

ТЕСТ КАК ОРУЖИЕ

А.Г. ШМЕЛЕВ



Шмелев Александр Георгиевич — доктор психологических наук, профессор факультета психологии МГУ им. М.В. Ломоносова.

Работает научным руководителем в ЗАО «Гуманитарные технологии». Член Российского психологического общества, член Европейской ассоциации прикладной психологии, член РОСПРО (Российского общественного совета по развитию образования). Разработка программно-методических средств для образовательного, профессионального и психологического тестирования в Центре тестирования ЗАО «Гуманитарные технологии» отмечена Патентом на технологию тестирования «Телестинг» (Роспатент, 1998, совместно с А.Г. Серебряковым и А.Г. Ларионовым), а также сертификатом на оболочку тестирования MAINTEST в 2004 году.

Контакты: ags04@ht.ru

Резюме

Глубинную основу дискуссий вокруг тестов составляют научно-методический кризис, незрелость профессионального сообщества и кризис кадровой политики в нашей стране. В статье предлагается метафорическое сравнение теста с оружием. Подобно оружию, тесты дают преимущество их обладателю и могут нанести вред человеку. Метафора оружия позволяет понятно изложить такие характеристики тестов, как надежность, валидность, достоверность.

Большая часть недоразумений, связанных с тестами, происходит от их неквалифицированного использования.

В последние годы полемика вокруг метода тестов в нашей стране заметно оживилась. Можно даже выразиться так: обострилась. Это вызвано рядом обстоятельств, среди которых на поверхности лежит внедрение единого государственного экзамена (ЕГЭ) для выпускников

общеобразовательной школы и абитуриентов (Болотов, 2002)¹. В результате появления такого общегосударственного проекта, как ЕГЭ, дискуссия по тестам выплеснулась в СМИ и стала объектом всевозможных рассуждений и домыслов со стороны непрофессионалов —

¹ Автору этих строк довелось разрабатывать по заказу Минобразования России в 2002–2003 гг. так называемый «Портал информационной поддержки ЕГЭ» (см. адрес в Интернете: www.ege.ru). На этом портале ежегодно размещались репродукции свыше 300 статей по тематике ЕГЭ, среди которых добрая половина содержала всегда остро дискуссионные доводы «за» или «против» самого метода тестирования в российском образовании.

лиц, не имеющих специальных знаний в области психологии и педагогики (прежде всего со стороны журналистов).

Истинные причины

В глубинной основе дискуссии вокруг тестов, по моему убеждению, кроются другие, более глобальные причины:

- а) научно-методический кризис;
- б) незрелость профессионального сообщества;
- в) кризис кадровой политики в стране.

Все три указанные причины тесно взаимодействуют друг с другом, так сказать, переплетены и взаимообусловлены.

На рубеже XX и XXI столетий в нашей стране обострился дефицит эффективных методов психологической и психолого-педагогической диагностики. Этот дефицит возник не вчера и даже не позавчера. Стало банальностью ссылаться на постановление ЦК ВКП(б) 1936 г. «О педологических извращениях в системе Наркомпроса», которое затормозило развитие отечественных работ в области психологических измерений на полстолетия, но факт остается фактом: к концу XX века Россия оказалась на задворках общемирового процесса развития тестологии, тестовых технологий, а главное — развития тестологической культуры среди пользователей тестов. Об этом говорит фактическое отсутствие востребованных практикой оригинальных отечественных тестовых методик, удовлетворяющих всем психометрическим требованиям — репрезентативности, надежности, валидности,

достоверности (см.: Бодалев, Столин, 1988; Шмелев и др., 1996). Только в 1982 г. появилось первое серьезное переводное издание на русском языке — книга А. Анастаси «Психологическое тестирование» (Анастаси, 1982). Однако как только запрет на тесты был снят, события стали повторяться в печальном соответствии с постановлением 1936 г. Это выразилось в том, что невзыскательные, неосведомленные, а проще говоря, невежественные практические работники уже в 80-е гг. немедленно кинулись тиражировать и использовать западные методики, не прошедшие экспериментальной проверки в российских условиях. Возникла массовая практика применения тестов в нижних слоях государственной и социальной пирамиды (в профотборе на массовые специальности, при проведении медико-психологических обследований населения и т. п.).

