

Методика и техника Социологических исследований

© 2002 г.

А.Ю. МЯГКОВ

ШКАЛЫ ЛЖИ ИЗ ОПРОСНИКА ММРІ: ОПЫТ ЭКСПЕРИМЕНТАЛЬНОЙ ВАЛИДИЗАЦИИ

МЯГКОВ Александр Юрьевич - кандидат философских наук, профессор, заведующий кафедрой социологии Ивановского государственного энергетического университета.

Постановка проблемы

Массовая социологическая практика свидетельствует, что исследования, базирующиеся на применении опросных методов, весьма уязвимы вследствие т.н. ситуативной лжи — результата использования респондентами защитных поведенческих стратегий (социальной желательности, конформизма, негативизма, регрессии поведения и др.). Для диагностики неискренних ответов в специальной литературе предлагаются разные методы и процедуры [1, р. 31-32; 2, с. 9-16]. Наиболее эффективными в этом отношении традиционно считаются шкалы лжи, заимствованные из различных личностных опросников.

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 проводившихся ими исследований с целью объяснения низкой предсказательной способности личностных тестов [3]. Многие авторы и по сей день используют ее для контроля искренности респондентов и отсеивания социально желательных ответов в опросах по сенситивной проблематике [4; 5].

В отечественной социологии и психологии, судя по литературе, в целях улучшения качества опросных данных исследователи чаще всего применяют шкалы лжи из тестов Айзенка [6], опросника Тейлор [7, с. 221-224], а также из Миннесотского многофазного личностного перечня (ММРІ). В работах многих психологов они оцениваются как высоко эффективное, хорошо разработанное психодиагностическое средство контроля искренности ответов опрашиваемых, позволяющее "автоматически отсеивать недостоверные протоколы" [8, с. 20; 9, с. 85; 10, с. 19-20]. Социологи также считают шкалы лжи "наиболее обоснованной методикой" измерения искренности респондентов [2, с. 17]. 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. В частности, в конце 1970-х годов Н. Брэдбери и С. Садман, тестируя шкалу Марлоу и Крауна, попытались ответить на вопрос, действительно ли она измеряет именно ту латентную переменную, для фиксации которой в свое время была создана¹. В результате валидационного эксперимента они пришли к выводу, что высокие баллы по результатам тестирования набирают не только респонденты, действительно склонные к социальной желательности, но и люди, имеющие высокие нравственные стандарты и неукоснительно следующие в повседневной жизни конвенциональным нормам. Следовательно, резюмируют авторы, шкала лжи в большей мере фиксирует не склонность испытуемых к искажению ответов на вопросы интервью, а реальные личностные характеристики индивидов, специфику их образа жизни, сознания и поведения [11, p. 85-106].

Методология и методы исследования

Объект исследования. Поскольку вопрос о диагностических возможностях шкал лжи по-прежнему остается открытым, мы решили выяснить, в какой мере данный инструмент соответствует своему функциональному назначению. В качестве объекта для изучения была выбрана L-шкала из опросника ММРІ, часто используемого в последнее время в отечественных психологических и педагогических исследованиях, а также в клинической диагностике и психотерапевтической практике.

Данный личностный опросник впервые был предложен в 1943 г. американскими психологами и клиницистами Дж. Маккинли и С. Хатауэем, а в 1989 г. модифицирован Дж. Баттлером и его коллегами (см.: [12, p. 447-449]). Сегодня ММРІ насчитывает в общей сложности 567 вопросов и, помимо 10 основных и множества вспомогательных "клинических" шкал, содержит еще 4 "контрольные" шкалы, предназначенные для выявления различных видов искажений в ответах испытуемых: "шкалу неискренности" (L), "надежности" (F), "коррекции" (K) и "неопределенных ответов" (?).

Одной из самых надежных с диагностической точки зрения считается L-шкала, ориентированная на фиксацию и измерение уровня социальной желательности. Она состоит из 15 вопросов-суждений, предполагающих однозначный ответ (по принципу "согласен - не согласен", "верно - неверно" и т.д.) в ситуации, отражающей житейские виды поведения, которые, как может показаться, осуждаемы строгой моралью (прил. 1). Большинство людей обычно легко признают те незначительные слабости и недостатки, которые фиксируются в вопросах, однако индивиды, намеренно стремящиеся подать себя в выгодном свете, отрицают их, отвечая не так, "как есть", а так, "как принято". В результате они получают высокие оценки по шкале, являющиеся основанием для отбраковки вопросников. Подробное описание этой и других "шкал валидности" из ММРІ можно найти в специальной литературе [13-16].

Выборка. Методы сбора данных. С целью проведения методологической экспертизы L-шкалы, предполагающей оценку ее надежности и валидности, в марте 2001 г. кафедрой социологии Ивановского государственного энергетического университета под руководством автора было предпринято специальное исследование экспериментального типа*. Всего в Иванове было опрошено 387 чел., предварительно отобранных на основе экспериментальной выборки и представляющих разные социально-демографические и профессиональные группы взрослого городского населения. В силу методологической направленности данного исследования выборка носила качественный характер. Строго пропорциональной репрезентации различных социальных категорий не требовалось. Поэтому с точки зрения социально-профессионального статуса в ней примерно в равных долях были представлены респонденты из пяти основных групп: рабочие, производственная и непроизводственная интеллигенция, работники торговли

*В сборе и обработке данных участвовали А.Н. Алексеичева, С.Л. Журавлева, А.О. Морозова, Н.А. Потемнина и О.В. Симакова

и сферы бытового обслуживания, студенты и безработные. По полу выборка точно воспроизводит статистические параметры генеральной совокупности. По возрастным характеристикам она несколько смещена в сторону большего представительства респондентов молодого и среднего возраста.

