

ШКАЛЫ ЛЖИ: ОПЫТ СОЦИОЛОГИЧЕСКОЙ РЕИНТЕРПРЕТАЦИИ

МЯГКОВ А.Ю., д-р социол. наук

Представлены результаты экспериментальной проверки шкал социальной желательности из теста MMPI (L и K) и шкалы лжи из опросника ЕРІ Г.И. Айзенка на конструктивную валидность и надежность измерения.

Ключевые слова: шкала лжи, социологическая реинтерпретация, социальная желательность

FALSEHOOD SCALES: THE ATTEMPT OF SOCIOLOGICAL REINTERPRETATION

MYAGKOV A.Yu., Ph. D.

The article gives the experimental testing results of social desirability scales according to MMPI (L and K) and falsehood scale of Izenk's questionnaire on construct validity and measurement reliability.

Key words: falsehood scales, sociological reinterpretaciya, social desirability.

Постановка проблемы

В современной социологии и психологии использование шкал лжи, заимствованных из различных личностных опросников, имеет давнюю традицию. Мало кто из исследователей сегодня рискнул бы усомниться в их пригодности для диагностики неискренних ответов респондентов. Статистические оценки, свидетельствующие о высокой надежности и валидности шкал лжи (социальной желательности, потребности в одобрении и т.д.) и подкрепляющие тем самым убеждение исследователей в безупречности этих инструментов, неоднократно приводились в специальной литературе. Однако работы, посвященные анализу идентификационных механизмов шкал лжи и переоценке их диагностических возможностей, по-прежнему крайне редки и малочисленны.

В западной исследовательской практике на протяжении многих десятилетий наиболее популярной остается «шкала потребности в одобрении» (МС), разработанная американскими психологами Д. Марлоу и Д. Крауном еще в 1960-е годы в рамках исследований, проводившихся ими с целью объяснения низкой предсказательной способности личностных тестов [1]. Многие специалисты и по сей день используют ее для контроля искренности респондентов и отсеивания социально желательных ответов в опросах по сенситивной проблематике [2–4]. Не менее известной является и шкала А. Эдвардса [5].

В отечественной социологии и психологии в целях улучшения качества опросных данных чаще всего применяются шкалы лжи из тестов Айзенка [6], из опросника Тейлор [7, с. 221–224], а также из Миннесотского многофазного личностного перечня (MMPI) [8]. В работах многих психологов они оцениваются как высоко эффективное, хорошо разработанное психодиагностическое средство контроля за искренностью ответов опрашиваемых, позволяющее «автоматически отсеивать недостоверные протоколы» [9, с. 20]. Социологи также считают шкалы лжи «наиболее обоснованной методикой» измерения искренности респондентов [10, с. 17]. На этом основании они рекомендуют включать их в социологические вопросники либо в полном, либо в сокращенном варианте особенно в тех случаях, когда исследователю предстоит работать с группами «повышенного риска»

(правонарушителями, наркоманами, трудными подростками и т.д.).

Вместе с тем в современной социологии известны лишь единичные случаи экспериментальной проверки шкал лжи на надежность и валидность измерений. Причем результаты этих исследований в целом неутешительны.

Предшествующие исследования

Наиболее серьезную попытку анализа шкал лжи предприняли К. Стокинг, Н. Брэдберн и С. Садман, которые в конце 1970-х гг. провели специальное исследование по валидизации шкалы, созданной Д. Марлоу и Д. Крауном для идентификации индивидов, подделывающих истинные реакции на вопросы личностных тестов [11, р. 85–106]. В ходе исследования Н. Брэдберн и С. Садман попытались выяснить, что же все-таки измеряет эта шкала: склонность респондентов к самопрезентации в расчете на одобрение со стороны интервьюера или специфические характеристики их действительного образа жизни, сознания и поведения, ошибочно принимаемые большинством экспериментаторов за проявление социальной желательности. В качестве эмпирических индикаторов второй группы переменных авторы исследования использовали вопросы, касающиеся употребления алкоголя и наркотиков, способов проведения досуга, частоты социальных контактов и т.п. В результате они пришли к выводу, что высокие баллы по шкале лжи чаще получают индивиды, имеющие жесткие стандарты оценок собственного поведения и поступков других людей и неукоснительно следующие в повседневной жизни нормам традиционной консервативной (и даже пуританской) морали. Поэтому вполне возможно, что, отвечая на тестовые вопросы необычным для большинства людей образом (например, «Я никогда не обманываю в игре»), они на самом деле не лгут, а выражают свои истинные установки, сообщают факты реального поведения, отличаясь тем самым от индивидов, имеющих релятивистские ценности и поведенческие стандарты. Исследователи пришли к выводу, что испытуемые с высокими баллами по МС-шкале менее социально активны (–0,218), имеют ограниченные социальные контакты, ведут более замкнутый образ жизни (–0,212), значительно реже других употребляют алкоголь (–0,279) и наркотики (–0,295) [11, р. 98].