Наряду с этим в те же 80-е гг. наблюдалась встречная тенденция — к отказу от применения метода тестов при отборе и аттестации влиятельных служащих госаппарата, в частности, руководящих работников силовых структур. Совсем недавно в «Российской газете» была опубликована статья М. Виноградова (Виноградов, 2003), в которой, конечно, присутствует изрядный элемент драматизации и сверхполитизации, но одновременно с этим и констатация определенных малоизвестных фактов — о том, какие причины привели к закрытию в 80-е гг. возглавляемой М. Виноградовым лаборатории психофизиологических исследований в системе МВД. Не имея возможности ни подтвердить, ни оп-

ровергнуть сведения и оценки, данные М. Виноградовым, я считаю тем не менее, что данная статья свидетельствует об объективной закономерности: тест как инструмент объективного контроля уровня развития профессиональных качеств, и особенно как инструмент выявления противопоказаний к работе, воспринимается коррумпированными высокопоставленными лицами как угроза личной безопасности и устойчивости их положения во властной иерархии. По-видимому, так было и так будет во все времена. Недаром подноготная постановления 1936 г. также окутана слухами и догадками о том, что практика тестирования воспринималась как угроза личным интересам лиц, принадлежавших в то время к правящей верхушке.

В 90-е гг. прошлого века положение усугубилось жесточайшим общесистемным и финансовым кризисом, в частности, кризисом финансирования отечественной науки. Производство отечественных тестов погибло, фактически не поднявшись из младенческой колыбели. Оказалось гораздо выгоднее пользоваться краденными западными методиками, чем оригинальными отечественными, за которые отечественные специалисты закономерно хотели получить компенсацию вложенного квалифицированного труда. Возникшую ситуацию можно в какой-то мере сравнить с вытеснением отечественных фильмов из кинопроката, хотя уровень развития отечественной киноиндустрии был к тому времени гораздо выше, чем уровень российской тестологии. Неоправданные массовые тиражи тестовых сборников (начиная с «Лучших психологических тестов»,

1992) не в последнюю очередь провоцировались потребностью научных работников в элементарном выживании «любой ценой» — ценой копеечных гонораров за разглашение профессиональных секретов и публикацию сырых исследовательских версий переводных методик, еще не подготовленных для практического использования. Редкие попытки выстраивания цивилизованных отношений с западными производителями тестов сталкивались на внешнем рынке с несоразмерными (с курсом рубля) ценами на лицензии в долларовом выражении, тогда как на внутреннем рынке такие попытки просто не могли выдержать конкуренции с пиратской практикой торговли краденой интеллектуальной собственностью (см. об этом мои заметки в «Психологической газете»: Шмелев, 1996, 1999). В этих условиях не только разработка оригинальных тестов, но и работа по серьезной психометрической адаптации западных тестов — наиболее популярных в России — оказалась экономически невозможной.

Но означает ли перечисленное выше, что само понятие «тест» опять следует предать анафеме и снова пойти по тому же замкнутому кругу — с возвратом в 1930-е гг., по которому мы (как страна, как гильдия профессионалов в области психологии и педагогики) уже ходили? Увы, попытки такой огульной критики метода тестов и сейчас наблюдаются. К подобной критике прибегают отнюдь не только движимые так называемым «гуманизмом» непрофессионалы, но и сами обладающие всеми академическими регалиями ученые-психологи (Холодная, 1999).

По моему убеждению, чтобы разорвать этот исторически порочный круг, надо прежде всего самим специалистам глубже разобраться в понятии «тест», осознать статус этого инструмента познания в широком социальном контексте, границы и правила его использования — не во вред, а во благо людей.

Основная метафора

Выражение «тест — оружие» — не более чем метафора, так как вовсе не претендует на полноту и строгость совпадения по всем признакам. Тем не менее, по моему мнению, эта метафора имеет определенный эвристический потенциал, по крайней мере помогает разъяснить ряд принципиальных моментов.

В чем сходство теста и оружия?

Оружие дает преимущество тому, кто им обладает. Любая процедура получения знания о человеке дает определенные преимущества обладателю этого знания. Психологический тест — не исключение. Сюда же можно отнести медицинский тест, профессиональный, образовательный тест и другие. Однако психологический тест воспринимается людьми даже с большей опаской, чем прочие. По-видимому, сказывается имплицитное (не вполне осознанное) распространённое представление о психологических свойствах, выявляемых психологическими тестами, как относительно более неизменных на протяжении жизни, чем такие временные явления, как болезнь (предмет медицинской диагностики) или отсутствие определенных знаний и

навыков (предмет педагогической диагностики).