Для сбора эмпирических данных использовался метод формализованного персонального интервью. Опросы проходили по месту жительства или работы респондентов. Об экспериментальном характере данного исследования испытуемым не сообщалось. Статистико-математическая обработка информации осуществлялась в SPSS.

Структура вопросника. Основные и дополнительные переменные. Социологическая информация собиралась с помощью специально разработанного вопросника, насчитывающего в общей сложности 69 вопросов и состоящего из четырех условных блоков. Первый блок был представлен 15-ю суждениями L-шкалы, разбросанными в пространстве анкеты в случайном порядке. Полученные ответы затем квалифицировались на предмет достоверности/недостоверности в соответствии с ключами, предусмотренными для данного теста, и сводились в интегральный показатель, названный нами "индексом социальной желательности". Он измерялся в интервальной шкале, в качестве значений которой выступали итоговые баллы, набранные респондентами по всем пунктам теста. С этой целью на стадии компьютерной обработки данных была построена дополнительная искусственная переменная.

Второй условный блок представлял собой серию из 16 основных и примерно такого же количества контрольных вопросов (проверочных, "ловушек", тестов на знания и др.), ответы на которые после сопоставления и соответствующей квалификации использовались в качестве относительно надежных эталонов достоверности/недостоверности. С учетом этого все респонденты были разделены на две группы: искренних и неискренних. При этом к неискренним мы относили тех испытуемых, которые либо демонстрировали неконсистентность в ответах на два-три сходных по смыслу вопроса, либо выбирали явно несуществующие подсказки в вопросах-"ловушках". Большинство вопросов контрольного блока были специально сформулированы таким образом, чтобы спровоцировать опрашиваемых на неискренние ответы (прил. 2). Режим персонального интервью исключал для респондентов возможность приведения в соответствие ответов на основные и проверочные вопросы. Полученные "эталонные", на базе которых были созданы 16 дополнительных переменных, использовались нами для проверки валидности L-шкалы.

На основе вопросов контрольного блока был сконструирован еще один сводный показатель, получивший условное название "эталонного индекса лжи". Он также измерялся в интервальной шкале и фиксировал распределение респондентов по количеству контрольных вопросов, на которые они дали неискренние ответы.

В третьем блоке вопросника присутствовали традиционные социально-демографические вопросы относительно пола, возраста, рода занятий, образования и брачного статуса респондентов. В анкету входили также вопросы функционально-психологического назначения (контактные, буферные, "глушители" и др.).

Методы измерения надежности. Поскольку L-шкала представляет собой классический образец тестовой методики, то для статистической оценки ее надежности мы использовали принятые в тестологии стандартные процедуры.

Во-первых, в ходе нашего исследования проводился анализ интеркорреляций между отдельными пунктами L-шкалы. В связи с тем, что коррелируемые переменные измерены на номинальном уровне и имеют дихотомическую природу ("верно-неверно"), то наиболее адекватной мерой корреляции в этом случае мы посчитали коэффициент контингенции, используемый обычно для измерения абсолютной связи в четырехклеточных таблицах [17, с. 223].

Во-вторых, изучались корреляции каждого из 15 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 данном случае является номинальной, а вторая — интервальной, для харак-

теристики силы связи между ними использовались Eta-коэффициенты как наиболее точно отвечающие природе и специфике измерения.

В-третьих, интересующий нас тест проверялся на внутреннюю консистентность посредством коэффициента Альфа Кронбаха. Данный коэффициент представляет собой оценку надежности, базирующуюся на гомогенности шкалы, и вычисляется как сумма корреляций между ответами испытуемых на вопросы внутри одной и той же тестовой формы. Его расчетная формула принимает во внимание количество вопросов, общую дисперсию оценок индивидов и сумму дисперсий баллов, полученных респондентами по каждому пункту шкалы [18, р. 502]. И, наконец, в-четвертых, производилась оценка т.н. надежности-согласованности теста [19, с. 114]. С этой целью осуществлялся расчет коэффициента Спирмена-Брауна в рамках статистической модели "split half". Ретестовая надежность в данном исследовании не проверялась.

Методы валидации шкалы. Вполне очевидно, что в случае с L-шкалой не может быть и речи о т.н. лицевой валидности ее вопросов-суждений в силу объективного характера теста и латентности цели осуществляемого измерения. Вряд ли возможно и установление содержательной валидности экспертным путем (например, с использованием метода "параллельных панелей"), поскольку ни один, даже самый опытный эксперт не рискнул бы дать однозначное заключение о степени соответствия или несоответствия того или иного суждения содержанию измеряемой латентной переменной. Тем более, что пункты L-шкалы отбирались ее авторами не на основе логического подхода, т.е. соотношения содержания суждений со значением заданного свойства, а посредством чисто эмпирических процедур [13, с. 6-7].

В данном случае, скорее всего, есть смысл попытаться установить конструктивную валидность данной шкалы и тем самым проверить, в какой мере все ее пункты вместе соответствуют содержанию конструкта в целом. Мы сочли возможным сделать это путем коррелирования "индекса социальной желательности", полученного по L-шкале, с ответами респондентов на целую серию "эталонных" (контрольных) вопросов, также фиксирующих неискренность, но иным, несомненно, более надежным образом. При этом мы исходили из того, что если в результате будут обнаружены высокие значимые корреляции между указанными переменными (по крайней мере, для тех случаев, которые обеспечивают максимально достоверные квалификации ответов), то можно говорить о валидности L-шкалы на уровне измеряемого конструкта.

