафедры социологии Ивановского государственного энергетического университета. E-mail: socio@sociol.ispu.ru.

механизм, позволяющий людям получать доступ к жизненным ресурсам» [1, р. 332]. Будучи «неотъемлемой частью человеческой природы», социальная желательность характерна для всех форм опросных коммуникаций [2, р. 210, 221]. Однако в чувствительных опросах вероятность ее появления резко возрастает; усиливается и угроза качеству данных. Пропуски вопросов и социально желательные ответы – две наиболее острые проблемы, возникающие в исследованиях по деликатной проблематике [3, р. 615; 4, р. 295]. Опыт изучения индивидуально острых тем (от наркотиков и алкоголя до преступности и сексуального поведения) показывает, что классические процедуры сбора данных, в общем виде сложившиеся еще в 20–30-е гг. прошлого столетия и опирающиеся на непосредственную интеракцию с отвечающими, систематически занижают искомую долю лиц с чувствительной характеристикой и не позволяют получить истинные оценки масштабов девиаций [5, р. 269–278]. Обычные напоминания об анонимности и призывы быть искренними с интервьюером не решают этой проблемы. Многие респонденты рассматривают их как простое проявление вежливости со стороны исследователей и дань этическим нормам проведения интервью [6, р. 465–466]. Прием проекции, опирающийся на использование безличных проективных вопросов, долгое время считавшийся весьма обнадеживающим средством профилактики неискренности, на поверку также оказался малоэффективным. Как свидетельствуют специальные методические исследования, выполненные в русле когнитивного анализа, вопросы, проецирующие социально стигматизированные виды поведения на «других» людей, дают совершенно нереалистичные оценки, резко завышая масштабы распространения девиаций и внося двусмысленность в процесс интерпретации полученных результатов [7; 8, с. 202–211].

В 1960–1970-е гг. был разработан целый ряд новых приемов и процедур, специально предназначенных для повышения субъективной анонимности и стимулирования искренних ответов

респондентов. К ним относится и техника «несовпадающих чисел» [9], и процедура *bogus pipeline* [10], и *метод рандомизированного ответа* (RRT), предложенный С. Уорнером в качестве альтернативы традиционным опросным процедурам [11] и являющийся целью нашего изучения. С технической точки зрения он достаточно прост и базируется на принципе «управляемой случайности». Респондентам предлагается на выбор ответить на один из двух взаимоисключающих вопросов-суждений (типа: «Я принадлежу к группе “А”»; «Я не принадлежу к группе “А”»), имеющих прямое отношение к сенситивной теме. Такой выбор респондент осуществляет, используя специальное рандомизирующее устройство (например, вращая волчок в центре круга, поверхность которого разбита на два сектора – «А» и «Б»)<sup>1</sup>. На случайным образом выбранный вопрос, не сообщая о выпавшем жребии интервьюеру, он должен ответить либо положительно, либо отрицательно («да» – «нет», «верно» – «неверно»). При этом иные ответы, раскрывающие содержание вопроса, исключаются. Интервьюер, в свою очередь, лишь бесстрастно регистрирует формальный ответ, не подозревая, к какому из двух суждений он относится. Зная вероятность выбора респондентом того или иного вопроса (она определяется площадью соответствующего сектора, на который укажет стрелка волчка), объем выборки и общее число ответов «да» («верно»), можно оценить истинную долю лиц, принадлежащих к группе «А». Формула, отвечающая модели С. Уорнера, имеет следующий вид [11, р. 66]:

$$\hat{\pi}_A = (P - 1 + \lambda) / (2P - 1), \quad (1)$$

---

<sup>1</sup> В качестве «генератора» случайных выборов можно использовать разные средства: карточки разных цветов, разноцветные пластиковые шарики, игральные кости, обычную книгу, листая которую, испытуемый случайным образом выбирает номер страницы, денежную купюру, телефонный номер респондента или его друга. В последних случаях рандомизирующей переменной для выбора вопроса будут выступать конечные цифры номеров телефонов, серий банкнот и т.д.

где  $\hat{p}_A$  – оценка доли лиц, принадлежащих к интересующей нас группе;  $P$  – вероятность выпадения сенситивного вопроса;  $\lambda$  – общая доля ответов «да», полученных в исследовании.

Этот метод, по мнению его автора, способствует повышению анонимности благодаря рандомизации вопросов (и ответов) и формализации, обезличиванию опросного общения. Однако в своей изначальной версии он позволял получать информацию лишь по вопросам, имеющим простейшую (дихотомическую) логическую структуру. В последующие годы и десятилетия базовая идея С. Уорнера об «управлении случайностью» получила дальнейшее развитие. Появился целый ряд более изощренных моделей RRT, позволяющих задавать испытуемым не только дихотомические, но и многозначные вопросы, использовать переменные с разным уровнем измерения, вместо двух исходных применять один единственный (сенситивный) вопрос и тем самым – существенно снизить когнитивную нагрузку на респондентов<sup>1</sup>.

Следует заметить, что в нашей стране валидационные испытания моделей RRT, за редким исключением [8], не проводились, отсутствует и опыт их полевого использования. Существующие возражения против метода исследователи склонны объяснять его высокой трудоемкостью и затратностью, низким аналитическим потенциалом и ограниченностью круга решаемых задач, а также сомнительной валидностью и неочевидной эффективностью по сравнению с традиционными методами сбора данных.

В зарубежной науке *RR-модели* достаточно широко используются не только в методических [12; 13], но и в прикладных исследованиях [14; 15]. Вместе с тем, за сорок лет интенсивных экспериментов среди социологов так и не было достигнуто единства мнений относительно их эффективности. По подсчетам

---

<sup>1</sup> Подробное описание различных модификаций уорнеровской модели RRT представлено в работе [8].