Тест — стандартизированное испытание, краткое в проведении и формализованное в обработке результатов. Это инструмент, подобный простому измерительному прибору: пустил в ход и получил результат. Пользователю кажется, что психологический тест не требует особой подготовки, в отличие от более сложных при проведении нестандартизованных психодиагностических методик. Особенно простым и привлекательным в этом смысле кажется компьютерный тест: нажал кнопку на клавиатуре компьютера, посадил испытуемого и получил на экране результат. Это напоминает выстрел из пистолета: достаточно дослать патрон и нажать спусковой крючок.

Конечно, полной аналогии между тестом и оружием нет и быть не может. Тест не приводит к физическому ущербу, не уничтожает испытуемого. Однако нанести весомый моральный ущерб с помощью теста можно, а если на основе теста принимаются серьезные кадровые решения (отсев кандидатов при профотборе, увольнение по результатам аттестации с использованием тестов), то ущерб оказывается не только моральным, но и вполне материальным.

Причины бесконтрольности

Основной смысл сравнения теста и оружия для меня заключается в том, чтобы на этой основе высветить проблему контроля за распространением тестов. Зададимся вопросом: почему круг лиц, имеющих разрешение на ношение оружия, ограничивается в нашей стране законода-

тельно, тогда как тест может использовать любой гражданин? Причин тому несколько.

1. **Забава.** Одна из причин выглядит весьма банальной и простой: благодаря популярным журналам многие граждане в нашей стране воспринимают тест скорее как забаву, как инструмент развлечения, и, где проходит грань между игрушкой и оружием, в данном случае многие просто не ведают.

2. **Незрелость профессионального сообщества психологов.** Эта вторая причина, на мой взгляд, более существенна. Государственные силовые ведомства, призванные контролировать распространение огнестрельного оружия, в нашей стране имеют гораздо более древнюю историю, весьма авторитетны, лучше институализированы, поэтому обеспечивают для своих работников особый статус — в отличие от профессионального психологического сообщества, которое фактически ни разу в истории России (или СССР) не издало перечня профессиональных тестов, право пользования которыми было бы ограничено кругом профессионалов (членов сообщества). Юридическое, законодательное оформление правил обращения с тестами пока не выглядит актуальной задачей для большинства российских законодательных институтов, включая Госдуму. Пожалуй, наиболее продвинутым здесь является Минздрав, издавший несколько внутриведомственных актов по вопросам медико-психологического тестирования, тогда как в системе Минобразования до сих пор, насколько это известно автору, не сформулировано ясного различия между педагогическими теста-

ми и психологическими — теми, которыми может пользоваться любой педагог, и теми, которыми может пользоваться только школьный психолог.

3. **«Великий уравниватель».** Третья причина заключается в позиции самих членов психологического сообщества. Вспомним про «великий уравниватель», которым американцы назвали в XIX веке кольт. Большинство влиятельных психологов, по моим наблюдениям, вовсе не хотят, чтобы в их профессиональной среде появлялись подобные «великие уравниватели»: мол, лучше уповать на собственный интеллектуальный потенциал, чем производить и распространять интеллектуальные инструменты, которые могут рано или поздно уравнивать тебя с другими пользователями подобного инструмента. Поэтому нечетко оформленная профессиональная гильдия психологов в нашей стране напоминает сообщество любителей боевых искусств, где все члены ранжируются по силе мышц и телесной ловкости, в разной мере развитой у разных людей, но не по уровню владения стандартным «табельным оружием».

4. **Доводы гуманистов.** Тест рассматривается как инструмент, якобы угрожающий суверенитету личности, ущемляющий права человека. При этом, по моему мнению, возникает грандиозное смешение понятий вплоть до переворачивания ценностной иерархии с ног на голову (об этом автор уже писал в популярной прессе несколько лет назад в ходе дискуссии с В. Аванесовым, С. Хайтуном и М. Холодной в «Независимой газете», см.: Шмелев, 1999). Сегодня, мне кажется, вопрос следует

сформулировать максимально ясно и остро: разве тестирование должно проводиться всегда в интересах самого испытуемого? Это, конечно, именно так в ситуации индивидуального консультирования, но в организационном контексте (т. е. когда тестирование проводится в так называемой ситуации экспертизы, а не в ситуации консультирования клиента) вовсе не так! Тест в организационном контексте — это инструмент защиты интересов третьих лиц от ущерба, который им может нанести испытуемый. Возьмем тест на водительские права. Если он показывает, что испытуемый не знает правил дорожного движения, то разве при этом интересы испытуемого не должны пострадать? Должны!² Иначе пострадают интересы пешеходов и других водителей. Ситуацию так просто понять в случае дорожного движения, но почему-то вовсе не просто в случае тестирования на пригодность к выполнению ответственной работы по экономическому или политическому управлению (крупной корпорацией, политической партией, страной). А разве тесты в этом случае должны проводиться в интересах тестируемого? Нет, они призваны защитить общество от про-

никновения на вершины общественной иерархии лиц с интеллектуальными и морально-личностными дефектами. В этом контексте по своему смыслу тесты превращаются в оружие общества против произвола отдельной личности³.

