

## СОВРЕМЕННЫЕ ТЕХНОЛОГИИ И ПРОБЛЕМА ОБЪЕКТИВНОСТИ ПСИХОЛОГИЧЕСКОГО ТЕСТИРОВАНИЯ

*Д.В. Ушаков*

Сегодня у тестируемого, например, в ситуации отбора существует немало способов уйти от объективной оценки: опросники часто понятны в своей направленности, а задачные тесты, имеющие 1–2 варианта, можно найти в магазине или Интернете и выучить. В статье разбираются современные способы решения этой проблемы: тесты, генерируемые компьютером, психофизиологическое тестирование и диагностика по речи.

*Ключевые слова: объективность тестирования; тесты, генерируемые компьютером; психофизиологическое тестирование, диагностика по речи.*

### **Проблема объективности психологического тестирования**

История наиболее мощного психологического инструментария – тестов интеллекта – перевалила уже за 100 лет, считая от пионерской работы Бине и Симона. В 2005 г. во Франции прошла конференция, посвященная этому событию. Долгую историю имеют и личностные опросники, и проективные тесты. За это время создано много полезных инструментов и сложилась разработанная психометрическая теория. Ни первое, ни второе, однако, не избавляют психологическое тестирование от некоторых серьезных проблем. Ниже речь пойдет об одной из них, особенно существенной в плане практического применения. Эта проблема заключается в том, что у тестируемого при существовании на то желания есть немало способов ввести тестирующего в заблуждение.

Начнем рассмотрение этой проблемы с личностных опросников. Для них отработаны способы внешней валидации по группам населения с высокой выраженностью соответствующего признака, для чего особенно часто используется психопатология. Например, шкала паранояльности оценивается по группе параноиков, а опросник эмоционального интеллекта – по группе алекситимиков. Способ этот, безусловно, весьма эффективен, однако не учитывает одного обстоятельства – мотивация заполнения опросника параноиком, находящимся в кабинете психолога психиатрической клиники, совсем не та, что, например, у кандидата на должность менеджера по продажам или сотрудника службы безопасности фирмы, который проходит тестирование для поступления на работу или в порядке ассесмента пер-

сонала. Во втором случае совершенно естественным является желание предстать перед работодателем в наиболее выигрышном свете, продемонстрировать наличие у себя тех качеств, которые обеспечивают эффективную работу на соответствующей должности (профессионально важных качеств, если воспользоваться более профессиональной терминологией). В первом случае аналогичное желание отнюдь не очевидно. Можем ли мы заранее отрицать здесь даже прямо противоположное – выставить в наиболее ярком виде симптомы своего недуга? Вряд ли. Во всяком случае, необходимо очень осторожно подходить к проблеме сравнения результатов тестирования при разной мотивации. Студенты-психологи, на которых проводится обкатка многих тестов, мотивированы принципиально иначе, чем тестируемые в бизнес-структурах. Отсюда возникает мнение о психологическом тестировании, выраженное одним моим знакомым, прошедшим тестирование при поступлении на работу в банк: «Мне дали тест, в котором на 100 разных ладов спрашивалось, не буду ли я воровать. Ну, конечно, ответил, что не буду. Что они, на дураков рассчитывают, что ли?» К сожалению, понять смысл психологических опросников часто не составляет проблемы для тестируемого. Возможно, в этом заключена основная причина невысоких корреляций показателей психологических опросников и индексов эффективности на рабочем месте.

Задачные тесты интеллекта и креативности, как и проективные личностные тесты, лишены отмеченного выше недостатка. Сколь хорошо ни понимает тестируемый, что оценивается, допустим, его интеллект, это не

поможет ему найти правильный ответ. Результаты тестов интеллекта не зависят от осознания тестируемым их направленности. Здесь, однако, есть другая проблема, особенно остро стоящая в нашей стране – проблема отсутствия защищенности теста от раскрытия. Стоит тестируемому получить заранее доступ к тому же тесту интеллекта и его ключам, как ничто не мешает показать ему тот результат, который он пожелает.