Дж. Фокса и П. Трэйси, применение RRT способствовало более высокой достоверности данных лишь в пяти из девяти полевых исследований (55,0%) и в одном из четырех тестов на валидность (25,0%) [16, р. 18]. Согласно более поздним оценкам У. Юмеша и Р. Петерсона, доводы за и против RRT в современной научной литературе встречаются с одинаковой частотой [17]. Вторичный анализ данных, проведенный голландскими социологами, показал, что модели рандомизированного ответа были более эффективны примерно в трех из пяти исследований (60,0%), использовавших для проверки метод критериальной валидации [18, р. 326].

На сегодняшний день дискуссионным остается и вопрос о том, какая из ныне существующих разновидностей RRT является наиболее эффективной. На наш взгляд, оптимальная модель «рандомизированного ответа» должна удовлетворять следующим критериям.

1. Ее процедура должна быть предельно простой и понятной для респондентов. В противном случае все потенциальные преимущества, заложенные в модели, будут перечеркнуты. Это важно, поскольку многие модели RRT настолько сложны, запутаны и трудоемки, что становятся практически нереализуемыми.

2. Процедура должна быть убедительной для респондентов и не вызывать сомнений в том, что их индивидуальные ответы не смогут быть идентифицированы и разглашены.

3. Модель должна превосходить традиционные разновидности опроса с точки зрения достоверности и качества собираемых данных.

4. Модель должна быть хорошо совместимой с разными вопросно-ответными процедурами. Жесткая привязка к персональному интервью, характерная для большинства ныне существующих версий RRT, не способствует раскрытию ее потенциала.

5. Исследователи должны иметь возможность задавать не только фактуальные или поведенческие вопросы, но и вопросы аттitudного характера, касающиеся мнений, отношений, оценок и намерений.

6. Модель должна предоставлять возможность осуществления межгруппового анализа, проведения повторных исследований и т.д.

Анализ результатов специальных исследований показывает, что подавляющее большинство ныне существующих моделей RRT не соответствуют указанным критериям. Единственным исключением из общего ряда может, по-видимому, считаться модель «вынужденного ответа», изначально разработанная Р. Борухом и модифицированная впоследствии канадским исследователем Г. Бегинном и его коллегами из Лавальского университета провинции Онтарио [19]. Она является, пожалуй, наиболее востребованной и популярной среди исследователей<sup>1</sup> и, по мнению некоторых авторов, «одной из самых статистически и психологически эффективных» разновидностей RRT [3, p. 615; 20, p. 253].

Между тем такого рода оценки чаще всего носят умозрительный характер и опираются лишь на теоретические доводы и аргументы. Эмпирического материала, который бы убедительно свидетельствовал о преимуществах этой модели, явно недостаточно. К тому же в отечественной социологии модель «вынужденного ответа» никем прежде не исследовалась, поэтому сведения о возможностях ее применения в российских условиях полностью отсутствуют.

### *Общая характеристика модели*

Процедура сбора и последующего анализа данных в рамках этой модели опроса заключается в следующем. Испытуемым предлагают самостоятельно заполнить письменную стандарти-

---

<sup>1</sup> По подсчетам голландских социологов модель «вынужденного ответа» использовалась исследователями более чем в половине всех известных случаях применения RRT (56,4%); около трети применений (30,8%) приходится на долю «двухвопросной техники несвязанных вопросов» Б. Гринберга – У. Симмонса; 5,1% – на метод Э. Кука и еще 7,7% – на иные модели, включая классическую версию С. Уорнера и модель Такахаси – Сакасагавы [18, p. 334–335].

зированной анкету, содержащую ряд сенситивных вопросов. Для выбора ответа на тот или иной вопрос респондент должен подбросить вверх шестигранный кубик, не показывая выпавшего жребия анкетеру. При этом в соответствии с условиями, оговоренными в специальной инструкции, опрашиваемых просят отметить (обвести кружком) позиции «да» или «нет» (независимо от их истинных мнений), если на кубике выпадут «5» или «6» соответственно. При выпадении любых других чисел они должны отвечать максимально честно, ничего не скрывая.

Математическая модель очень проста. Для определения удельного веса лиц, обладающих сенситивной характеристикой, достаточно из общей доли утвердительных ответов по всей выборке вычесть ожидаемую долю ответов «да» (1/6) и полученную разность разделить на вероятность получения честных ответов (4/6) [19, р. 57]:

$$\hat{\pi}_s = \frac{\lambda - \theta}{P}, \quad (2)$$

где  $\lambda$  – общая доля утвердительных ответов по всей выборке;  $\theta$  – ожидаемая доля ответов «да» (1/6);  $P$  – вероятность получения честных ответов (4/6).

По сравнению с другими разновидностями RRT модель «вынужденного ответа» имеет три очевидных преимущества.

*Во-первых*, в отличие от многих прочих моделей, жестко привязанных к персональному интервью, она может успешно работать и в режиме раздаточного анкетирования. Респонденты самостоятельно выполняют все предписанные им действия, включая работу с рандомизатором, выбор ответа и заполнение письменного вопросника<sup>1</sup>.

---

<sup>1</sup> Между тем следует признать, что метод «вынужденного ответа» не полностью универсален. Его применение к телефонным интервью и почтовым опросам затруднительно. Однако абсолютно универсальных моделей RRT не существует. Таковой можно было бы считать лишь модель Такахаси – Сакасагавы (1977 г.),

*Во-вторых*, в анкете вполне допустим социально-демографический блок вопросов, на которые опрашиваемые должны ответить без использования рандомизирующего устройства. Тем самым исследователь получает возможность межгруппового анализа и применения любых методов математической статистики для работы с эмпирическим материалом. Кроме того, доступ к персональным (социально-демографическим) данным открывает возможность проведения повторных замеров в панельном режиме.