«Шкала Марлоу и Крауна, – резюмируют Н. Брэдберн и С. Садман, – крайне интересная переменная, но она не может быть полезной для идентификации людей, склонных искажать свои ответы в опросных исследованиях» [11, р. 106].

Неутешительные данные были получены также относительно шкалы социальной желательности Эдвардса (ESDS). В результате специальных исследований, проведенных в 1980-е гг. с целью валидации тестовых методик, предназначенных для измерения различных аспектов качества жизни, выяснилось, что ESDS плохо диагностирует ситуативную ложь в ответах респондентов. Высокий уровень корреляций между значениями шкал «психологического благополучия» и социальной желательности (0,58–0,70), по мнению А. Козмы и М. Стоунса [12; 13], объясняется не фальсификацией испытуемыми ответов на тесты «социального самочувствия», а значительным сходством содержания вопросов, задаваемых в обоих типах шкал. К тому же, как оказалось, шкала Эдвардса намного слабее коррелирует со шкалой Марлоу и Крауна (0,38–0,42), чем со шкалами качества жизни (0,58–0,82) [13, р. 6]. В результате А. Козма и М. Стоунс приходят к выводу, что конструкт, измеряемый с помощью ESDS, правильнее было бы определить как «благополучие», а не как «социальную желательность». Шкала Эдвардса, по их мнению, вряд ли может служить хорошей мерой смещений в ответах респондентов; она более подходит для измерения социального самочувствия индивидов, чем их склонности к самопрезентации [13, р. 10].

Серьезные сомнения в способности шкал лжи измерять тенденцию респондентов к социальной желательности высказывают сегодня и многие другие авторы [14, р. 317; 15].

Методология и методы исследования

Объекты исследования. Поскольку вопрос о диагностических возможностях шкал лжи остается открытым, мы решили выяснить, в какой мере данные инструменты соответствуют своему функциональному назначению. В качестве объектов для изучения были выбраны шкалы L и K из опросника MMPI, все чаще используемого в последнее время в отечественных психологических, педагогических и этносоциологических исследованиях [8], а также знаменитая шкала лжи из теста EPI Г.И. Айзенка.

Методическая экспертиза L-шкалы проходила в марте 2001 г. (N = 387, экспериментальная выборка – из взрослого населения г. Иванова; метод сбора данных – персональное формализованное интервью по месту жительства или работы респондентов). K-шкала проверялась на надежность в методическом эксперименте, предпринятом в январе 2002 г. (N = 182, экспериментальная выборка – студенты ИГЭУ, метод сбора данных – групповое очное анкетирование по месту учебы испытуемых). Тестирование шкалы Айзенка осуществлялось в рамках специального исследования, проведенного нами среди жителей областного центра в феврале 2002 г. в режиме face-to-face интервью (N = 200, выборка – квотная). Статистико-математическая обработка данных всех трех иссле-

дований производилась в программно-аналитическом комплексе SPSS.*

L-шкала из MMPI считается одной из самых надежных с диагностической точки зрения и ориентирована на фиксацию и измерение уровня социальной желательности. Она состоит из 15 вопросов-суждений, предполагающих однозначную вербальную реакцию («согласен – не согласен», «верно–неверно» и т.п.) на ситуации, отражающие житейские виды поведения, которые, как может показаться, осуждаемы строгой моралью [16; 17]. Большинство людей обычно легко признают в себе те незначительные слабости и недостатки, которые фиксируются в вопросах, однако индивиды, намеренно стремящиеся подать себя в выигрышном свете, отрицают их, отвечая не так, «как есть», а так, «как принято». В результате они получают высокие баллы по шкале, являющиеся основанием для отбраковки вопросников.

K-шкала (или «шкала коррекции»), состоящая из 30 пунктов, также направлена на измерение установок опрашиваемых на намеренное и неосознанное улучшение результатов исследования. В нее включены утверждения, фиксирующие часто встречающиеся в жизни проявления, которые могут восприниматься испытуемыми как признаки болезни или какого-то личностного недостатка [16]. Человек, желающий показаться здоровым или благополучным, часто отрицает наличие у себя таких симптомов, привычек, реакций на окружающие обстоятельства. K-шкала считается более эффективным вследствие большей завуалированности вопросов-суждений инструментом диагностики неискренности, поскольку, по мнению специалистов, она лучше, чем L-шкала фиксирует неосознанный контроль поведения и неосознаваемую самоидентификацию индивидов с социально желательным образом вследствие повышенной конформности. Получаемые по ней высокие баллы исследователи связывают с защитным подходом к заполнению теста, в то время как низкие – с прямой и самокритичностью опрашиваемых.