Некоторые авторы называют такой вид валидности "содержательной (корреляционной) валидностью для конструкта", в том смысле, что ее установление возможно посредством анализа корреляций между ответами испытуемых на близкие по смыслу вопросы [20, р. 81]. П. Каттенс, вслед за Д. Кэмпбелом и Д. Фиске [21], определяет ее как "конструктивную (конвергентную)", не без оснований полагая при этом, что если ответы респондентов по тесту в целом и на вопросы контрольного блока близки ("конвергентны"), т.е. характеризуются одной и той же тенденцией, то мы имеем дело именно с конструктивной валидностью тестируемой шкалы. "Для нефактуальных вопросов, - пишет автор, - валидация единственно возможна лишь на основе оценки консистентности серии данных, полученных с помощью различных вопросов, каждый из которых имеет целью измерить тот же самый ненаблюдаемый теоретический конструкт, т.е. посредством оценки конструктивной (конвергентной) валидности" [22, р. 31].

Вместе с тем ничуть не менее важной задачей в данном исследовании мы считали проверку L-шкалы на т.н. дискриминантную валидность. Если шкала отчетливо дифференцирует группы искренних и неискренних респондентов по числу набранных ими баллов, при этом устойчиво демонстрируя данную способность в целом ряде внетестовых ситуаций, то она без сомнений может считаться высоко валидным инструментом для измерения заданного свойства.

При установлении валидности этого типа мы использовали процедуру t-теста с целью анализа средних оценок, полученных искренними и неискренними респондентами по шкале "социальной желательности", на предмет статистической значимости различий.

Результаты и их обсуждение

L-шкала: оценка надежности. Анализ интеркорреляций между пунктами L-шкалы на основе критерия χ^2 свидетельствует, что лишь 37 из 105 проверенных взаимосвязей (35,2%) являются значимыми на конвенционально приемлемом уровне ($p < 0,05$). Рассмотрение каждого суждения в отдельности показало, что ни одно из них не коррелирует значимо со всеми остальными 14-ю. Даже пункты №№ 45 и 90, имеющие максимальное число статистически значимых корреляций (по 8), взаимосвязаны лишь с половиной тестовых субшкал (табл. 1).

Средний показатель коэффициента контингенции для значимых связей по всему тесту составляет лишь 0,157 при общей вариации значений от 0,003 до 0,292. В целом по матрице связей он едва превышает 0,10. Подобные значения весьма типичны для шкал с невысоким уровнем надежности. В исследовании Н. Брэдберна и С. Садмана, посвященном реинтерпретации шкалы Марлоу-Крауна, средний коэффициент межвопросной корреляции оказался равным 0,09 [11, р. 88], а в исследовании Д. Смита - 0,18 [23, р. 91], что было расценено авторами как веское основание для отрицательного заключения относительно пригодности тестируемого инструмента.

Наиболее сильными переменными L-шкалы могут, по-видимому, считаться суждения №№ 225, 90, 15 и 45. Однако усредненные значения коэффициентов взаимной корреляции по этим вопросам невысоки (0,12-0,13) и не позволяют считать данный вывод окончательным. Наиболее слабыми являются суждения №№ 195, 150, 60 и 165. Однако и это заключение нуждается в дальнейшей проверке и уточнении.

По сути дела все эти данные означают, что в тесте нет высоко скоррелированных пунктов. Следовательно, практически все суждения в нем индивидуальны, не взаимозаменяемы, а процедура измерения столь сложного конструкта, как неискренность, не может быть сведена к постановке одного или даже нескольких единичных вопросов. Даже если взять четыре указанных выше суждения с наивысшими средними значениями корреляций, то можно заметить, что не все из них значимо коррелируют между собой: № 225, например, выбивается из общей тенденции, хотя и имеет в среднем самый высокий показатель силы взаимосвязей.

Чтобы понять, в какой мере все пункты L-шкалы "работают" на одну и ту же латентную переменную, подлежащую измерению, нами был проведен анализ корреляций каждого ее пункта с интегральным аддитивным показателем, т.е. с суммарным баллом, набранным всеми респондентами по тесту в целом. Результаты, полученные в рамках данной процедуры, чуть более оптимистичны, чем в предыдущем случае, хотя и они свидетельствуют о недостаточной надежности и, в частности, о неполной внутренней консистентности анализируемой шкалы. Все изученные нами корреляции значимы на высоком уровне, однако коэффициенты, характеризующие силу связей, вновь невелики и могут быть интерпретированы скорее как умеренные³. Средний показатель Eta для всего теста равен 0,330, при этом максимальное значение данного коэффициента не достигает 0,5, а минимальное - 0,16 (табл. 2).

Вместе с тем полученные данные позволяют заключить, что как минимум для двух пунктов шкалы (№№ 75 и 150) интенсивность связи с суммой баллов весьма незначительна (0,155 и 0,180, соответственно). Их вклад в значение итогового индекса минимален. Однако вопрос о том, можно ли рассматривать эти суждения в качестве кандидатов на удаление из теста с целью повышения его гомогенности, пока неясен. Для этого необходимо более детально проанализировать влияние данных пунктов (как, впрочем, и всех остальных) на общий показатель надежности L-шкалы. Тем более что при оценке надежности измерительных инструментов важнее ориентироваться не на высокие значения интеркорреляций или корреляций отдельных суждений с итоговой суммой баллов, а на показатель "их общей внутренней согласованности" традиционно оцениваемый посредством коэффициента Альфа Кронбаха [24, с. 56].