В нашей стране проблема приобретает дополнительную остроту. Во-первых, у нас очень невелико количество тестов интеллекта, используемых на практике. В 80% случаев это «Векслер», «Равен», «Амтхауер», «Кеттелл» или «КОТ» (о тестах Г. Айзенка, публикуемых с советского времени в популярных книжках самого автора, говорить не стоит). Стоит познакомиться с этим инструментарием, и Вы будете поражать воображение потенциальных работодателей «заоблачными высотами» интеллекта. Во-вторых, наши психологи в 1990-е годы сами лишили себя инструментального оснащения, опубликовав сборники тестов с ключами, о чем уже неоднократно говорилось как в научных, так и популярных изданиях. А.Г. Шмелев приводит следующий типичный пример: «В 90-е гг. ... в журнале «Космополитен» были опубликованы материалы, раскрывающие сотням тысяч читателей этого журнала секретные ключи к таким методикам, как «Рисунок несуществующего животного». Узкий круг разработчиков этого теста мог десятилетиями нарабатывать 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 тест тут же из инструмента профессионала превратился в игрушку-забаву для любителей салонных развлечений. Интересно, что когда я позвонил в редакцию журнала с попыткой узнать, кто именно из наших коллег представил этот материал в редакцию, то наткнулся 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 других профессиональ-

ных групп. Именно к психологам как нельзя более приложим один из постулатов общей теории конкуренции: «переживание суженного пространства для выживания (суживающейся экологической ниши), дефицит адаптивности в межвидовой конкуренции приводит к обострению внутривидовой конкуренции» (Шмелев, 2004, с. 46).

Итак, проблема может быть сформулирована следующим образом: у психологов есть сегодня достаточно серьезный репертуар тестов, в отношении которого, однако, тестируемый при желании может пустить в ход «противооружие». Этот тезис особенно актуален в сфере бизнес-тестирования, где у тестируемого есть все основания к приукрашиванию своего результата.

Психолог, применяя сегодня кажущийся самым надежным и валидным тест и получая результат путем корректных и тщательных процедур, не может быть уверен в том, что против его теста не было использовано «противооружие», и в том, что полученный в конечном счете результат не сохнет с толку работодателя.

Сегодня появляются несколько путей к разрешению этой проблемы, причем все пути основаны на развитии компьютерных технологий.

#### **Тесты интеллекта, генерируемые компьютером**

Представим себе, что имеется несколько параллельных версий теста интеллекта. Другими словами, у нас есть несколько тестов интеллекта, построенных по единому принципу и измеряющих одно и то же свойство (например, общий интеллект), но реализуемых разными способами. В этом случае попытка испытуемого заранее подготовиться к тестированию будет существенно затруднена. Эта попытка станет невозможной, если психолог располагает бесконечным количеством вариантов. К сожалению, разработка даже двух параллельных вариантов теста представляет непростую проблему, поскольку для каждого варианта необходимо проведение трудоемких исследований, связанных с установлением эквивалентности заданий и т.д. Иметь хотя бы десяток вариантов – крайне дорогое и фактически недоступное удовольствие.

Однако есть и другой путь – попытка испытуемого подготовиться к тестированию невозможна также и в том случае, если до начала тестирования «теста» «не существу-

ет», поскольку он генерируется в момент, когда тестируемый только приступает к работе. В этом состоит идея тестов интеллекта, генерируемых компьютером. Представьте, что испытуемый садится перед компьютером, который предъявляет ему сгенерированную им задачу, причем компьютеру заранее известен уровень сложности этой задачи. В зависимости от ответов испытуемого он генерирует и предъявляет ему новые задачи, определяя доступную испытуемому степень трудности. Этим, с одной стороны, открывается возможность многоразового использования теста одним испытуемым и снимается вопрос о его защищенности. С другой стороны, уровень интеллекта определяется при помощи предъявления минимального числа заданий, поскольку их сложность корректируется в зависимости от работы испытуемого и отпадает необходимость предъявления лишних, слишком простых и слишком сложных, заданий.