И, наконец, *в-третьих*, статистическая модель, лежащая в основе метода, позволяет резко сократить минимальный объем выборки, необходимый для его надежного применения. Если, согласно расчетам С. Уорнера, для успешной работы предложенной им техники требуется от 400 до 2000 наблюдений [11, р. 66], то для модели «вынужденного ответа», по оценке ее авторов, вполне достаточно 100 опрошенных [19, р. 57]<sup>1</sup>. При этом она остается простой, понятной и хорошо исполнимой респондентами.

---

которая, по замыслу авторов, не требует применения какого-либо специального рандомизирующего устройства, а потому может использоваться в привязке к любой из известных опросных процедур (персональному и телефонному интервью, анкетированию и почтовому опросу). Однако, судя по оценкам специалистов, указанная модель статистически ущербна и несовершенна в психологическом отношении. В методических тестах она не смогла продемонстрировать положительных результатов [21, р. 106–107], а потому не получила сколь-нибудь заметного распространения в современной исследовательской практике.

<sup>1</sup> Дело в том, что модель С. Уорнера добавляет в общую дисперсию дополнительный компонент, увеличивая тем самым размер стандартной ошибки и требуя большого объема выборки для ее нейтрализации [18, р. 321–322]. Метод «вынужденного ответа» свободен от этого недостатка. По мнению Б. Уильямса и его коллег, он привносит в результаты только случайную ошибку, связанную с ответом на сенситивный вопрос, и дисперсию, обусловленную применением рандомизирующего устройства [22, р. 234].



## *Методология и методы исследования*

В январе-феврале 2008 г. в ходе специального методического эксперимента<sup>1</sup> мы тестировали модель «вынужденного ответа», относящуюся к классу моделей RRT. В рамках нашего исследования решались три взаимосвязанные задачи: оценка эффективности данной модели по сравнению с традиционным анкетным опросом; уточнение возможностей и перспектив ее использования в условиях российской социокультурной среды; выявление субъективных оценок респондентов относительно этой разновидности RRT и особенностей ее восприятия опрашиваемыми.

**Экспериментальный план. Выборка и процедуры.** Методический эксперимент строился по принципу «разветвленного опроса», получившему в западной литературе название *split-ballot*. Такой экспериментальный план был выбран с целью сравнительного анализа ответов, полученных разными опросными методами. Были сформированы две одинаковые по объему (по 100 чел.) и по социально-демографическим характеристикам *подвыборки*.

В качестве *испытуемых* выступали студенты ивановских вузов, а также представители производственной и непромышленной интеллигенции города (учителя, врачи, инженеры, менеджеры и др.), отобранные на основе маршрутной выборки с последующим квотным отбором респондентов<sup>2</sup>.

Предлагавшийся респондентам вопросник представлял собой *тематический омнибус* и включал, помимо социально-демографического и «постэкспериментального» блоков, 23 сенси-

---

<sup>1</sup> Авторы признательны Т.В. Черновой за участие в проведении опросов и обработке эмпирических данных.

<sup>2</sup> По данным мета-анализа, предпринятого Г. Ленсвелт-Малдерс и ее коллегами и обобщающего результаты многих десятков экспериментальных работ, выполненных в мире за период с 1965-го по 2000 г., только треть исследований в области RRT (34%) основывается на случайных выборках, в то время как большинство из них опирается на выборки конформного типа (*convenience sample*) [18, p. 335].

тивных вопроса, 18 из которых были поведенческими и 5 – аттитюдными<sup>1</sup>. Они касались различных стигматизированных тем (см. табл. 1). Опросы в обеих группах проводились одновременно.

*Метод опроса* респонденты выбирали случайным образом с помощью обычного шестигранного кубика. При выпадении на нем единицы, двойки или тройки предлагалось самостоятельно заполнить анкету в обычном режиме, во всех остальных случаях применялся метод «вынужденного ответа». Тем самым образовались две совокупности респондентов, соответственно, *группа К* (контрольная) и *группа Э* (экспериментальная), объем которых оказался равным.

Процедура выбора метода исключила систематические смещения, нередко возникающие в ситуации свободного выбора респондентом метода опроса. По данным А. Вейсмана и его коллег, примерно 3/4 опрошиваемых при такой свободе выбора отдают предпочтение RRT. Это те, кому есть что скрывать. Респонденты же, не имеющие отношения к обсуждаемым в исследовании видам поведения, а потому не считающие их сенситивными, как правило, выбирают «прямые» методы опросных коммуникаций [23].

*Инструкция* для заполнения анкеты в *группе К* предьявлялась опрошиваемым в напечатанном виде. В ней кратко объяснялось, как следует работать с вопросником, и особо подчеркивалось, что все ответы останутся анонимными, а собранные исследователями данные будут использованы только в научных целях. Предварительная инструкция для *группы Э* также была напечатана – респонденты знакомились с ней самостоятельно. Кроме того, в ходе устной беседы анкетер дополнительно объяснял правила работы с рандомизатором, механизм и принципы его действия, а также специ-

---

<sup>1</sup> В качестве аттитюдных задавались вопросы типа: «Считаете ли вы для себя приемлемой практику группового секса?» Поведенческие вопросы имели традиционную формулировку и звучали, например, следующим образом: «Приходилось ли вам хоть раз в жизни попробовать наркотики?»

ально акцентировал их внимание на том, что метод гарантирует абсолютную конфиденциальность ответов.