Шкала Айзенка (форма «Б») состоит из 9 тестовых вопросов и призвана измерять склонность индивидов к управлению имиджем [18, с. 121–124]. По мнению исследователей, она фиксирует сознательное стремление людей отрицать присущие им «социально девiantные импульсы» [19, р. 1027] и тем самым создавать у окружающих впечатление о себе как о нравственных, добродетельных, законопослушных и социально ориентированных гражданах [20, р. 371–372]. Эту тенденцию Д. Паулюс и О. Джон называют «моралистической» (в отличие от «эгоистической»), поскольку она отражает преувеличенное желание индивида «видеть себя хорошим человеком, неизменно следующим социальным нормам и избегающим антиобщественного поведения» [19, р. 1038–1039].

Меры надежности. Поскольку шкалы лжи представляют собой классический образец тестовых методик, то для статистической оценки их надежности мы использовали принятые в тестологии стандартные процедуры.

* Помимо автора в этих исследованиях участвовали также А.Н. Алексеичева, С.Л. Журавлева, А.О. Морозова, Н.А. Потемина и О.В. Симакова.

Во-первых, проводился анализ интеркорреляций между отдельными пунктами шкал. В связи с тем, что коррелируемые переменные во всех трех шкалах измерены на номинальном уровне и имеют дихотомическую природу («верно-неверно»), наиболее адекватной мерой корреляции мы посчитали коэффициент контингенции, используемый обычно для измерения абсолютной связи в четырехклеточных таблицах.

Во-вторых, изучались корреляции отдельных пунктов теста с «индексом социальной желательности», построенным в результате подсчета общей суммы баллов, полученной всеми респондентами по всем вопросам шкалы. Поскольку одна из переменных в каждом из наших случаев является номинальной, а вторая – интервальной, для характеристики силы связи между ними использовались Eta-коэффициенты, как наиболее точно отвечающие природе и специфике измерения.

В-третьих, интересующие нас тесты проверялись на внутреннюю согласованность посредством коэффициента Альфа Кронбаха. Данный коэффициент, как известно, представляет собой оценку надежности, базирующуюся на гомогенности шкалы, и рассчитывается как сумма корреляций между ответами испытуемых на вопросы внутри одной и той же тестовой формы.

В-четвертых, производилась оценка «надежности–согласованности» тестов. С этой целью осуществлялся расчет коэффициентов Спирмена-Брауна в рамках статистической модели «split half».

Ретестовая надежность указанных шкал не проверялась.

Методы валидации. Проверка L-шкалы и шкалы лжи Айзенка на конструктивную (конвергентную и дискриминантную) валидность производилась с использованием процедур, базирующихся на идеях Д. Кэмпбела и Д. Фиске [21].

Конвергентная валидность устанавливалась посредством коррелирования «индекса социальной желательности», построенного на аналитической стадии исследования по каждой из двух тестируемых шкал, с «эталонной шкалой лжи», состоящей из серии контрольных вопросов, также фиксирующих неискренность, но иным, несомненно, более надежным образом. В первом случае контрольная шкала состояла из 16 пунктов, во втором – из 10. При этом мы полагаем, что при обнаружении высоких значимых корреляций между указанными переменными можно будет считать доказанным факт, что шкалы лжи действительно измеряют именно тот социально-психологический конструкт, для фиксации которого они изначально создавались.

При установлении дискриминантной валидности использовалась процедура одновыборочного *t*-теста для независимых подвыборок с целью анализа средних баллов, набранных по шкалам лжи искренними и неискренними респондентами, получившими такую квалификацию на основе ответов на «вопросов-ловушки». Статистическая значимость различий между средними определялась с учетом теста Ливиня для условия о равенстве дисперсий. В данном случае мы исходили из предположения о том, что если по всем или хотя бы по большинству сравниваемых пар средних будут получены значимые положительные

t-оценки, то тестируемые шкалы лжи могут считаться инструментами, достоверно дифференцирующими группы искренних и неискренних респондентов, а потому обладающими достаточно высокой дискриминантной валидностью. Подробное обоснование правомерности изложенных гипотез и релевантности описанных аналитических процедур представлено в нашей работе [17].

Результаты и их обсуждение

Показатели надежности. Результаты исследований показывают, что уровни надежности всех трех анализируемых нами шкал, измеренные по разным критериям, явно не дотягивают до минимально приемлемых значений (табл. 1).

Таблица 1. Показатели надежности шкал лжи, коэффициенты корреляции

Показатели надежности	L-шкала	K-шкала	Шкала Айзенка
Средний коэффициент интеркорреляций	0,094	–	0,083
Средний коэффициент корреляции пунктов с общей суммой баллов	0,330	0,271	0,349
Коэффициент Альфа Кронбаха	0,4441	0,4744	0,1657
Коэффициент Спирмена-Брауна	0,3324	0,5585	0,4376

Данные, представленные в табл. 1, свидетельствуют о том, что самый высокий коэффициент Альфа Кронбаха (0,4744) зафиксирован у шкалы «коррекции», однако необходимым порогом надежности для α считается показатель, равный 0,6 [20, р. 377].