Четвертая причина — «ущерб интересам личности» — гораздо чаще указывается критиками тестов, чем другие. А третью причину вообще вряд ли кто-то из самих психологов назовет вслух. Будет приводиться масса других аргументов, включая трогательную заботу об интересах испытуемых, но, по моему мнению, самая важная причина сводится к интересам самих психологов, осмысленных ими таким образом. Психологам кажется предпочтительнее социальная роль «кустарей-одиночек» (чтобы избежать уничижительного тона, давайте назовем это ролью «неподражаемых маэстро»), чем роль членов хорошо структурированного профессионального сообщества, наработавшего собственный инструментарий и ограничившего круг его распространения собственными границами.

Границы распространения — этот вопрос в случае с психологическими тестами стоит с особой остротой.

² При этом не ближайшие, но отдаленные интересы испытуемого, конечно же, не страдают, но оказываются защищены: ведь за преступления на дороге, совершенные из-за незнания правил, самому водителю также приходится строго отвечать.

³ Может быть, и не надо было бы царедворцам Бориса Годунова брать страшный грех на душу и подстраивать в Угличе самоубийство царевича Дмитрия, если бы в то время в России существовал защищенный демократической властью институт медико-психиатрической экспертизы наследников престола, который просто установил бы диагноз «патологически педагогически-некурабельный генетически-детерминированный садизм», после которого царевич смог бы жить спокойной частной жизнью и периодически истязать только мелких домашних животных, но никак не подданных огромной страны.

Ведь дело в том, что в отсутствие таких «охраняемых границ» большинство психологических тестов просто не работают, перестают быть полезным инструментом. Если вся популяция потенциальных испытуемых знает о тесте то, что положено знать лишь профессиональным пользователям, то тест лишается необходимого элемента секретности, перестает работать. В этом плане особой профессиональной группой, враждебной психологам по своим интересам, фактически являются журналисты. В 90-е гг., например, в журнале «Космополитен» были опубликованы материалы, раскрывающие сотням тысячам читателей этого журнала секретные ключи к таким методикам, как «Рисунок несуществующего животного». Узкий круг разработчиков этого теста мог десятилетиями набирать тонкий слой проверенных экспериментально диагностических признаков («размеры зубов», «размеры глаз» и т. п.), но достаточно всего одной публикации, чтобы ключи к тесту были разглашены и он тут же из инструмента профессионала превратился в игрушку-забаву для любителей салонных развлечений. Интересно, что когда я позвонил в редакцию этого журнала с попыткой узнать, кто именно из наших коллег представил этот материал в редакцию, то натолкнулся на стойкое сопротивление журналистов, не выдавших имя осведомителя-перебежчика. Увы, мне трудно вспомнить пример подобной профессиональной стойкости у самих психологов ради сохранения в неприкосновенности собственных корпоративных профессиональных секретов. По-моему, психологи с гораздо

большим усердием готовы разоблачать (и предавать) именно своих коллег, чем представителей каких-то других профессиональных групп. Именно к психологам как нельзя более приложим один из постулатов общей теории конкуренции: переживание суженного пространства для выживания (сжимающейся экологической ниши), дефицит адаптивности в межвидовой конкуренции приводит к обострению внутривидовой конкуренции (Шмелев, 1997).

Свойства теста

Какие другие важные следствия мы можем вывести из метафоры «тест — оружие»? Эта метафора позволяет нам точнее и глубже осознать ряд инструментальных требований к тестам, которым тесты должны соответствовать, а также нормативы применения тестов. Я вовсе не собираюсь здесь перечислять все психометрические свойства тестов, но все-таки некоторые наиболее важные стоит упомянуть — хотя бы не строго, хотя бы чисто метафорически.