**Способы валидации данных.** Существуют два основных подхода к оценке эффективности RR-моделей: индивидуально валидационный, получивший в специальной литературе статус «золотого стандарта» [18, р. 341], и компаративный. Если известен истинный статус каждого респондента относительно изучаемой сенситивной проблемы, то у исследователя появляется возможность сопоставить опросные оценки (ответы респондентов) с объективными (истинными) данными. Такого рода исследования обладают высокой внутренней валидностью, однако, они крайне редкие и дорогие, поскольку требуют получения труднодостижимой (а зачастую и вовсе недостижимой) информации о поведении или установках индивидов, участвующих в исследовании и, соответственно, доступа к милицейской статистике, медицинским файлам и т.д. В качестве меры эффективности RRT здесь используется «процент ошибочных ответов», который определяется как разность между истинной оценкой, принимаемой за 1 или за 100%, и опросными данными. При этом маленькая разность указывает на лучшее качество RR-модели.

В компаративном исследовании истинные оценки неизвестны социологам. Результаты RRT сравниваются с данными, полученными посредством тех или иных стандартных методов (персональных интервью, анкетирования и т.д.)<sup>1</sup>. Результаты таких сравнений обычно интерпретируются по принципу «чем

---

<sup>1</sup> По данным мета-аналитического исследования Г. Ленсвелт-Малдерс и ее коллег, результаты RRT в методических тестах чаще всего сравнивались с параллельно полученными данными персональных и телефонных интервью (по 57,9%). Примерно в каждом третьем исследовании (34,2%) сравнения производились с анкетными опросами. Анкетирование с компьютерной поддержкой (CASI) было объектом компаративного анализа в 5,3% всех обследованных авторами случаях и еще 7,9% подобных работ приходится на прочие опросные процедуры [18, р. 334].

больше, тем лучше»: более высокая случаевость сенситивной характеристики рассматривается как индикатор более валидных ответов и большей эффективности той или иной техники сбора данных. Компаративные исследования проще и дешевле индивидуально валидационных, однако, они могут быть информативны только в том случае, если их результаты столь же релевантны, как и в исследованиях с критериальной валидизацией [18, р. 323–324].

**Интерпретационная стратегия.** Из-за отсутствия индивидуальной информации о респондентах для оценки качества тестируемой модели мы использовали компаративный подход. Эффективность RRT оценивалась путем попарного сравнения долей утвердительных ответов, полученных от участников опроса посредством двух разных методов. Значимость различий определялась по  $\phi^*$ -критерию углового преобразования Фишера [24, с. 157–163, 330–332]. При этом значимые процентные превышения в пользу модели «вынужденного ответа» по отдельным вопросам интерпретировались как показатели ее превосходящей валидности. Максимально эффективным мы считали тот метод, который обеспечивал наиболее высокий удельный вес ответов «да», т.е. признаний в причастности к тем или иным формам социально неодобряемого поведения. Принимая такую модель интерпретации результатов, мы исходили из следующих соображений (см.: [25, с. 63–64; 8, с. 245–247]).

Поскольку все испытуемые закреплялись за тем или иным тестируемым методом случайным образом, то истинная доля лиц, обладающих сенситивной характеристикой, должна быть одинаковой в *группе К* и *группе Э*, а различия (в случае их появления) могут считаться функцией используемого *метода опроса*.

Если респонденты, опрошенные посредством одного метода, дадут большее число утвердительных ответов (условно *группа 1*), чем опрошенные другим методом (*группа 2*), то возможны три альтернативные схемы интерпретации различий:

1) испытуемые в обеих группах в равной мере лгут при том, что значение истинного масштаба девиаций находится в

интервале эмпирически наблюдаемых различий, т.е. в *группе 1* реальные масштабы девиаций могут быть завышены, а в *группе 2* – занижены;

2) чаще лгут представители *группы 1*, явно преувеличивая фактический уровень девиантности, а респонденты из *группы 2* дают более правдивые ответы;

3) в *группе 1* люди отвечают честнее, а в *группе 2* – склонны скрывать факты социально неодобряемого поведения.

Принятие первых двух вариантов интерпретации означает согласие с тем, что люди, отвечающие утвердительно на вопросы, касающиеся юридически преследуемых и нравственно осуждаемых видов поведения, склонны к самооговору, т.е. к добровольному признанию в совершении таких действий и поступков, которых они на самом деле не совершали. Это маловероятно, учитывая крайнюю стигматизированность и интимность большинства исследовательских вопросов. Как известно, факты сенситивного поведения значительно чаще умалчиваются или отрицаются респондентами [26, р. 99; 27, р. 266]. Тем самым, можно с большей долей уверенности утверждать, что если респонденты лгут, то делают это, отрицая свою причастность к девиантным формам поведения. Тогда правомерным следует считать третий вариант интерпретации различий: группа с повышенным уровнем утвердительных ответов более искренна, чем противоположная, а метод, обеспечивающий этот результат, является более эффективным. Процентное превышение в удельном весе признаний можно, на наш взгляд, рассматривать в качестве эмпирического индикатора эффективности того или иного *режима опроса*.

Несмотря на то, что индивидуальная валидизация несомненно сильнее с методологической точки зрения, компаративный подход, как показывают специальные исследования [18, р. 341], может считаться вполне достойной и обоснованной заменой.

### *Результаты эксперимента*

Результаты проведенного исследования свидетельствуют, что метод «вынужденного ответа» существенно эффективнее с точки зрения качества опросных данных. В 16 из 23 задававшихся вопросов (а это 69,6%) были зафиксированы статистически значимые различия (на уровне не ниже  $p \leq 0,05$ ), свидетельствующие об этом. Применение метода позволило увеличить число достоверных ответов почти в два раза: RRT-фактор<sup>1</sup>, усредненный по 16 успешным вопросам, составил 1,8. При этом средний прирост сенситивной информации, достигнутый благодаря применению модели «вынужденного ответа», по этой совокупности вопросов составил 15,0% при вариации значений по отдельным вопросам от 7 до 25%. В табл. 1 в первых двух столбцах показаны доли (в % от числа ответивших) утвердительных ответов.