«Надежность–согласованность», измеренная посредством процедуры *split-half*, также крайне невысока. Наивысшее значение коэффициента Спирмена-Брауна (для той же шкалы), по расчетам, составило всего 0,5585, в то время как данный показатель, по мнению специалистов, не должен быть ниже 0,8 [22, с. 184].

Средние коэффициенты Eta, полученные в результате корреляции отдельных пунктов шкал с итоговой суммой баллов, по всем трем тестам следует считать весьма умеренными. Учитывая естественный диапазон их вариации (от 0 до 1), можно говорить лишь о слабой взаимосвязи между указанными переменными и, соответственно, о недостаточной согласованности надежности анализируемых шкал.

Анализ интеркорреляций также убеждает, что надежность всех трех шкал лжи находится на совершенно неудовлетворительном уровне.

Исходя из закона основного психометрического соотношения, согласно которому валидность теста не может быть выше его надежности [23, с. 62], можно с достаточной долей уверенности утверждать, что все три шкалы не являются валидными инструментами для измерения социальной желательности. Полученные данные позволяют заключить, что предельно возможные показатели валидности для анализируемых тестов, определяемые как корень квадратный из надежности [20, р. 373], варьируют в диапазоне 0,41–0,69 и в лучшем случае составляют лишь 2/3 от своего теоретического максимума.

Оценки валидности. Валидации подвергались две шкалы – L из MMPI и шкала лжи из теста Айзенка.

L-шкала. При установлении конструктивной валидности данной шкалы проверялось наличие значимых корреляций между переменными, фиксирующими соотношение искренних и неискренних ответов респондентов на 16 контрольных вопросов анкеты, и общей суммой баллов, набранных ими по L-шкале. При этом мы исходили из предположения, что если шкала на уровне индекса высоко коррелирует с «эталонными» оценками неискренности, полученными посредством контрольной техники, то она действительно измеряет тот социально-психологический конструкт, для фиксации которого она изначально создавалась.

В результате проведенного анализа оказалось, что лишь 5 из 16 тестируемых зависимостей были статистически значимыми на уровне не ниже 0,05. Из четырех «эталонных» переменных только две значимо коррелировали с суммарным «индексом социальной желательности». Все остальные проведенные нами тесты дали основание для отрицательных заключений. Вместе с тем следует заметить, что даже в тех случаях, когда наблюдались значимые связи, их интенсивность была незначительной: в среднем для 5 переменных, коррелирующих с суммарным баллом, значение η^2 -коэффициента составило лишь 0,206 при максимальном значении, равном 0,253, и минимальном – 0,159. В принципе, та же картина наблюдается и для двух значимо коррелирующих «эталонных» вопросов (см. подр. [17, с. 124]).

Проверка на дискриминантную валидность показала, что лишь в половине всех тестируемых сравнений имеются статистически значимые различия между средними баллами, набранными искренними и неискренними респондентами по шкале лжи из MMPI (см. об этом [17, с. 125]). Следовательно, вероятность правильной идентификации респондентов, склонных к социальной желательности, для данного инструмента не превышает 50%. Фактически это означает, что производимые по L-шкале измерения достигают своей изначальной цели лишь в каждом втором случае. Однако более внимательный анализ данных показывает, что даже эти крайне неутешительные показатели являются сильно завышенными.

В этой связи обращает на себя внимание большое число отрицательных t -статистик. В 9 из 16 всех сравниваемых пар средних и в 5 из 8 случаев значимых различий между ними t -оценки свидетельствуют о наличии обратных тенденций, идущих вразрез с теми, которые могли бы быть предсказаны, исходя из гипотезы о том, что высокие баллы по L-шкале характерны для респондентов, склонных к самопрезентации. То есть, по сути дела, шкала лжи из MMPI позволяет достоверно дифференцировать группы искренних и неискренних испытуемых только в 3 из 16 тестируемых нами случаев. К тому же в достоверности одного из них имеются большие сомнения, поскольку в группе неискренних респондентов здесь оказалось всего 4 человека. Следовательно, вероятность безошибочных квалификаций, основанных на результатах L-тестирования, на самом деле не превышает 19%. Во всех остальных случаях шкала либо необоснованно пропускает большое количество

неискренних респондентов, либо незаслуженно отсеивает «достоверных информаторов» во внетестовых ситуациях. При этом последняя из указанных тенденций встречается, по-видимому, значительно чаще (см. подр. [17, с. 125]).

Эта гипотеза находит, похоже, и еще одно статистическое подтверждение. Проведенный нами корреляционный анализ свидетельствует о существовании слабой обратной взаимосвязи между двумя интегральными индексами: индексом «эталонной лжи» и индексом «социальной желательности», измеренной по L-шкале. Коэффициент корреляции Пирсона (R) составляет $-0,091$ при $p = 0,074$. Следовательно, чем большее число раз респондент «попался» на контрольных вопросах, тем ниже его балл по L-шкале и, соответственно, меньше вероятность того, что его интервью будет признано невалидным. И наоборот, чем искреннее человек отвечает на «эталонные» вопросы, тем больше у него шансов не пройти испытание по шкале из MMPI.