Хотя в предшествующем изложении идея тестов, генерируемых компьютером, может показаться простой, для их создания потребовалась огромная работа. Вся проблема заключается в том, чтобы по структуре задачи заведомо определять степень ее трудности. Ведь компьютер, конструируя задачу, должен заранее, ни разу не предъявив ее испытуемому, знать, насколько она сложна и за какой промежуток времени ее должен решить (или вообще не решить) испытуемый с определенным уровнем интеллекта. Существующие на сегодняшний день способы создания таких тестов восходят к работе Карпентер и Джаста, которые на материале теста «Прогрессивные матрицы Равенна» вывели закономерности связей между структурой матричных заданий и их сложностью. В работах С. Эмбретсон были уточнены некоторые психометрические аспекты проблемы, в частности произведено соотношение с так называемой IRT-теорией тестирования, учитывающей сложность заданий. Лидером по созданию таких тестов на сегодня является группа немецких психологов во главе с Х. Холлингом.

Тесты, генерируемые компьютером – это замечательная (однако, к сожалению, не универсальная) возможность в плане создания объективных, нефальсифицируемых методов тестирования. Вряд ли возможно сконструировать этим путем личностные тесты, поскольку на сегодня не существует способа автоматического создания вопросов,

отражающих определенную личностную черту. Даже в сфере тестирования интеллекта применение принципов компьютерного генерирования достаточно ограничено. Сегодня эти принципы можно применять только в отношении тестов, сложность заданий которых определена их структурой, а не содержанием. Данное положение требует своего пояснения. Сложность заданий теста Равена определена их структурой, т.е. количеством элементов, образующих фигуры, правилами их появления и т.д. Сложность этих заданий не зависит (точнее сказать, зависит в минимальной, контролируемой степени) от того, какие именно элементы (линии, фигуры, комбинации и т.д.) применяются в задании. Однако, в случае использования теста на осведомленность ситуация оказывается прямо противоположной. Результативность выявления диагностируемого конструкта зависит от (во многом определяется) содержания. Сложность одного и того же по структуре задания (например, определить понятие) решающим образом зависит от содержания определяемого понятия. Очевидно, что более конкретное понятие «треугольник» определить намного проще, чем более абстрактное понятие, например, «демократия».

Компьютер может справляться с генерированием тестовых заданий в том случае, если их сложность определена их структурой. Для этого в коды компьютерных программ нужно заложить определенные правила, связывающие сложность и структуру тестовых заданий. Однако компьютерная программа самостоятельно не может генерировать задания, сложность которых определяется содержанием. Сложность содержания единичного задания нельзя установить априорно, а лишь путем проведения эмпирических исследований. Поэтому сегодня можно научить компьютер генерировать задания по типу равновесных, но не удастся добиться от него, например, векслеровских. А многошкальные тесты типа векслеровского имеют немало преимуществ, что, впрочем, составляет предмет для отдельного разговора.

#### **Психофизиологическое тестирование**

Психофизиологическое тестирование также может быть основано на внедрении компьютерных технологий. Так же как и генерируемые компьютером тесты, этот вид тестирования обладает столь же полной объективностью и возможностью противостоять фальсификации, но является в то же время

более универсальным. Основоположником компьютерного способа регистрации электрофизиологических данных и их соотношения с психологическими характеристиками является «живой классик» американской психофизиологии Е. Рой Джон. В наши дни персональный компьютер при помощи соответствующего программного обеспечения может не только провести Фурье-анализ электромагнитной активности мозга и построить на его основе спектрограммы, но и соотнести особенности электроэнцефалограммы человека с его психологическими особенностями. Например, Рой Джон установил в обширном исследовании такой любопытный факт: частота альфа-ритма человека положительно коррелирует с его интеллектом. Другими словами, чем выше у человека частота альфа-ритма, тем больше вероятность, что он продемонстрирует высокий интеллект.