1. **Надежность теста.** Может ли быть надежным оружие, изготовленное в кустарной полуподвальной мастерской, как говорится, «на коленках»? Это оружие будет стрелять куда попало — иногда в цель, но чаще вбок, а иной раз может и просто разорваться в руках стреляющего. Тут уместно напомнить следующее: надежные тесты не создаются в крошечных лабораториях (и тем более за письменным столом автором-одиночкой). Надежность теста не только проверяется на репрезентативной

(массовой) выборке, но просто не вырабатывается без обширной статистики. Репрезентативная выборка для стандартизации теста — это своеобразный полигон для обстрела нового оружия. Только после таких полевых испытаний конструктор теста может внести целенаправленные («зрячие») коррективы в первоначальную конструкцию своего оружия. Тем самым уже на примере этого одного свойства теста — надежности — мы видим, что нам дает в этом контексте метафора «тест — оружие». Плохое оружие не усиливает, а, наоборот, ослабляет пользователя, подвергает его риску. Но разве можно судить о качестве оружия вообще по образцам кустарного оружия? Плохими являются не тесты вообще, а ненадежные тесты.

2. Валидность теста. Напомним, что это мера пригодности теста целям психодиагностики, мера соответствия измеряемому свойству. Куда будет стрелять оружие? Это зависит не только от надежности самого теста, но и от пользователя. Ненадежный тест не может быть валидным. Эту аксиому теории измерения в данном контексте легко понять: если вы не попадаете с пяти шагов в силуэт, то о какой валидности, о каком соответствии теста измеряемому свойству может идти речь, ведь вы же можете попасть с помощью такого «теста» не во врага, а в «своего» — того, кто рядом стоит, то есть «цепляете» с помощью теста не целевое, а иное психическое свойство. Но если сам стрелок слепой, если он дальтоник, который не различает цвета мундиров, в которые одеты свои и чужие, если он к тому же паникер, то будет в панике палить даже из на-

дежного стрелкового оружия и по своим, и по чужим. Таким образом, мы легко формулируем важное следствие: тест не может быть валидным в руках непрофессионала. Вот вам и еще одна аксиома тестологии, которую, увы, так трудно бывает объяснить не только массовой аудитории, но и самим психологам, ибо при словах «надежность» и «валидность» в их сознании всплывают страшные и непонятные психометрические формулы. Поэтому эти понятия кажутся им скорее математическими, чем психологическими, то есть чуждыми их «гуманитарному интеллекту».

Опять же вернемся в этом контексте к критике тестов. Можно ли судить о тесте и, тем более, тестах вообще, если даже вполне качественное фабричное оружие передано в руки новобранцев-паникеров, которые то стреляют из пушки по воробьям (например, применяют тяжелую батарею IQ вроде теста Векслера для диагностики дефицита внимания), то бросаются с пистолетиком тщетно обстреливать бронированный танк (пытаются понять природу и содержательный смысл внутреннего конфликта по цветовым предпочтениям в тестике Люшера, пригодного, по моему убеждению, лишь для грубой оценки фона настроения). Любому мало-мальски сведущему в военном деле человеку как дважды два понятно: нет универсального оружия, и в разных условиях боя надо применять разное. Но психика человека — более тонкая, невидимая стороннему взгляду реальность, чем поле боя. И вот мы путаем все на свете: вялую позиционную перестрелку, активную артподготовку и яростную

штыковую атаку в полный рост, когда пора доставать из-за пояса гранаты. Когда вы проводите какую-нибудь очень краткую пробу из нескольких заданий (несколько спрятанных фигур из теста Готтшальда, несколько чернильных пятен Роршаха), то следует все-таки отдавать себе отчет в том, что вы с такой же вероятностью наткнетесь на диагностически ценную информацию, с какой можно поразить стальной ДОТ с помощью легкой пехотной гранаты. Результата, скорее всего, не будет никакого! Только следует ли после этого делать вывод о том, что все тесты неэффективны? Я бы сказал, что многие одиночные психологические тесты — это очень слабое оружие против хорошо замаскированных укреплений, против глубоко эшелонированной обороны многоэтажной человеческой психики, которая нарабатывает ко времени социальной зрелости многие слои из весьма изощренных механизмов психологической защиты. Тут мы подходим к проблеме достоверности — проблеме соотношения сознательных и неосознаваемых механизмов психологической защиты от тестирования. Р. Кэттелл назвал это в свое время проблемой мотивационных искажений. Звучит красиво, хотя речь идет о некрасивых вещах — о более или менее осознаваемой лжи.