По шести переменным различия в ответах, полученных двумя разными методами, статистически незначимы. В четырех из них положительные знаки также свидетельствуют в пользу модели RRT. И лишь по двум переменным (о злоупотреблении алкоголем и о добрачном половом опыте) наблюдается номинальное превосходство альтернативного метода. Процентные превышения здесь минимальны, находятся за пределами статистической значимости, поэтому могут быть интерпретированы как равенство позиций обоих методов. Следует подчеркнуть, что явного ухудшения качества данных не наблюдается ни по одной из переменных.

Общие результаты эксперимента весьма оптимистичны и указывают на более высокую валидность данных, полученных методом «вынужденного ответа». Анализ результатов исследова-

---

<sup>1</sup> Этот коэффициент рассчитывался как соотношение процентных долей утвердительных ответов, полученных в группах Э и К [ 26, р. 96]. Его значения показывают, во сколько раз чаще респонденты дают искренние ответы в RRT по сравнению с обычным анкетированием и наоборот.

Таблица 1

ХАРАКТЕРИСТИКИ МЕТОДОВ ОПРОСА

Переменные	Группа Э	Группа К	RRГ-фактор	Разница долей
<b>Поведенческие переменные:</b>				
Мелкие бытовые кражи	35,7	18,0	1,98	+17,7**
Меркантилизм в отношениях с друзьями	38,6	30,0	1,29	+8,6*
Кражи из магазина	29,7	27,0	1,10	+2,7
Злоупотребление алкоголем	45,9	47,0	0,98	-1,1
Привязанность к алкоголю	26,8	17,0	1,58	+9,8**
Сексуальные контакты в состоянии алкогольного опьянения	50,4	45,5	1,11	+4,9
Сексуальные отношения за последнюю неделю	65,1	53,4	1,22	+11,7**
Секс в общественных местах	37,5	35,7	1,05	+1,8
Супружеская неверность	48,2	35,4	1,36	+12,8**
Вынужденные сексуальные отношения («секс против воли»)	19,9	17,7	1,12	+2,2
Добрачный половой опыт	75,1	78,4	0,96	-3,3
Гомосексуальный опыт	10,1	3,0	3,37	+7,1**
Случайные сексуальные связи	56,8	39,2	1,45	+17,6**
Использование платных интимных услуг	32,4	13,3	2,44	+19,1**

\* Для  $\phi^* p \leq 0,1$ .

\*\* Для  $\phi^* p \leq 0,05$ .

Окончание табл. 1

Переменные	Группа Э	Группа К	RRГ-фактор	Разница долей
Покупка порнографических фильмов	37,8	25,5	1,48	+12,3**
Проба наркотиков	42,0	27,0	1,55	+15,0**
Несообщение о преступлении	47,5	23,0	2,06	+24,5**
Аборты (женщины)	14,0	5,2	2,69	+8,8**
<b>Античуждые переменные:</b>				
Нечестность в денежных делах	63,4	48,9	1,30	+14,5**
Практика группового секса	42,2	21,7	1,94	+20,5**
Практика орального секса	68,2	56,0	1,22	+12,2**
Отношение к абортам (мужчины)	57,3	35,3	1,62	+22,0**
Отношение к абортам (женщины)	47,8	32,9	1,45	+14,9**



ния показывает, что чем острее и деликатнее изучаемые темы, тем существеннее различия в ответах респондентов и выше эффективность модели RRT. В случае с обычным анкетированием действует обратная закономерность: валидность ответов снижается по мере роста чувствительности рассматриваемых в исследовании тем. При этом, правда, стоит заметить, что результаты RRT от остроты и деликатности задаваемых вопросов зависят в меньшей степени, чем это происходит в традиционном анкетном опросе. Этот вывод хорошо согласуется с данными исследования, проведенного голландскими социологами [18, с. 339–340].

Интересно отметить также, что модель «вынужденного ответа» продемонстрировала 100%-ную успешность в работе с аттитюдными вопросами, для всех пяти переменных существенно улучшив качество данных. По поведенческим переменным эффективность модели оказалась ниже (67% переменных).

В специальной литературе, начиная с работ Н. Брэдберна и С. Садмена, считается, что метод «рандомизированного ответа» предназначен исключительно для преодоления диссимуляции и сокращения числа преуменьшений, но при этом совершенно бессилён против преувеличений в ответах респондентов [28, с. 8]. Результаты нашего исследования показывают, что это не совсем так. Полученные нами данные свидетельствуют, что из 14 вопросов с наибольшей вероятностью преувеличений эффект браважирования девиантным поведением был обнаружен лишь в двух случаях, которые касались злоупотреблений алкоголем и добрачного полового опыта. При этом, как и следовало ожидать, этот эффект характерен исключительно для молодежи и практически не встречается в старших возрастных группах. Вместе с тем наблюдаемые процентные превышения в пользу обычного анкетирования по этим переменным крайне малы (от 1 до 3%) и незначимы в статистическом отношении. Следовательно, тестируемая нами модель RRT позволяет минимизировать

склонность респондентов к бравированию девиантностью и существенно снизить число «фальшиво позитивных» ответов. Этот вывод также хорошо согласуется с результатами некоторых предыдущих исследований [7], однако нуждается в более детальной эмпирической проверке.