Шкала лжи из теста Айзенка. Для проверки конструктивной валидности данной шкалы мы предприняли конвергентную и дискриминантную валидацию посредством тех же методов, что и в случае с L-шкалой из MMPI.

В результате измерений, проведенных по шкале лжи Айзенка, не было идентифицировано ни одного согласованного респондента, который бы набрал в сумме 6 «пороговых» баллов и выше. При этом 10,5% опрошенных получили от 3 до 5 баллов, 19,5% – по 2 балла, 35,5% – по 1 баллу и 34,5% набрали 0 баллов. Между тем применение контрольной (эталонной) шкалы из 10 вопросов дало иные результаты: в среднем удельный вес неискренних ответов составил 18,2% при вариационном размахе от 4,1 до 43,8%. Интересно заметить, что в нашем эксперименте с L-шкалой из MMPI использование 16-пунктного эталона позволило выявить 18,6% неискренних респондентов в аналогичном по содержанию вопроснике. Это свидетельствует о внешней валидности использованных контрольных шкал.

Анализ корреляций между контрольными переменными и итоговой суммой баллов, полученной по шкале лжи, показал, что лишь две из тестируемых нами зависимостей являются статистически значимыми для условия $p \leq 0,05$. Среднее значение коэффициента η^2 крайне низкое и составляет всего 0,0652 при вариации частных показателей в диапазоне от 0,004 до 0,196.

Не было обнаружено и зависимости между двумя индексами: индексом «эталонной лжи» (по всем 10 контрольным вопросам) и индексом «социальной желательности», измеренной на основе суммы баллов по тесту Айзенка: коэффициент Пирсона $R = 0,100$ при $p = 0,160$. Это значит, что интересующая нас шкала лжи недостаточно валидна в конвергентном аспекте.

Анализ результатов t -теста, проведенного с целью установления дискриминантной валидности шкалы Айзенка, показывает, что лишь в 2 из 10 тестируемых сравнений имеются статистически значимые различия между средними баллами, набранными искренними и неискренними испытуемыми по шкале лжи. Следовательно, вероятность безошибочных ква-

лификаций в данном случае примерно такая же, как и у L-шкалы из MMPI, и составляет всего 20%. В остальных 80% случаев шкала лжи из теста Айзенка пропускает лгущих респондентов, квалифицируя их как искренних. Об этом свидетельствуют и тот факт, что 4 из 10 полученных t-статистик имеют отрицательный знак. Таким образом, полученные данные свидетельствуют о явно недостаточной дискриминантной валидности тестирующей шкалы лжи.

Идентификационные механизмы L-шкалы.

Для уяснения механизмов, на основе которых интересующая нас шкала позволяет идентифицировать социально желательные ответы, мы использовали две эмпирические процедуры. Во-первых, мы сравнили доли респондентов, чьи интервью были забракованы по результатам L-тестирования и по «эталонной» шкале лжи, а во-вторых, провели анализ кростабуляций, где в качестве взаимосвязанных переменных выступали 16 контрольных вопросов и итоговая сумма баллов по шкале из MMPI.

При сравнении результатов измерений уровня искренности, произведенных с помощью двух разных инструментов, оказалось, что L-шкала чрезмерно завышает число анкет, подлежащих отбраковке. Так, при использовании в качестве критерия нормы, равной 1 баллу, мы должны отсеять около 86% всех собранных интервью. Измерение искренности на основе «эталонной» шкалы лжи дало совершенно иной результат: в среднем по всем 16 контрольным вопросам неискренние ответы, как оказалось, характерны лишь для 18–19% испытуемых. Это, безусловно, более точный показатель, поскольку он установлен критериальным путем.

Анализ кростабуляций свидетельствует, что даже при использовании самого жесткого критерия отбраковки анкет (1 балл – ложь) шкала улавливает лишь 80% неискренних респондентов, но 20% все равно пропускает. В свете этого вполне можно согласиться с выводом ряда специалистов о том, что «L-шкала весьма нечувствительна к некоторым типам тестовых искажений» [24, с. 17]. Вместе с тем она попутно «захватывает» и большое количество искренних испытуемых: до 88% в этой группе также должны быть незаслуженно отсеяны. Вне зависимости от жесткости квалификационного контроля на одну правомерную квалификацию в среднем приходится четыре-пять ошибочных (табл. 2).

Таблица 2. Соотношение правильных и ошибочных квалификаций в массиве анкет, забракованных L-шкалой, в зависимости от нормы отсеивания (%)

Квалификации анкет	1 балл и более – ложь	2 балла и более – ложь	3 балла и более – ложь	4 балла и более – ложь	5 баллов и более – ложь	В среднем по массиву
Правильные	18,2	17,8	18,1	19,3	18,9	18,5
Ошибочные	81,8	82,2	81,9	80,7	81,1	81,5

В результате в оставшемся после фильтрации массиве лишь около 76% анкет пригодны для дальнейшего анализа, а остальные 24% содержат социально желательные ответы. К тому же в этом случае в распоряжении исследователя останется лишь около

14% от первоначально собранного количества вопросников.