В нашей стране в этой области на протяжении нескольких десятилетий работает А.Н. Лебедев. В его первых исследованиях в центре внимания был все тот же альфа-ритм, частоту которого он связывал с объемом рабочей памяти, предложив при этом объяснительную модель (Забродин, Лебедев, 1977; Лебедев, 2004). В более поздних работах А.Н. Лебедев стал привлекать большее количество характеристик электрофизиологической активности мозга, но теоретическое обеспечение этих исследований оставляет желать лучшего. В исследованиях используется множественный регрессионный анализ, в котором независимыми переменными выступают частотные и амплитудные характеристики ЭЭГ, а также синхронизированность отведений, а зависимыми переменными являются различные психологические характеристики. В результате для каждой психологической характеристики составляются так называемые «диагностические уравнения», которые связывают различные психологические характеристики с параметрами ЭЭГ. В то же время характер этой множественной связи остается неясным. Более того, интуитивно многим психологам и психофизиологам кажется странным, что и каким образом электромагнитные волны мозга (природа которых и сегодня еще во многом неизвестна) несут информацию о психологических свойствах.

Если связь частоты альфа-ритма с уровнем интеллекта еще получает какое-то объяснение, то объяснение механизм эффекта

синхронизации колебаний вызывает вопросы, на которые сегодня не существует ответов. Тем не менее, факт отсутствия достаточных научно-обоснованных объяснений не отменяет возможного практического значения психофизиологической диагностики. Интуиция исследователей не может быть критерием принятия или непринятия научного метода. Таким критерием могут быть только факты. Например, полученные Е. Рой Джоном факты о наличии корреляционных связей между частотой альфа-ритма и интеллекта, безусловно, свидетельствуют об отражении некоторых психологических свойств в электромагнитной активности мозга. Оценить, насколько это отражение может быть реально использовано в целях психодиагностики, — дело дальнейших исследований.

В контексте данной статьи важно отметить, что психофизиологическая диагностика является полностью защищенной от попыток каких-либо манипуляций со стороны испытуемого. Знакомство с этим видом диагностики, опыт ее предыдущего прохождения не помогут испытуемому что-либо изменить в своих результатах, поскольку все получаемые данные являются результатом работы энцефалографа и компьютера, обрабатывающий результаты. К тому же психофизиологическая диагностика не имеет тех ограничений по сфере применения, которые имеют генерируемые компьютером тесты. Она в принципе может использоваться как в плане психологического исследования интеллекта, так и личностных особенностей.

Получение психодиагностических данных в форме математических уравнений предполагает проведение большой предварительной работы. Использование в процедурах множественного регрессионного анализа более 300 параметров ЭЭГ выдвигает высокие требования к оценке статистической значимости связей. Это делает необходимым использование больших выборок испытуемых — в 400 и более человек. Объем такой работы представляется весьма значительным, учитывая, что речь идет не просто о психологическом исследовании, а об аппаратной регистрации электрофизиологических параметров.

### Диагностика по речи

Еще одним перспективным источником информации о тестируемом является его речь. Разумеется, речь производится человеком в большей или меньшей степени про-

извольно. Однако человек отнюдь не контролирует все, что проявляется в его речи. С середины 1950-х гг. в США Готтшалком и Глезер начала разрабатываться методика диагностики осознанных и неосознанных аффектов по речи. Использовались как спонтанные тексты испытуемых, так и тексты, порождаемые в соответствии с определенной инструкцией. Основным предметом в исследовании было проявление агрессии, хотя метод может использоваться и в отношении изучения иных эмоций. Оцениваются: частота появления определенных категорий в речи, степень интенсивности проявления аффекта и степень личного отношения? участия? испытуемого (Михайлова, 2006). Исходным методом анализа была работа экспертов по отнесению высказываний к тем или иным категориям, однако постепенно все большую силу набирают методы компьютерного анализа. Компьютер на основании составленного исследователями словаря категорий автоматически регистрирует присутствие тех или иных категорий в тексте и их взаимное расположение. Использование компьютера резко снижает трудоемкость проведения исследования и позволяет разрабатывать новые варианты ее практического использования.