Перейдем к рассмотрению другого показателя эффективности, такого как *восприятие метода респондентами*. Исследование показало, что метод «вынужденного ответа» получает значительно более высокие субъективные оценки. В частности, более 75% опрошенных в *группе Э* высказали благоприятное отношение к нему, в то время как обычное анкетирование получило лишь 46% одобрительных отзывов. К тому же, как выяснилось, испытуемые из *группы Э* воспринимали предложенные им вопросы как менее чувствительные по сравнению с респондентами *группы К*. Это свидетельствует о том, что тестируемая модель RRT более психологична для опрашиваемых и опросная ситуация характеризуется меньшей тревожностью и травматичностью, чем в обычном анкетировании.

Кроме этого, модель «вынужденного ответа» способствует уменьшению количества неответивших на отдельные вопросы. В табл. 2 представлены данные по трем типам вопросов (где  $n$  – число вопросов определенного типа) разной степени чувствительности. В *группе Э* количество пропусков и неответов почти в 4,5 раза меньше, чем в *группе К*: соответственно 14 и 62 ( $p \leq 0,000$ ).

Таблица 2

ДОЛИ НЕОТВЕТОВ НА ВОПРОСЫ РАЗНОЙ СТЕПЕНИ  
ЧУВСТВИТЕЛЬНОСТИ, % от числа всех ответов в каждой группе

Степень чувствительности	Группа Э	Группа К	Разность
Очень чувствительные ( $n = 8$ )	1,1	5,6	4,5**
Не очень чувствительные ( $n = 10$ )	0,5	1,7	1,2*
Нечувствительные ( $n = 5$ )	0,0	0,0	0,0
Все вопросы ( $n = 23$ )	0,6	2,7	2,1**

\* Для  $\phi^* p \leq 0,01$ .

\*\* Для  $\phi^* p \leq 0,000$ .

Обращение к данным табл. 2, позволяет заметить еще две интересные тенденции. *Во-первых*, они свидетельствуют, что показатели неотчетов заметно варьируют в зависимости от степени чувствительности обсуждаемых в исследовании вопросов и тем. Чем острее вопрос, тем выше вероятность его пропуска респондентами. Эта закономерность не зависит от способа сбора данных. Следовательно, неотчеты – надежный индикатор чувствительности и предиктор возможных смещений в ответах опрашиваемых. *Во-вторых*, в *группе К* задаваемые вопросы воспринимаются опрашиваемыми как более острые, деликатные и смущающие, в *группе Э*, напротив, снижается степень их остроты в восприятии респондентами. Тем самым модель «вынужденного ответа» обеспечивает респондентам большую анонимность, чем классический вариант опроса.

Высокий уровень кооперации и доброжелательное отношение к модели «вынужденного ответа» – это верные симптомы доверия к ней, показатели ее простоты и хорошей «исполнимости». Следовательно, существующие опасения по поводу дополнительной когнитивной нагрузки на респондентов, предположительно связанной с трудностями понимания и соблюдения процедуры RRT [18, р. 322], скорее всего несколько преувеличены.

За счет чего произошло существенное улучшение оценок в RRT? Главная причина, на наш взгляд, коренится в самой природе и объективных характеристиках применявшейся нами модели. Дело в том, что модель Г. Бегина относится к числу так называемых «частичных» версий RRT в отличие от «полных» версий, представленных, в частности, моделями уорнеровского типа. «Полными» называют такие модели, в которых ответы *всех* респондентов являются результатом случайного выбора. В «частичных» версиях в жеребьевке участвует лишь малая часть испытуемых (в нашем случае – не более одной трети), большинство же фактически работает в обычном режиме с установкой отвечать честно [29, р. 215–217]. Поэтому «ошибка сообщения», привносимая респондентами в

связи с вынужденным искажением их истинных ответов, здесь значительно ниже, чем в иных моделях RRT.

### *Ограничения модели*

Следует признать, что идеальных результатов мы не получили. Преимущества RRT очевидны, но не безоговорочны. Методика «вынужденного ответа» заметно сокращает, но не элиминирует полностью ошибки сообщения. У 30% из всех задававшихся нами вопросов (7 из 23) не было обнаружено различий, явно свидетельствующих в пользу RRT; около 25% опрошенных из группы Э сочли альтернативный способ сбора данных менее привлекательным, чем обычное письменное анкетирование с «самозаполнением». Эта ситуация обусловлена, на наш взгляд, несколькими причинами.

*Во-первых*, несмотря на то, что RR-модели надежно защищают частную жизнь опрашиваемых и обеспечивают конфиденциальность их ответов, тем не менее некоторые респонденты из группы Э все равно не верят в гарантии своей анонимности и не доверяют анкетерам. При этом они полагают, что у исследователей есть какие-то скрытые механизмы идентификации отвечающих и не считают возможным давать искренние ответы на сенситивные вопросы социологов. По данным исследования П. ван дер Хейдена и его коллег, только 72% респондентов полностью доверяли модели «вынужденного ответа», в то время как 22% – лишь отчасти и еще 4% не доверяли совсем [30, р. 520]. Недоверие респондентов к RRT, неизбежно вызывающее редактирование ответов, может быть дополнительным источником ошибок и смещений в результатах проведенного исследования.

*Во-вторых*, как и во многих других аналогичных исследованиях, в нашем эксперименте участвовали представители социально благополучных слоев населения. В большинстве своем это были законопослушные и не склонные к девиациям люди, не причаст-

ные к тем видам поведения, которые изучались в ходе исследования. Им нечего было скрывать от анкетеров и исследователей, поэтому они отвечали честно вне зависимости от способа сбора данных. Это в значительной мере могло способствовать выравниванию результирующих показателей, полученных с помощью двух использованных методов, и нейтрализации существующих различий.