Данный коэффициент, рассчитанный для L-шкалы с помощью программы "Reliability" из SPSS, составил 0,4441, что явно свидетельствует о низком уровне

Матрица интеркорреляций между пунктами L-шкалы

№	15	30	45	60	75	90	105	120	135	150	165	195	225	255	285	Среднее значение коэффициента контингенции	Число значимых связей
15		0,270*	0,224*	0,058	0,035	0,158*	0,096	0,143*	0,148*	0,042	0,041	0,102	0,149	0,152*	0,217*	0,131	7
30	0,270*		0,203*	0,027*	0,196*	0,200*	0,088	0,034	0,220*	0,011	0,006	0,047	0,071	0,056*	0,118	0,110	7
45	0,224*	0,203*		0,128*	0,063	0,128*	0,102*	0,166*	0,165*	0,070	0,131*	0,017	0,056	0,095	0,168	0,122	8
60	0,058	0,027*	0,128*		0,006	0,087	0,012	0,043	0,061	0,003*	0,041	0,006	0,230	0,035*	0,028	0,055	4
75	0,035	0,196*	0,063	0,006		0,211*	0,292*	0,036	0,074	0,061	0,114*	0,048	0,094	0,042	0,050	0,094	4
90	0,158*	0,200*	0,128*	0,087	0,211*		0,176*	0,090	0,145*	0,019	0,078	0,088	0,250*	0,128*	0,071	0,131	8
105	0,098	0,088	0,102*	0,012	0,292*	0,176*		0,007	0,006	0,034	0,002	0,033	0,185*	0,019	0,001	0,075	4
120	0,143*	0,034	0,166*	0,043	0,036	0,090	0,007		0,122*	0,077	0,093	0,013	0,052	0,220*	0,170*	0,090	5
135	0,148*	0,220*	0,165*	0,061	0,074	0,145*	0,006	0,122*		0,048	0,144*	0,014	0,068	0,142*	0,074	0,102	7
150	0,042	0,011	0,070	0,003*	0,061	0,019	0,034	0,077	0,048		0,040	0,054	0,116	0,068	0,119*	0,054	2
165	0,041	0,006	0,131*	0,041	0,114*	0,078	0,002	0,093	0,144*	0,040		0,006	0,058	0,044	0,086	0,063	3
195	0,102	0,047	0,017	0,006	0,048	0,088	0,033	0,013	0,014	0,054	0,006		0,173*	0,082	0,060	0,053	1
225	0,149	0,071	0,056	0,230	0,094	0,250*	0,185*	0,052	0,068	0,116	0,058	0,173*		0,219	0,124	0,132	3
255	0,152*	0,056*	0,095	0,035*	0,042	0,128*	0,019	0,220*	0,142*	0,068	0,044	0,082	0,219		0,116*	0,101	7
285	0,217*	0,118	0,168	0,028	0,005	0,071	0,001	0,170*	0,074	0,119*	0,086	0,060	0,124	0,116*	0,097		4

*Коэффициенты контингенции, значимые на уровне $p < 0,05$.

Таблица 2

Корреляции всех 15 пунктов L-шкалы с индексом "социальной желательности"

№№ вопросов	χ^2	df	P	Eta*
15	92,229	5	0,000	0,482
30	20,942	5	0,001	0,226
45	75,923	5	0,000	0,380
60	28,663	5	0,000	0,214
75	22,718	5	0,001	0,155
90	85,150	5	0,000	0,401
105	39,532	5	0,000	0,282
120	70,627	5	0,000	0,417
135	63,595	5	0,000	0,401
150	16,194	5	0,006	0,180
165	45,231	5	0,000	0,300
195	21,950	5	0,001	0,224
225	83,103	5	0,000	0,433
255	102,176	5	0,000	0,487
285	64,741	5	0,000	0,374

*Для направленной связи, где "индекс социальной желательности" - зависимая переменная

надежности интересующего нас вопросника. Судя по литературе, приемлемым в данном случае можно было бы считать такой показатель корреляции, который превышает 0,8. По сообщению Д. Крауна и Д. Марлоу, коэффициент надежности для их 33-пунктной шкалы социальной желательности составляет 0,88 [11, p. 88]. Ф. Ример считает допустимым значение Альфа, равное 0,69 для 7-пунктной шкалы [18, p. 502]. Тест, сконструированный Е. Головахой, Н. Паниной и А. Горбачиком из 44 вопросов,

в эксперименте продемонстрировал общую надежность на уровне 0,867, что было признано авторами весьма высоким показателем [24, с. 56]. Между тем, Н. Брэдбери и С. Садман, обнаружив, что тестированный ими "усеченный" вариант МС-шкалы, сформированной из 10 пунктов, характеризуется коэффициентом, равным 0,497, сделали однозначный вывод о том, что шкала является "относительно слабым" инструментом для измерения неискренности [11, р. 88]. Процедура последовательного исключения пунктов из L-шкалы при одновременном контроле Альфа Кронбаха, результаты которой представлены в табл. 3, показала, что три вопроса в данной шкале (№№ 90, 225 и 165) значимо отрицательно влияют на показатель ее общей надежности, снижая гомогенность теста. Эти суждения, по-видимому, иррелевантны содержанию и структуре вопросника и не отражают истинного значения измеряемой переменной. Одновременное (совместное) удаление всех этих трех вопросов заметно повышает значение Альфа (до 0,5165), хотя и не настолько, чтобы можно было считать его приемлемым. Вместе с тем измерение того же показателя только для этих пунктов дает явно отрицательный результат: Альфа Кронбаха для искусственно созданной 3-пунктной шкалы составляет - 0,0462. Для сравнения отметим, что значение Альфа для шкалы, сформированной из трех высоко надежных вопросов (№№ 30, 75, 150) составило 0,4163. Следовательно, при условии высокой гомогенности L-шкалы был бы смысл в применении сокращенного (12-пунктного) ее варианта для повышения надежности измерения неискренности респондентов. С другой стороны, при решении задачи, связанной с отбором суждений для кон-струирования "комбинированной шкалы искренности" [2, с. 17], целесообразно было бы использовать именно те из указанных пунктов, которые обладают повышенной надежностью (№№ 15, 45, 285, 255, 120, 135 и 30).