3. Достоверность. Это проблема фальсификаций. Сформулируем в этом контексте такой несколько парадоксальный профессиональ-

но-этический норматив: «Испытуемый имеет право на ложь». В самом деле, если тест — это оружие проникновения в человеческую психику, то испытуемый имеет право на самозащиту — на то, чтобы сопротивляться этому проникновению. В конце концов, можно оправдать испытуемого, сумевшего скрыть свои проблемы, свои дефекты, мобилизовавшись на социально-желательное выполнение теста: таким способом он проявляет в момент тестирования силу своих компенсаторных механизмов, умение решать задачки на моральное развитие, умение решать задачки на интеллектуальное развитие и т. п.⁴, хотя, возможно, в повседневной жизни он ведет себя во все и не так. Прочность бронированного корпуса его судна, обеспечивающая ему непотопляемость, оказалась сильнее того удара, который психолог нанес из своего оружия. Честь и хвала такому испытуемому. Но этот тезис имеет и такое важное следствие: положительные результаты тестирования имеют меньшую ценность, меньшую предсказательную силу, чем негативные результаты.

Таким образом, если мы наконец разберемся в базисных представлениях о сущности теста, мы научимся адекватным образом его применять в социальной практике. До тех пор, пока мы неверно трактуем сущность теста, не видим адекватным образом ограничений в практике его использования, мы совершаем серьезные

⁴ При этом, конечно же, не имеется в виду право испытуемого на совершение подлога — путем использования, например, тех же шпаргалок при выполнении интеллектуальных и образовательных тестов.

ошибки. Нужно ли запрещать распространение оружия в обществе, где никто толком не умеет грамотно им пользоваться? По-видимому, все-таки разумнее не запрещать вовсе, а разумно ограничить более узким кругом подготовленных аттестованных пользователей! И предоставлять им следует только сертифицированные инструменты, а не какие попало. Если горе-строители возводят многоэтажные здания на болотах или зыбучих песках без закладки прочного фундамента, т. е. нарушают все правила технологии безопасного строительства, то таким образом здание не следует строить вообще; это не значит, что следует запретить и архитектурные институты, и все заводы по производству стройматериалов, и сами строительные организации. Если кто-то использует определенные медицинские препараты не по назначению, превращая их в наркотики, то это не значит, что следует запретить фармацевтическую промышленность, хотя строгость в контроле за распространением опасных

медицинских препаратов повысить, конечно, придется.

Тесты и экспертные оценки

По моему убеждению, стандартизованные тесты не дают основания для окончательного положительного диагноза (т. е. диагноза о пригодности к определенной деятельности), для этого они должны быть дополнены экспертными оценками (или другими менее стандартизованными диагностическими процедурами, включающими экспертные оценки в той или иной мере, как это, например, происходит в проективных методиках).

Таким образом, позитивный исход тестового испытания выступает логически необходимым, но не достаточным условием для окончательного позитивного заключения. Поскольку мне как тестологу, увы, хорошо известно, что с элементарной логикой у наших сограждан порой возникают серьезные проблемы⁵, схематизируем сказанное в виде следующей таблички.

	Вывод о пригодности	Вывод о непригодности
Позитивный исход теста	Нельзя сделать	Нельзя сделать
Негативный исход теста	Нельзя сделать	Можно сделать

⁵ Это показали, в частности, тестологические обследования десятков тысяч наших выпускников школ и абитуриентов вузов в рамках проекта компьютеризованного тестирования «Телестинг» (Шмелев, 2000): наши российские учащиеся в массе своей демонстрируют поразительную беспомощность при решении простейших текстовых логических задач, хотя при этом «щелкают как орешки» сложные математические уравнения (правда, опять-таки не имея ни малейшего представления о том, какие физические процессы могут быть описаны с помощью этих формально-математических моделей).

Поясним это на содержательном примере. Вначале возьмем самый тривиальный случай, далекий от психологии — уже упомянутый экзамен на знание правил дорожного движения. Если кандидат сдал тест по правилам, то ему еще нельзя выдавать права: он должен пройти после этого менее формализованный экзамен на практическое вождение. Если же кандидат провалил тест, то он не допускается до следующего испытания. В этом контексте самое время также сделать и такую оговорку: отрицательный результат тестирования — не приговор. Всем понятно, что правила можно выучить, прийти снова и пересдать экзамен.

Возьмем теперь менее очевидную (не оформленную пока нормативно) процедуру тестирования кандидата при приеме на работу на предмет уровня так называемой «корпоративной лояльности». Предположим, что испытуемому при этом предъявляется совершенно примитивный тест-опросник, содержащий лобовые вопросы типа «Вам никогда не доводилось обманывать учителей при сдаче экзаменов в школе?» Как мы говорили выше, испытуемый в этом случае использует свое право на фальсификацию и отвечает: «Вер-

но, не доводилось». И какой вывод в данном случае мы сделаем? Никакой! Но если испытуемый вдруг сам в порыве откровенности отвечает: «Неверно, доводилось», — то по крайней мере насторожиться следует.