*В-третьих*, отсутствие ожидаемых результатов связано также с некоторыми объективными характеристиками самой тестируемой модели и, в частности, с установленными для респондентов «правилами игры». Очевидно, что в интересующей нас разновидности RRT озвучиваемый респондентами ответ зависит как от их истинного статуса (мнений или поведенческого опыта), так и от выпавшего им жребия. При определенном исходе событий испытуемые вынуждены давать искаженные, т.е. противоположные истинным, ответы. И если ответы «нет» в целом психологически комфортны для опрашиваемых, то утвердительные являются источником сильной идиосинкразии, поскольку сопряжены с необходимостью вынужденной самоинкриминации. В результате респонденты используют защитные стратегии поведения, ограждая себя от дискомфорта, связанного с самооговором. Как показывают специальные исследования, до четверти всех опрашиваемых (25%), которые должны в соответствии с выпавшим жребием отвечать «да» на чувствительные вопросы, дают отрицательные ответы, нарушая тем самым установленные «правила игры» и не соблюдая процедурные требования инструкций [27, р. 267].

Вместе с тем и само наше исследование также было не-свободно от ряда *ограничений* организационно-методического характера. *Первое* из них связано, возможно, с объемом выборки. Ста опрошенных посредством RRT даже для предварительного методического эксперимента, скорее всего, недостаточно. Тем более, что, согласно расчетам голландских социологов, при применении модели «вынужденного ответа» выборочная совокупность долж-

на почти вдвое превышать по численности ту, что считалась бы приемлемой для опроса с использованием стандартных методов сбора данных [20, p. 261].

*Второе* ограничение касается способов отбора респондентов. Известно, что все разновидности RRT основаны на вероятностных (статистических) моделях, что в свою очередь предполагает использование тех или иных видов случайных выборок. Между тем в нашем исследовании выборка не полностью соответствует этому требованию. Присутствие в ней элементов псевдослучайного отбора вполне могло ослабить статистические закономерности и смазать ожидаемые результаты.

*Третий* фактор, накладывающий ограничение на интерпретацию полученных данных, имеет отношение к специфике эмпирической базы и объекта нашего эксперимента. Обращение к социально благополучным слоям населения и как следствие – смещение в социальной структуре выборочной совокупности также могло сказаться на конечных результатах исследования и привести к элиминации реально существующих различий в ответах, полученных двумя разными методами.

И, наконец, еще одно обстоятельство, которое также могло повлиять на итоговые результаты нашего эксперимента. Известно, что эффективность RRT во многом зависит от выбранного исследователями соотношения вероятностей. Г. Ленсвелт-Малдерс и ее коллеги экспериментально доказали, что оптимальное значение вероятности получения честного ответа ( $P$ ) для моделей Р. Боруха и Г. Бегина должно находиться в диапазоне 0,7–0,85 [20, p. 261–263]. При приближении  $P$  к 1 резко возрастает риск нарушения субъективной анонимности респондентов; когда же наблюдается противоположная тенденция ( $0,5 \leftarrow P$ ), появляется опасность снижения эффективности опросной модели с точки зрения баланса требуемых затрат и валидности получаемых оценок [20, p. 262]. В нашем исследовании показатель вероятности честных ответов был установлен ниже этого критериального значения, что, вероят-

но, отразилось и на качественных характеристиках проведенного нами эксперимента. Вероятность получения честных ответов функционально связана «с правилами игры», т.е. с заданием для респондентов. У нас она была зафиксирована на уровне 0,666 (ибо связана с выпадением граней 1-4).

На основании вышеизложенного можно сделать вывод, что результаты проведенного исследования показывают, что модель «вынужденного ответа» имеет серьезные преимущества перед классической формой анкетирования. И хотя в эксперименте она не смогла продемонстрировать стопроцентного результата, тем не менее, опыт ее практической апробации можно считать весьма успешным.

Модель позволяет получить более высокие оценки «случаемости» изучаемых сенситивных характеристик, заметно сократить число пропусков и неответов на деликатные вопросы, существенно улучшить кооперацию с респондентами как на предварительной стадии исследования, так и в ходе самого опроса. Тем самым она способствует повышению общей валидности данных, характеризующих масштабы социальных девиаций, что позволяет считать ее вполне адекватной при изучении вопросов и тем, отягощенных бременем социальной стигматизации.

Модель «вынужденного ответа» достаточно проста, психологична и убедительна для респондентов, не провоцирует чрезмерного числа ошибок, не требует больших объемов выборки, необходимых для нейтрализации возникающих смещений. Кроме того, она эффективна с точки зрения качества получаемых данных, хорошо совместима с традиционными методами опроса и имеет широкие аналитические возможности.

Совершенствование исследовательского дизайна за счет увеличения объема выборки, применения случайных методов отбора респондентов, использования оптимальных рандомизаторов позволит в дальнейшем уточнить сделанные нами выводы относительно эффективности модели «вынужденного ответа».