Для измерения степени согласованности отдельных вопросов L-шкалы с латентной характеристикой в рамках модели split-scale нами был также рассчитан коэффициент

Таблица 3

Влияние отдельных пунктов L-шкалы на изменение значения коэффициента Альфа Кронбаха

Удаляемые пункты шкалы	Коэффициент Альфа Кронбаха	Удаляемые пункты шкалы	Коэффициент Альфа Кронбаха
15	0,3750	135	0,4131
30	0,4176	150	0,4428
45	0,3951	165	0,4526
60	0,4400	195	0,4350
75	0,4419	225	0,4913
90	0,4481	255	0,4069
105	0,4293	285	0,4022
120	0,4133	90+225+165	0,5165

Спирмена-Брауна, фиксирующий меру корреляции между двумя частями теста, каждая из которых включает половину пунктов единого вопросника. По мнению специалистов, чем выше этот коэффициент, тем согласованнее оценки истинного значения переменной, получаемые с помощью данного набора индикаторов [19, с. 114]. Судя по полученным в исследовании данным, вопросы L-шкалы, к сожалению, не обладают указанным свойством: коэффициент Спирмена-Брауна, равный 0,3324, свидетельствует о низкой надежности анализируемого теста по критерию согласованности его субшкал между собой и с латентной переменной.

L-шкала: проблема валидности. Валидность шкалы тестировалась нами в двух отношениях. С одной стороны, нас интересовал вопрос о том, измеряет ли она неискренность как изначально заданное свойство (конструктивная валидность). А с другой, мы хотели знать, обладает ли данный инструмент способностью значимо дифференцировать искренних и неискренних респондентов и тем самым продуктивно выполнять свою главную функцию (дискриминантная валидность).

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

В результате проведенного анализа оказалось, что лишь 5 из 16 тестируемых зависимостей были статистически значимыми на уровне не ниже 0,05. Из четырех указанных выше "эталонных" переменных ("визит В. Путина в США", "дата основания Москвы", "лоббинг" и "куадрин"⁴) только две последние значимо коррелировали с суммарным "индексом социальной желательности". Все остальные проведенные нами тесты дали основание для отрицательных заключений. Вместе с тем даже в тех случаях, когда наблюдались значимые связи, их интенсивность была незначительной: в среднем для 5 переменных, коррелирующих с суммарным баллом, значение Ета-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 (табл. 4).

Для проверки L-шкалы на дискриминантную валидность мы воспользовались процедурой "двухвыборочного" t-теста. С этой целью в SPSS первоначально были рассчитаны средние баллы, набранные по каждой из 16 контрольных переменных искренними и неискренними респондентами, получившими такую квалификацию на основе ответов на вопросы-"ловушки", а затем эти показатели сравнивались между собой посредством t-статистик. Значимость различий между средними определялась с учетом теста Ливиня для условия о равенстве дисперсий.

Таблица 4

Корреляции между контрольными переменными и индексом "социальной желательности" по L-шкале

№	Контрольные переменные*	Eta**	P
1.	Визит В.В. Путина в США	0,116	0,663
2.	Симпатии к политическим партиям	0,132	0,082
3.	Участие в выборах	0,019	0,998
4.	Участие в выборах мэра г. Иваново	0,116	0,199
5.	Дата основания Москвы	0,086	0,114
6.	Интерес к культурной жизни города	0,127	0,132
7.	Интерес к политической жизни города	0,064	0,275
8.	Интерес к экономическим проблемам города	0,080	0,437
9.	Интерес к социальным проблемам города	0,078	0,677
10.	Допустимость лжи	0,159	0,033***
11.	Обман государства	0,064	0,406
12.	Потерянный бумажник	0,052	0,671
13.	Англоязычные термины ("лоббинг")	0,161	0,032***
14.	Наркотики ("куадрин")	0,210	0,000***
15.	Гомосексуальные отношения	0,248	0,000***
16.	Супружеская неверность	0,253	0,000***

*Здесь приводятся лишь условные названия-дескрипторы переменных. Полные версии вопросов формулировок представлены в приложении 2.

**Для ненаправленной связи.

***Корреляции, значимые на уровне $p < 0,05$.

В данном случае мы предполагали, что если по всем или хотя бы по большинству сравниваемых пар средних будут получены значимые положительные t-оценки, то L-шкала может считаться инструментом, достоверно дифференцирующим группы искренних и неискренних респондентов, а потому обладает достаточно высокой дискриминантной валидностью.

В результате анализа полученных данных выяснилось, что лишь в половине всех тестируемых сравнений имеются статистически значимые различия между средними баллами, набранными испытуемыми по шкале лжи из ММРІ (табл. 5). 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 внимательное рассмотрение данных показывает, что даже эти крайне неутешительные показатели сильно завышены.