В еще большей степени этот принцип относится к базовым тестам на элементарные профессиональные знания. Если кандидат в бухгалтеры не может ответить на вопрос конкурсного теста-опросника о том, что такое «план счетов», то следует ли дальше заниматься с этим кандидатом? Следует ли тратить дорогое время квалифицированных экспертов на детальное интервьюирование подобного кандидата? Конечно, нет⁶.

Таким образом, я предлагаю буквально везде, во всех отраслях практики использовать тест как первичный дешевый и формализованный фильтр, предшествующий применению более сложных и дорогих экспертных процедур. В какой-то мере подобной логикой руководствуются в настоящее время специалисты по оценке персонала, которые применяют технологию «Ассесмент-центр» (2003).

Таким образом, приведенную выше табличку следует изменить таким образом:

	Позитивный исход экспертной оценки	Негативный исход экспертной оценки
Позитивный исход теста	Вывод о пригодности	Вывод о непригодности
Негативный исход теста	Вывод о непригодности	Вывод о непригодности

⁶ Только не следует примитивно трактовать подобные примеры. Автор, конечно, не предлагает делать выводы на материале одного вопроса. Выводы следует делать путем подсчета ответов на несколько десятков вопросов, обеспечивая статистическую достоверность.

Как видим, для позитивного общего заключения требуется конъюнкция (логическое «И») двух независимых событий — позитивного тестового исхода и позитивного исхода экспертной оценки. Отсутствие хотя бы одного из позитивных исходов не дает возможности делать общий позитивный вывод.

Качество такой двухфилтровой системы отбора в любом случае выше, чем любой однофилтровой — основанной только на экспертных оценках или только на тестах. А разговоры о том, что в нашей стране результаты тестирования очень легко купить (увы, такие разговоры часто затевались, например, на дискуссионном форуме портала ЕГЭ), носят либо сознательно демагогический характер, либо опять-таки выявляют дефект логического мышления. Там, где можно купить результаты тестирования, как правило, можно купить и результаты экспертной оценки, и нужно еще специально изучать, какой из фильтров по факту является менее продажным. Если даже при широком распространении теста происходит утечка ключей, негативный исход тестирования продолжает сохранять свою ценность, но особенно важно, чтобы после позитивного исхода в дело вступали неподкупные эксперты. Если мы связываем результаты двух процедур логическим «И», то числовые результаты теста и экспертной оценки правильнее не суммировать, но умножать, то есть агрегировать не аддитивно, а мультипликативно:

$$O = T \times E,$$

где T — результат теста, E — результат экспертной оценки,

O — общая оценка. Если любой из сомножителей принимает нулевое значение (оказывается ниже минимального порога), то общий результат оказывается нулевым независимо от значения второго сомножителя.

Хороший тест лучше среднего интервьюера

К сожалению, в наших дебатах вокруг тестов редко апеллируют к результатам исследований сравнительной прогностической эффективности тестов и нетестовых диагностических процедур. Нет ничего удивительного в том, что наши специалисты просто не располагают такими данными, ведь подобные исследования очень дорогостоящи. За рубежом подобные исследования неоднократно проводились на весьма репрезентативных выборках испытуемых. Таблицы сравнительной эффективности (для прогнозирования профессиональной деятельности) различных тестов и такой процедуры, как интервью, публикуются практически в любом западном учебном пособии по организационному поведению (Fincham, Rhodes, 1998).

Суммируя различные источники, я бы так оценил в настоящее время выявленные коэффициенты прогностичности:

Процедура	Прогностическая валидность
Интервью	0,1–0,3
Психологическое тестирование	0,3–0,5
Квалификационное тестирование	0,4–0,6
Ассессмент-центр	0,5–0,7

Почему же столь низкой оказывается валидность обыкновенного интервью? Все дело в том, что извлечение диагностически значимой информации из интервью требует не только значительного опыта интервьюирования (личного проведения сотен и тысяч интервью), но и определенного искусства, куда входит умение владеть собственными эмоциями, не выражать их резко, чтобы не индуцировать какое-то ситуационное атипичное ответное эмоциональное состояние у интервьюируемого (настороженность и замкнутость, лицемерную любезность, развязность, восторженно-прекраснодушное воодушевление и т. п.). Живое общение людей (каким бы объективным профессионалом ни старался быть интервьюер) — это всегда вольный или невольный обмен эмоциями, который создает весьма различный фон в зависимости от ситуационного коммуникативного контекста, возникшего здесь и теперь в той или иной диаде (группе) общающихся друг с другом людей. На практике выясняется, что все разговоры о том, что «точнее всего понимаешь человека, когда посмотришь ему в глаза», как правило, лишь прикрывают стремление отстоять режим личной власти над судьбами.