ЛИТЕРАТУРА

1. Hogan R. In Defense of Personality Measurement: New Wine for Old Whiners // Human Performance. 2005. Vol. 18. No. 3. P. 331–341.
2. Dilchert S., Ones D.S., Viswesvaran C., Deller J. Response Analysis in Personality Measurement: Born to Deceive, Yet Capable of Providing Valid Self-assessments? // Psychology Science. 2006. Vol. 48. No. 3. P. 209–225.
3. Hox J., Lensvelt-Mulders G. Randomized Response Analysis in Mplus // Structural Equation Modeling. 2004. Vol. 11. No. 4. P. 615–620.
4. Carr J.W., Marascuilo L.A., Busk P. Optimal Randomized Response Models and Methods for Hypothesis Testing // Journal of Educational and Behavioral Statistics. 1982. Vol. 7. No. 4. P. 295–310.
5. Tourangeau R., Rips L., Rasinski K. The Psychology of Survey Response. Cambridge, Engl.: Cambridge Univ. Press, 2000.
6. Methods of Data Collection, Perceptions of Risks and Losses, and Motivation to Give Truthful Answers to Sensitive Survey Questions / K.A. Rasinski, G.B. Willis, A.K. Baldwin et al. // Applied Cognitive Psychology. 1999. Vol. 13. No. 4. P. 465–484.
7. Begin G., Boivin M. Comparison of Data Gathered on Sensitive Questions via Direct Questionnaire, Randomized Response Technique, and a Projective Method // Psychological Reports. 1980. Vol. 47. No. 7. P. 751–760.
8. Мягков А.Ю. Искренность респондентов в чувствительных опросах: Методы диагностики и стимулирования / Иван. гос. энэрг. ун-т. Иваново, 2007.
9. Dalton J.R., Wimbush J.C., Daily C.M. Using the Unmatched Count Technique (UCT) to Estimate Base Rates for Sensitive Behavior // Personnel Psychology. 1994. Vol. 47. No. 8. P. 817–828.
10. Jones E.E., Sigall H. The Bogus Pipeline: A New Paradigm for Measuring Affect and Attitude // Psychological Bulletin. 1971. Vol. 76. No. 3. P. 349–364.
11. Warner S.L. Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias // Journal of the American Statistical Association. 1965. Vol. 60. No. 309. P. 63–69.
12. Guerriero M., Sandri M.F. A Note on the Comparison of Some Randomized Response Procedures // Journal of Statistical Planning and Inference. 2007. Vol. 137. No. 7. P. 2184–2190.
13. Lensvelt-Mulders G.J., Boeije H.R. Evaluating Compliance with a Computer Assisted Randomized Response Technique: A Qualitative Study into the Origins of Lying And Cheating // Computers in Human Behavior. 2007. Vol. 23. No. 1. P. 591–608.
14. Himmelfarb S. The Multi-item Randomized Response Technique // Sociological Methods and Research. 2008. Vol. 36. No. 4. P. 495–514.
15. Bockenholt U., Barlas S., Heijden P. van der. Do Randomized Response Designs Eliminate Response Biases?: An Empirical Study of Non-compliance Behavior // Journal of Applied Econometrics. 2009. Vol. 24. No. 3. P. 377–392.



16. *Fox J.A., Tracy P.E.* Randomized Response: A Method for Sensitive Surveys. Beverly Hills (Calif.): Sage Publications, 1986.
17. *Umesh U.N., Peterson R.A.* A Critical Evaluation of the Randomized Response Method: Applications, Validation, and Research Agenda // *Sociological Methods and Research*. 1991. Vol. 20. No. 1. P. 104–138.
18. *Lensvelt-Mulders G.J., Hox J.J., Heijden P. van der, Maas C.* Meta-analysis of Randomized Response Research: Thirty-five Years of Validation // *Sociological Methods and Research*. 2005. Vol. 33. No. 3. P. 319–348.
19. *Begin G., Boivin M., Bellerose J.* Sensitive Data Collection Through the Random Response Technique: Some Improvements // *Journal of Psychology*. 1979. Vol. 101. No. 1. P. 53–65.
20. *Lensvelt-Mulders G.J., Hox J.J., Heijden P. van der.* How to Improve the Efficiency of Random Response Designs // *Quality and Quantity*. 2005. Vol. 39. No. 2. P. 253–265.
21. *Beldt S.F., Daniel W.W., Garsha B.S.* The Takahasi-Sakasegawa Randomized Response Technique: A Field Test // *Sociological Methods and Research*. 1982. Vol. 11. No. 1. P. 101–111.
22. *Williams B.L., Suen H.K., Baffi C.R.* A Controlled Random Response Technique // *Evaluation and the Health Professions*. 1993. Vol. 16. No. 2. P. 225–245.
23. *Weissman A.N., Steer R.A., Lipton D.S.* Estimating Elicit Drug Use Through Telephone Interviews and the Randomized Response Technique // *Drug and Alcohol Dependency*. 1986. Vol. 18. No. 2. P. 225–233.
24. *Сидоренко Е.В.* Методы математической обработки в психологии. СПб.: Речь, 2001.
25. *Мягков А.Ю.* Техника «рандомизированного ответа»: Опыт полевого тестирования // *Социологический журнал*. 2002. № 4. С. 60–77.
26. *Donovan J.J., Dwight S.A., Hurtz G.M.* An Assessment of the Prevalence, Severity, and Verifiability of Entry-level Applicant Faking Using the Randomized Response Technique // *Human Performance*. 2003. Vol. 16. No. 1. P. 81–106.
27. *Cruyff M.J., Hout A. van den, Heijden P. van der, Böckenholt U.* Log-linear Randomized-response Models Taking Self-protective Response Behavior into Account // *Sociological Methods and Research*. 2007. Vol. 36. No. 2. P. 266–282.
28. *Bradburn N.M., Sudman S.* Improving Interview Method and Questionnaire Design. San Francisco: Jossey-Bass, 1979.
29. *Thornton B., Gupta S.* Comparative Validation of a Partial (Versus Full) Randomized Response Technique: Attempting to Control for Social Desirability Response Bias to Sensitive Questions // *Individual Differences Research*. 2004. Vol. 2. No. 3. P. 214–224.
30. *Heijden P. van der, Gils G. van, Bouts J., Hox J.J.* A Comparison of Randomized Response, Computer-assisted Self-interview, and Face-to-face Direct Questioning: Eliciting Sensitive Information in the Context of Welfare and Unemployment Benefit // *Sociological Methods and Research*. 2000. Vol. 28. No. 4. P. 505–537.