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, существование которых противоречит гипотезе о том, что высокие баллы по L-шкале характерны для респондентов, склонных к самопрезентации. Т.е., по сути дела шкала лжи из ММРІ смогла достоверно дифференцировать группы искренних и неискренних испытуемых только в 3 из 16 протестированных нами случаев. К тому же в достоверности одного из них (вопрос № 2) 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 информантов" во внетестовых ситуациях. При этом последняя из указанных тенденций встречается, по-видимому, значительно чаще.

Таблица 5

Значимость различий между средними баллами по шкале "социальной желательности" неискренних респондентов по контрольным вопросам

№	Контрольные переменные	Средние баллы по L-шкале		t-оценки	P
		Искренние респонденты*	Неискренние респонденты*		
1.	Визит В.В. Путина в США	2,547	2,976	2,063	0,040**
2.	Симпатии к политическим партиям	2,628	4,000	16,032	0,000**
3.	Участие в выборах	2,648	2,600	-0,229	0,819
4.	Участие в выборах мэра г. Иваново	2,594	3,222	1,884	0,060
5.	Дата основания Москвы	2,654	2,545	-0,404	0,686
6.	Интерес к культурной жизни города	2,751	2,354	-2,109	0,036**
7.	Интерес к политической жизни города	2,638	2,638	-0,001	0,999
8.	Интерес к экономическим проблемам города	2,590	2,778	0,961	0,337
9.	Интерес к социальным проблемам города	2,703	2,443	-1,325	0,186
10.	Допустимость лжи	2,485	3,009	2,820	0,005**
11.	Обман государства	2,629	2,727	0,266	0,790
12.	Потерянный бумажник	2,625	2,681	0,283	0,777
13.	Англоязычные термины ("лоббинг")	2,810	2,295	-2,880	0,004**
14.	Наркотики ("квадрин")	2,754	1,906	-3,476	0,001*
15.	Гомосексуальные отношения	2,785	1,971	-3,388	0,001**
16.	Супружеская неверность	2,749	1,830	-3,077	0,003**

Эта гипотеза находит, похоже, и еще одно статистическое подтверждение. Проведенный нами корреляционный анализ свидетельствует о существовании слабой обратной взаимосвязи между двумя интегральными индексами: "эталонной лжи" и "социальной желательности", измеренной по L-шкале. Коэффициент корреляции Пирсона (R) составляет - 0,091 при $p = 0,074$. Следовательно, чем чаще респондент "попался" на контрольные вопросы, 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.

Заключение

Наш опыт экспериментального исследования позволяет сделать неутешительные выводы. Мы вынуждены констатировать, что тестируемая нами шкала лжи из опросника MMPI является недостаточно надежным и слабо валидным инструментом для диагностики неискренних ответов респондентов. Ни один из четырех проведенных нами тестов на надежность не дал положительных результатов, которые могли бы убедительно свидетельствовать о состоятельности L-шкалы. Все полученные в исследовании статистические показатели оказались ниже конвенционально приемлемых значений.

Использованные процедуры валидации также не позволяют говорить о "чистоте" производимых измерений. Вопреки общепринятому мнению, L-шкала представляет собой довольно слабый диагностический инструмент. Как свидетельствуют полученные нами данные, она не дифференцирует значимо искренних и неискренних респондентов, а потому не обладает свойством дискриминантной валидности. С одной стороны, шкала пропускает большое количество социально желательных ответов, квалифицируя их как вполне достоверные, а с другой, ошибочно идентифицирует многих искренних информантов как имеющих склонность к искажению результатов. Отсутствие достаточной конструктивной и дискриминантной валидности делает L-шкалу несоответствующей ее изначальному предназначению. В результате общий уровень ее эффективности не превышает 19%. Все это свидетельствует о серьезных нарушениях в идентификационном механизме данной шкалы и не позволяет рекомендовать ее в качестве надежного инструмента для выявления неискренних ответов респондентов в социологических и психологических исследованиях.

Кроме того, у обсуждаемой проблемы есть еще один очень важный, хотя и менее очевидный аспект. Поскольку в MMPI "контрольные" шкалы, как отмечают специалисты, тесно коррелируют с "клиническими" [13, с. 8, 13-14; 17, с. 109], то любые неверные квалификации, сделанные на стадии фильтрации негодных анкет, неизбежно обернутся ошибками в общей и специальной диагностике личности испытуемых. Поэтому выводы о низкой надежности и валидности L-теста следует учитывать не только исследователям, стоящим перед выбором того или иного шкального метода с целью его использования для улучшения качества собираемых данных, но и специалистам, работающим с MMPI в клинко-диагностических и терапевтических целях.

Вместе с тем, сказанное нами относительно L-шкалы неправомерно механически экстраполировать на весь класс подобных методик. Шкалы лжи, созданные в рамках многочисленных личностных опросников, заметно различаются по своим свойствам и характеристикам, способам конструирования и процедурам валидации, что не дает оснований для генерализации выводов. Учитывая общность исходных принципов их организации и функционирования, нельзя исключить, что изъяны, присущие L-шкале, весьма характерны и для других аналогичных инструментов. Исследователям еще предстоит найти наиболее эффективный способ диагностики неискренних ответов респондентов в опросных исследованиях. Однако для этого необходима целая серия специальных испытаний.