Что измеряют шкалы лжи. Если шкалы лжи не в состоянии зафиксировать склонность индивидов к социальной желательности, то что же тогда они измеряют? В поисках ответа на этот вопрос мы проанализировали социально-демографические характеристики искренних и неискренних респондентов, идентифицированных по сумме баллов, набранных ими по L-шкале, а также формы досуговой активности представителей этих категорий. Изучение полученных в исследовании данных позволяет констатировать следующие факты:

1. Испытуемые, набравшие высокие баллы по шкале лжи (4 и более), старше респондентов с низкими оценками (0–3 балла) в среднем на 6 лет.
2. Они имеют более высокий уровень образования, как правило принадлежат к интеллигенции (32,3% vs 25,7%).
3. Среди них значительно выше удельный вес людей вдовых, разведенных и разошедшихся (20,6% vs 10,5%).

Данные, представленные в табл. 3, свидетельствуют о наличии в подавляющем большинстве случаев статистически значимых взаимосвязей между количеством баллов, полученных респондентами, и их социально-демографическими характеристиками (за исключением гендерной принадлежности).

Таблица 3. Показатели взаимосвязи между социально-демографическими характеристиками респондентов и суммой баллов по L-шкале

Социально-демографические характеристики	Коэффициенты корреляции	Уровни значимости (p)
Пол	0,037 (Eta)	0,281
Возраст	0,237 (Sommer's d)*	0,000
Образование	0,246 (Eta)	0,038
Род занятий	0,272 (Eta)	0,005
Брачный статус	0,299 (Eta)	0,000

* Для направленной связи, где L-шкала – зависимая переменная.

Кроме того, судя по способам проведения досуга, индивиды, имеющие высокие дискредитирующие баллы, ведут здоровый образ жизни, соответствующий традиционным представлениям о нравственных формах поведения.

Между тем приведенные выше данные, хотя и укладываются в гипотезу Н. Брэдберна и С. Садмана, носят косвенный характер. Для более строгой статистической проверки сформулированных ранее предположений мы использовали следующую методику. На стадии обработки данных была построена интегральная аддитивная шкала, условно названная нами «шкалой нравственных стандартов». Она состояла из 8 вопросов–суждений на моральные темы: о допустимости/недопустимости лжи в межличностных отношениях, супружеской измене и т.д. Каждый из этих пунктов выступал в качестве самостоятельного эмпирического индикатора нравственного сознания и поведения людей в различных жизненных ситуациях.

Данная шкала была проверена на надежность посредством двух известных процедур, предпола-

гающих оценку ее внутренней гомогенности и консистентности. Коэффициент Альфа Кронбаха для 8-пунктной шкалы нравственности составил 0,5799, а коэффициент Спирмена-Брауна – 0,5436. И хотя эти показатели, к сожалению, не «дотягивают» до конвенционально приемлемого уровня, их значения выше, чем у шкал лжи.

Сведение данных, полученных по частным подшкалам, в интегральный «индекс нравственности» производилось на основе методики Гуттмана [25, с. 109–112]. В результате была создана 9-пунктная ординальная шкала размерностью от 0 до 8 баллов, которая затем была вновь преобразована в дихотомическую номинальную шкалу с позициями: 1) высокий стандарт нравственности; 2) низкий стандарт нравственности. При этом к первой позиции приписывались индивиды, набравшие по сумме всех вопросов от 5 до 8 баллов, а ко второй – от 0 до 3 баллов. Ответы респондентов, получивших 4 балла, в силу специфической метрики 9-пунктных шкал исключались из последующего анализа.

В результате построения индекса были выделены две группы респондентов: с высоким и низким нравственным стандартом. К первой из них мы относили тех испытуемых, которые ответили: 1) во всех жизненных ситуациях нужно говорить только правду; 2) они никогда не скрывают свои доходы от налогообложения; 3) всегда платят за проезд в общественном транспорте; 4) вернули бы найденный ими кошелек его владельцу (два ответа); 5) считают неприемлемым для себя обманывать государство; 6) никогда не изменяли жене (мужу) (два ответа). Во вторую группу вошли респонденты, давшие иные (противоположные) ответы.

Как свидетельствует проведенный нами корреляционный анализ, все изученные взаимосвязи отдельных пунктов шкалы нравственности с суммой баллов по L-тесту высоко значимы ($p = 0,000$). Коэффициенты Eta варьируют в диапазоне от 0,323 до 0,577, составляя в среднем 0,488 (табл. 4).