Соотношение эффективности формализованных и неформализованных процедур в области психодиагностики можно сравнить с ситуацией противостояния шахматиста и шахматной программы. Хотя гроссмейстеры успешно конкурируют с шахматными суперкомпьютерами, средней руки разрядники (не говоря про новичков и любителей)

проигрывают хорошей компьютерной программе весьма устойчиво. Имеется в виду программа, которая опирается, например, на компьютерный банк отработанных дебютов. Ведь подобный дебютный репертуар выработал коллективный разум сотен и тысяч профессиональных шахматистов в течение всей писаной истории шахмат.

Так получается и с тестами. Плохой тест может уступать среднему интервьюеру, но хороший тест, прошедший психометрическую апробацию на сотнях испытуемых, показавший прогностическую валидность, превосходит среднего одиночного интервьюера (подчеркнем здесь слово «одиночного», так как коллектив интервьюеров всегда работает эффективнее). Особенно это касается многофакторных тестов. Сколько параметров (характеристик) индивидуальности способен оперативно отслеживать интервьюер? Опросы показывают, что всего лишь чуть больше, чем обычный человек, — в пределах 5–7 (что-то похожее на магическое число 7 ± 2), и в этом случае уже батарея из 12–16 факторов дает значительное преимущество.

Более высокое положение в указанной таблице методики «Ассесмент-центр» вполне объяснимо как раз в логике совмещения тестов и экспертных оценок, о которой мы говорили в предыдущем параграфе: здесь имеет место сочетание методов тестирования и интервьюирования, плюс деловые игры с множественной экспертной оценкой (привлекаются независимые судьи, способные сообща отследить больше параметров).

Выводы

Подводя итоги этой статьи, еще раз подчеркну главный тезис: нынешние дискуссии по поводу метода тестов, по моему убеждению, часто оказываются непродуктивными из-за того, что участники дискуссии (причем неважно, сторонники это тестов или противники) не вполне адекватно воспринимают статус теста как инструмента. Тест является инструментом не только научного исследования, не только индивидуального консультирования и помощи, но и социальной практики в малых и больших организациях. Большая часть недоразумений связана даже не с качеством (увы, подчас низким) разработанных тестов, но с их неправильным, некачественным использованием. Значительная часть пользователей тестов в России до сих пор руковод-

ствуется нетехнологичными представлениями о возможностях тестов, можно даже сказать, мифологическими. Значимость тестовых результатов либо резко переоценивается, либо недооценивается, причем некоторые пользователи приписывают ответственность за ошибки при тестировании разработчикам, не понимая, что это они сами — прежде всего сами — несут ответственность за выбор и адекватное применение инструмента.

Тест — это оружие. И пользоваться им надо умело. Надо правильно заряжать это оружие и нацеливать в нужную сторону. Как нет универсального оружия, пригодного во всех ситуациях, так нет и универсальных тестов. Тест больше полезен, когда включен в контекст более широких процедур, использующих, кроме тестирования, экспертные оценки.

Литература

Анастаси А. Психологическое тестирование. В 2 т. М.: Педагогика, 1982.

Бодалев А.А., Столин В.В. (ред.) Общая психодиагностика. М.: Изд-во МГУ, 1987.

Болотов В.А. (ред.) Единый государственный экзамен. Научные основы, методология и практика организации эксперимента. М.: Логос, 2002.

Виноградов М.В. Кто не станет оборотнем // Российская газета. 2003. 11 дек.

Лучшие психологические тесты. Петрозаводск, 1992.

Холодная М.А. Тестирование и права личности // Независимая газета. Приложение «Наука». 1999.

Шмелев А.Г. (ред.) Основы психодиагностики. Ростов н/Д: Феникс, 1996.

Шмелев А.Г. Дилетанты и шарлатаны // Психологическая газета. 1996. № 4.

Шмелев А.Г. Продуктивная конкуренция. М.: Открытое общество, Магистр, 1997.

Шмелев А.Г. «Каша из топора», или История адаптации 16PF в России // Психологическая газета. 1999. № 5.

Шмелев А.Г. Тесты: не угроза, а защита прав личности // Независимая газета. 1999. 17 нояб.

Шмелев А.Г. Телетестинг. Тесты для старшеклассников и абитуриентов. М.: Первое сентября, 1999–2000.

Fincham R., Rhodes P. Principles of organisational behavior. Oxford: Univ. Press, 1998.