ПРИМЕЧАНИЯ

- ¹ В нашей методологической литературе результаты этого исследования остались практически незамеченными. Исключение составляет, пожалуй, лишь работа И.Ф. Девятко [19, с. 101].
- ² В современной литературе существуют разные подходы к пониманию валидности. Нередко этот термин используется в предельно широком смысле, как интегральная характеристика качества эмпирических данных [25, р. 169]. В нашей работе под валидностью мы понимаем соответствие измерения его изначальной цели [26, р. 208; 28, с. 25] и "доказательность смысла замеренных переменных" [28, с. 19].
- ³ Между тем некоторые авторы считают значения корреляций такого уровня крайне низкими. Так, например, И.Ф. Девятко, анализируя методы оценки надежности измерений, пишет, что если для группы опрошенных коэффициент корреляции между отдельным вопросом и суммарным индексом оказался равен 0,3, "то можно предположить, что названный вопрос не отражает истинного значения переменной... и может быть исключен из опросника" [19, с. 115].
- ⁴ Этот вымышленный термин мы обнаружили в статье Т. Маккая и Я. Макаллистера [25, р. 185].

ПРИЛОЖЕНИЕ 1

Вопросы L-шкалы из теста ММРІ* (с сохранением аутентичной нумерации)

- № 15. Временами мне в голову приходят такие нехорошие мысли, что о них лучше не рассказывать.
- № 30. Иногда мне хочется выругаться
- № 45. Я не всегда говорю правду.
- № 60 Я не каждый день читаю передовицы в газетах.
- № 75 Иногда я бываю сердитым.
- № 90 Иногда я откладываю на завтра то, что можно сделать сегодня.
- № 105 Иногда, когда я плохо себя чувствую, я бываю раздражительным.
- № 120 Мои манеры за столом у себя дома не так хороши, как в гостях.
- № 135 Я перехожу улицу в неполюженном месте, когда уверен(а), что меня не заметит милиционер.
- № 150 В игре мне приятнее выигрывать, чем проигрывать.
- № 165 Мне нравится иметь значимых людей среди моих знакомых, т.к. это повышает мой престиж.
- № 195 Не все, кого я знаю, мне нравятся.
- № 225 Иногда я могу немного посплетничать.
- № 255 Иногда я даю хорошую оценку людям, о которых знаю очень мало.
- № 285 Бывает, что я смеюсь по поводу неприличной шутки.

ПРИЛОЖЕНИЕ

Формулировки основных и вспомогательных вопросов контрольного блока интервью.

№	Основные вопросы	№	Вспомогательные вопросы / Тип контрольной техники
1.	"Сейчас все говорят о предстоящем визите В.В. Путина в США. Слышали ли Вы об этом?"	1.	Вопрос-ловушка
2.	"Считаете ли Вы себя сторонником какой-либо политической партии, организации?". [Да;нет]	2.	"Если да, то какой именно? Укажите, пожалуйста, ее название" (Проверочный вопрос-фильтр). "Кто является лидером данной организации? Укажите, пожалуйста, фамилию". (Проверочный вопрос).
3.	"Говорят, что участие в выборах является долгом каждого гражданина. Согласны ли Вы с этим?". [Да; нет].	3.	"Участвовали ли Вы в выборах мэра г. Иваново в декабре 2000 г.?" [Да; нет]. (Контрольный вопрос)
4.	"Участвовали ли Вы в выборах мэра г. Иваново в декабре 2000 г.?" [Да; нет]	4.	"Если да, то за кого Вы проголосовали? Укажите, пожалуйста, фамилию". (Проверочный вопрос-фильтр)
5.	"Считается, что историю Отечества должен знать каждый. А Вы знаете, в каком году была основана столица нашей Родины?". [Да; нет]	5.	"Укажите, пожалуйста, дату основания Москвы": 1223 г. 1240 г. 1112 г. 1147 г. Иной год. (Тест на знание)
6.	"Следите ли Вы за событиями культурной жизни нашего города?". [Да; нет]	6.	"Какие наиболее яркие события, на Ваш взгляд, произошли в культурной жизни нашего города за последние два месяца? (Укажите, пожалуйста)". (Проверочный вопрос)
7.	"Следите ли Вы за событиями политической жизни нашего города?". [Да; нет]	7.	"Как Вы считаете, какие проблемы в политической жизни нашего города являются наиболее острыми, требующими первоочередного решения? (Укажите, пожалуйста)". (Проверочный вопрос)
8.	"Следите ли Вы за событиями в экономической жизни нашего города?". [Да; нет]	8.	"Как Вы считаете, какие проблемы в экономике нашего города являются наиболее острыми, требующими первоочередного решения? Укажите, пожалуйста". (Проверочный вопрос)
9.	"Следите ли Вы за событиями, происходящими в социальной сфере нашего города?". [Да; нет]	9.	"Как Вы считаете, какие проблемы в социальной сфере нашего города являются наиболее острыми, требующими первоочередного решения? Укажите, пожалуйста". (Проверочный вопрос)
10.	"В настоящее время в нашем обществе стали преобладать отношения недоверия и взаимного обмана. Как Вы думаете, допустимо ли в некоторых ситуациях солгать, или нужно при любых обстоятельствах говорить правду?". [В некоторых ситуациях допустимо солгать; при любых обстоятельствах нужно говорить правду; затрудняюсь ответить]	10.	"Приходилось ли Вам когда-либо лгать и обманывать?" [Да; нет]. (Контрольный вопрос).
11.	"Я никогда не скрываю свои доходы от налогообложения". [Верно; неверно]	11.	"Считаете ли Вы себя возможным обманывать государство?" [Да; нет].