Таблица 4. Показатели корреляций каждого из 8 пунктов шкалы «нравственных стандартов» с индексом «социальной желательности» по L-шкале

№№ вопросов	χ^2	df	p	Eta
1	57,131	8	0,000	0,323
2	75,490	8	0,000	0,440
3	70,394	8	0,000	0,434
4	143,808	8	0,000	0,577
5	104,618	8	0,000	0,530
6	137,910	8	0,000	0,515
7	111,623	8	0,000	0,514
8	126,353	8	0,000	0,573

* Для направленной связи, где шкала «нравственных стандартов» – зависимая переменная.

С другой стороны, сами интегральные шкалы («нравственных стандартов» и социальной желательности) также тесно положительно коррелируют между собой на уровне индексов: $R = 0,238$ для $p = 0,000$.

И наконец, t-тест на статистическую значимость различий между средними баллами по шкале лжи, полученными респондентами с высоким и низким

нравственными стандартами, также оказался положительным (табл. 5).

Результаты теста, равно как и данные корреляционного анализа, убедительно свидетельствуют о том, что, с одной стороны, баллы, набираемые испытуемыми по L-тесту, значимо дифференцируют респондентов по уровню нравственности, а с другой, – индивиды, обладающие высоким нравственным стандартом, получают по шкале лжи значительно более высокие оценки, чем представители противоположной группы. Следовательно, чем большее количество баллов человек набирает по L-шкале, тем более высокую ступеньку он занимает в иерархии нравственных стандартов.

Таблица 5. Значимость различий между средними баллами по шкале «социальной желательности» (L) для испытуемых с высоким и низким «нравственным стандартом»

Категории испытуемых	Средние баллы по L-шкале	T-оценка	df	p
С высоким нравственным стандартом	2,9831	3,079	258	0,002
С низким нравственным стандартом	2,3451			

Сказанное дает основание утверждать, что повышенные оценки по шкале лжи репрезентируют не склонность испытуемых к социальной желательности, а специфику их реального мышления и поведения, повышенный стандарт нравственных оценок, а также особый способ восприятия социальных норм.

Эмпирическая проверка гипотезы о том, что шкала лжи из опросника Айзенка измеряет не склонность испытуемых к социальной желательности, а характеристики их нравственного сознания и поведения, не дала однозначных результатов. Свидетельством в пользу этого предположения выступает тот факт, что респонденты с высокими нравственными стандартами действительно набирают большее количество баллов по шкале лжи, чем испытуемые из противоположной группы (1,1065 и 0,800 соответственно). Однако различия между средними, фиксируемые посредством t-оценки, статистически не значимы ($p = 0,116$).

Заключение

Результаты исследований позволяют констатировать, что тестированные нами шкалы лжи из опросников MMPI и Айзенка являются недостаточно надежными инструментами для диагностики неискренних ответов респондентов. Ни один из проведенных нами тестов на надежность не дал положительных результатов, которые могли бы убедительно свидетельствовать о состоятельности этих шкал. Все полученные в исследовании статистические показатели оказались ниже конвенционально приемлемых значений.

Результаты валидизации также не позволяют говорить о «чистоте» производимых измерений. Как свидетельствуют полученные нами данные, шкалы лжи не дифференцируют значимо искренних и неискренних респондентов, а потому не обладают свойст-

вом дискриминантной валидности. С одной стороны, шкалы пропускают большое количество социально желательных ответов, квалифицируя их как вполне достоверные, а с другой, – ошибочно идентифицируют многих искренних информантов как имеющих склонность к искажению результатов. Отсутствие достаточной конструктивной (конвергентной и дискриминантной) валидности делает шкалы лжи не соответствующими их предназначению. Вместо измерения социальной желательности они фиксируют личностные характеристики респондентов, обусловленные спецификой их образа жизни, мышления и поведения, а также нравственные ценности и стандарты испытуемых. В результате общий уровень их эффективности не превышает 20%. Все это свидетельствует о серьезных нарушениях в идентификационном механизме данных шкал и не позволяет рекомендовать их в качестве надежного инструмента для выявления неискренних ответов респондентов в социологических и психологических исследованиях.

Вместе с тем данные выводы, скорее всего, не универсальны и могут быть экстраполированы лишь на так называемые «старые» шкалы лжи, большинство из которых были разработаны еще в 40–60-е гг. прошлого столетия. Они имеют сходные принципы конструирования, содержат много пересекающихся, почти одинаковых по содержанию и формулировкам вопросов. Но главное, пожалуй, состоит в том, что все они создавались исходя из представления о социальной желательности как едином, одномерном конструкте. Шкалы лжи нового поколения (например, BIDR–шкала Д. Паулюса [26]), напротив, основаны на идее многомерной природе измеряемой переменной, образующей двухфакториальное пространство. Согласно Д. Паулюсу, эти факторы, называемые им Альфа и Гамма, репрезентируют два разных измерения социальной желательности: тенденцию индивидов к самопревознесению, связанную с самообманом, и склонность к управлению впечатлением, ориентированную на обман других. Базирующаяся на этом подходе «сбалансированная шкала социально желательных ответов» (BIDR) требует самого серьезного анализа и дальнейших валидационных исследований.