12	“Я могу оставить себе найденный кошелек с деньгами”. [Верно; неверно]	12	“Представьте себе, что Вы случайно обнаружили бумажник, в котором оказалась крупная сумма денег. Как бы Вы поступили в этом случае?” [Оставил бы находку себе; постарался бы вернуть владельцу; поступил бы иначе] (Контрольный вопрос).
13	“В последнее время в СМИ и в повседневной жизни стало модным употребление терминов английского происхождения. Приходилось ли Вам слышать следующие термины?” [Да;нет]: шопинг. Шейпинг. Лоббинг. Маркетинг. Мониторинг. Дансинг.	13	Вопрос –ловушка (“лоббинг”)
14	“Слышали ли Вы когда –либо о следующих видах наркотиков?” [Да;нет]: гашиш. Опиум. Морфин. Кодеин в таблетках. Куадрин. Экстази. Героин. Кокаин. ЛСД	14	Вопрос –ловушка (“куадрин”)
15	“Если ли бы Ваш близкий друг (подруга) Вашего пола предложил (а) бы Вам вступить в сексуальные отношения, как бы Вы поступили?” [Согласился (ась) бы; отказался (ась) бы; затрудняюсь ответить]	15	“Я думал (а) о том, чтобы вступить в интимные отношения с представителями своего пола”. [Да; нет]. (Контрольный вопрос).
16	“Приходилось ли Вам, уже будучи в браке, вступать в сексуальные отношения с кем-нибудь, помимо супруга?” [Да; нет]	16	“Были случаи, что я изменял (а) своей жене (мужу)”. [Верно; неверно]. (Контрольный вопрос).

СПИСОК ЛИТЕРАТУРЫ

1. *Obligation F.R.* Managing perceived deception among respondents: a traveler's tale // *Journal of Contemporary Ethnography*. 1994. Vol. 23. № 1. P. 29-50.
2. *Давыдов А.А., Давыдова Е.В.* Измерение искренности респондента. М.: Изд-во ИС РАН, 1992.
3. *Crowne D.P., Marlowe D.* The Approval Motive: Studies in Evaluative Dependence. N.Y.: Wiley, 1964.
4. *Nederhof A.J.* A survey on suicide: using a mail survey to study a highly threatening topic // *Quality and Quantity*. 1985. Vol. 19. № 3. P. 293-302.
5. *Tracy P.E., Fox J.A.* The validity of randomized response for sensitive measurements // *American Sociological Review*. 1981. Vol. 46. № 1. P. 187-200.
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.: Алетей, 1999.
8. *Общая психодиагностика* / Под ред. А.А. Бодалева, В.В. Столина. М.: Изд-во Моск. ун-та, 1987.
9. *Блейхер В.М., Бурлачук Л.Ф.* Психологическая диагностика интеллекта и личности. Киев: Вища школа, 1978.
10. *Богомолов Ю.П., Воронкин А.И., Куст В.П., Соколов В.П.* Краткие сведения по применению некоторых медико-психологических тестов // *Вопросы психической адаптации*. Новосибирск, 1974.
11. *Bradburn N.M., Sudman S.* Improving Interview Method and Questionnaire Design. San-Francisco: Jossey-Bass, 1979.
12. *Wortman CB., Jofus E.F., Marshal M.E.* Psychology. N.Y.: Mc Graw Hill inc., 1992.
13. *Пособие по применению ММПИ* / Сост. А.А. Рукавишников, Н.Г. Рукавишникова, М.С. Соколова. Ярославль: НПЦ "Психодиагностика", 1993.
14. *Березин Ф.Б., Мирошников М.П., Рожанец Р.В.* Методика многостороннего исследования личности. М.: Медицина, 1976.
15. *Кабанов М.М., Личко А.Е., Смирнов В.М.* Методы психологической диагностики и коррекции в клинике. Л.: Медицина, 1983.
16. *Мельников В.М., Ямпольский Л.Т.* Введение в экспериментальную психологию личности. М.: Просвещение, 1985.
17. *Толстова Ю.Н.* Анализ социологических данных. Методология, дискриптивная статистика, изучение связей между номинальными признаками. М.: Научный мир, 2000.
18. *Reamer F.G.* Protecting research subjects and unintended consequences: the effect of guarantees of confidentiality // *Public Opinion Quarterly*. 1979. Vol. 43. № 4. P. 497-506.
19. *Девятко И.Ф.* Методы социологического исследования. Учебное пособие для вузов. Екатеринбург: Изд-во Урал, ун-та, 1998.

20. *Jacoby J.* The construct validity of opinion leadership // *Public Opinion Quarterly*. 1974. Vol. 38. № 1. P. 81-85.
21. *Campbell D.T., Fiske D.W.* Convergent and discriminant validation by the multitrait-multimethod matrix // *Psychological Bulletin*. 1959. Vol. 56. № 1. P. 81-105.
22. *Cuttance P.F.* Towards a typology of information to aid reliability of response in social surveys // *Quality and Quantity*. 1986. Vol. 20. № 1. P. 27-52.
23. *Smith D.H.* Correcting for social desirability response sets in opinion-attitude survey research // *Public Opinion Quarterly*. 1967. Vol. 31. № 1. P. 87-94.
24. *Головаха Е.И., Панина И.В., Горбачик А.П.* Измерение социального самочувствия: тест ИИСС // *Социология*: 4М. 1998. № 10 С. 45-72.
25. *Makkai T., Mcallister I.* Measuring social indicators in opinion surveys: a method to improve accuracy on sensitive questions // *Social Indicators Research*. 1992. Vol. 27. № 2. P. 169-186.
26. *Bishop G.F., Oldendick R.F., Tuchfarber A.J., Bennet S.E.* Pseudo-opinions in public affairs // *Public Opinion Quarterly*. 1980. Vol. 44. № 2. P. 198-209.
27. *Паниотто В.И.* Качество социологической информации. (Методы оценки и процедуры обеспечения). Киев: Наукова думка, 1986.
28. *Саганенко Г.И.* Надежность результатов социологического исследования. Л.: Наука, 1983.