Список литературы

1. **Crowne D.P., Marlowe D.** The Approval Motive: Studies in Evaluative Dependence. – N.Y.: Wiley, 1964.
2. **Andrews F.M.** Construct validity and error components of survey measures: A structural modeling approach // *Public Opinion Quarterly*. – 1984. – Vol. 48. – № 2. – P. 409–442.
3. **Dijkstra W.** Interviewing style and respondent behavior: An experimental study of the survey-interview // *Sociological Methods and Research*. – 1987. – Vol. 16. – № 2. – P. 309–334.
4. **Reinecke J., Schmidt P.** Explaining interviewer effects and respondent behavior: Theoretical models and empirical analysis // *Quality and Quantity*. – 1993. – Vol. 27. – № 3. – P. 219–247.
5. **Edwards A.L.** The Measurement of Personality Traits by Scales and Inventories. – N.Y. etc.: Rinehart and Winston, 1970.
6. **Гурко Т.А.** Особенности развития личности подростков в различных типах семей // *Социол. исслед.* – 1996. – № 3. – С. 81–90.
7. **Знаков В.В.** Психология понимания правды. – СПб.: Алетейя, 1999.
8. **Касьянова К.** О русском национальном характере. – М.: Ин-т национальной модели экономики, 1994.
9. **Общая психодиагностика** / Под ред. А.А. Бодалева, В.В. Столина. – М.: Изд-во Моск. ун-та, 1987.
10. **Давыдов А.А., Давыдова Е.В.** Измерение искренности респондента. – М.: ИС РАН, 1992.
11. **Bradburn N.M., Sudman S.** Improving Interview Method and Questionnaire Design. – San Francisco: Jossey-Bass, 1979.
12. **Kozma A., Stones M.J.** Social desirability in measures of subjective well-being: A systematic evaluation // *Journal of Gerontology*. – 1987. – Vol. 42. – № 1. – P. 56–59.
13. **Kozma A., Stones M.J.** Social desirability in measures of subjective well-being: Age comparisons // *Social Indicators Research*. – 1988. – Vol. 20. – № 1. – P. 1–14.
14. **Mastekaasa A., Kaasa S.** Measurement error and research design: A note on the utility of panel data in quality of life research // *Social Indicators Research*. – 1989. – Vol. 21. – № 3. – P. 315–335.
15. **Piedmont R.L., McCrae R.R., Riemann R., Angleitner A.** On the validity of validity scales: Evidence from self-reports and observer ratings in volunteer samples // *Journal of Personality and Social Psychology*. – 2000. – Vol. 78. – № 3. – P. 582–593.
16. **Собчик Л.Н.** Стандартизированный многофакторный метод исследования личности СМИЛ. – СПб.: Речь, 2002.
17. **Мягков А.Ю.** Шкалы лжи из опросника MMPI: Опыт экспериментальной валидизации // *Социол. исслед.* – 2002. – № 7. – С. 117–130.
18. **Обозов Н.Н.** Психодиагностика личности. – СПб.: Академия психологии, предпринимательства и менеджмента «Облик», 1998. – Ч. 1.
19. **Paulhus D.L., John O.P.** Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives // *Journal of Personality*. – 1998. – Vol. 66. – № 6. – P. 1025–1060.
20. **Stüber J., Dette D.E., Musch J.** Comparing continuous and dichotomous scoring of the balanced inventory of desirable responding // *Journal of Personality Assessment*. – 2002. – Vol. 78. – № 2. – P. 370–389.
21. **Campbell D.T., Fiske D.W.** Convergent and discriminate validation by the multitrait-multi-method matrix // *Psychological Bulletin*. – 1959. – Vol. 56. – № 1. – P. 81–105.
22. **Гречихин В.Г.** Лекции по методике и технике социологических исследований. – М.: Изд-во Моск. ун-та, 1988.
23. **Девятко И.Ф.** Диагностическая процедура в социологии: Очерк истории и теории. – М.: Наука, 1993.
24. **Пособие по применению MMPI** / Сост. А.А. Рукавишников, Н.Г. Рукавишникова, М.С. Соколова. – Ярославль: НПЦ «Психодиагностика», 1993.
25. **Ядов В.А.** Социологическое исследование: Методология, программа, методы. – Самара: Изд-во Самарского ун-та, 1995.

26. Paulhus D.L. The Balanced Inventory of Desirable Responding. – Toronto, Ontario: Multi-Health Systems, 1998.

Мягков Александр Юрьевич,
ГОУВПО «Ивановский государственный энергетический университет имени В.И. Ленина»,
доктор социологических наук, профессор, зав. кафедрой социологии,
телефон (4932) 26-97-82,
e-mail: socio@sociol.ispu.ru