Исходно методика была рассчитана на диагностику ситуативных эмоций, однако есть надежда применить ее в отношении исследования стабильных свойств личности. В 2005 г. в Институте психологии РАН Г.А. Малковой под руководством Н.А. Алмаева была защищена кандидатская диссертация, в которой изучалось проявление личностных характеристик испытуемых в текстах, сочиняемых ими на заданную тему. Пока что показатели взаимосвязанности объективных характеристик содержания речи с выявленными в процессе традиционной психодиагностики личностными особенностями, оказываются весьма скромными. В то же время сохраняется надежда, что путем совершенствования методов обработки, а также оптимизации процедуры получения текста удастся достичь больших результатов. Если это будет сделано, то возникнет новая ветвь психологической диагностики – психолингвистическая. Таким образом, это направление является перспективным, однако эта перспектива вполне реальна, дополняет непростую картину высокотехнологичных психодиагностических методов будущего.

### **Сравнительная характеристика новых подходов к тестированию**

Для сравнения перечисленных новых подходов к тестированию в плане их практического применения следует оценивать по нескольким критериям. К таким критериям можно отнести: широта применения, объективность (защищенность от фальсификации), простота проведения, доступность на рынке и, наконец, стоимость.

Достоинствами тестов, генерируемых компьютером, являются полная на сегодняшний день защита от фальсификации и простота проведения. Для проведения теста достаточно посадить испытуемого за компьютер, и через полчаса – час можно получить надежный результат. Однако, как уже отмечалось, область применения этих тестов, хотя и очень важная, но пока ограниченная – оценка общего интеллекта. Общий интеллект является часто самым важным предиктором профессиональной успешности человека (Ушаков, 2004), однако о многих аспектах поведения человека он не дает информации. Например, работодатель может интересоваться информацией о надежности, верности потенциального сотрудника. Тест общего интеллекта, к сожалению, здесь не скажет решающего слова.

Тесты, генерируемые компьютером, на нашем рынке еще не появились, хотя при большом желании доступ к ним получить уже можно. Стоимость еще не определена, но это продукт современных западных технологий, который при соблюдении прав интеллектуальной собственности достаточно дорог.

Психофизиологическая диагностика так же, как и генерируемые компьютером тесты, обладает полной защитой от фальсификации, но при этом имеет более широкую сферу приложения. Дополнительным преимуществом психофизиологической диагностики является возможность ее параллельного проведения с психологическим тестированием. В этом случае возникает дополнительный контроль – расхождение результатов служит сигналом к дополнительной проверке на предмет возможной необъективности результатов. Недостатком психофизиологического тестирования является сложность проведения. Хотя процедура и не очень длительна, она достаточно трудоемка. Тестируемому на голову нужно наложить электроды, что предполагает уже определенную психофизиологическую квалификацию, не

всегда доступную «чистому» психологу. Энцефалограф стоит на порядок дороже компьютера. Кроме того, все это железо не будет работать без специальных программ, составляющих «ноу-хау» разработчиков. В стандартный набор поставки энцефалографов эти программы естественно не входят. Как следствие, психофизиологическая диагностика оказывается существенно дороже психологической. Хотя она уже доступна на отечественном рынке, споры о ее валидности отнюдь не утихли.

Диагностика по речи пока что является наименее отработанной, однако перспективы ее вполне реальны. Она имеет в принципе широкую область применения и в значительной степени защищает от фальсификации.

В целом можно сказать, что психология постепенно подходит к этапу, когда перед высокотехнологичной диагностикой в разных ее формах открываются перспективы практического внедрения. Вероятно, в ближайшие десятилетия и в научных, и в практических целях будет происходить расширение сферы применения высокотехнологичных тестовых методов, хотя, по-видимому, место для традиционных опросников и по-

добных им методов будет сохраняться. При этом в конечном счете основу всего прогресса в этой сфере составляет развитие вычислительной техники.

#### **Литература**

1. Забродин, Ю.М. *Психофизиология и психофизика* / Ю.М. Забродин, А.Н. Лебедев. – М.: Наука, 1977.

2. Лебедев, А.Н. *Нейронный код* / А.Н. Лебедев // *Психология. Журнал Высшей школы экономики*. – 2004. – Т. 1. – №3. – С. 18–36.

3. Михайлова, Н.Б. *Речевая диагностика* / Н.Б. Михайлова // *Психолингвистика* / под ред. Т.Н. Ушаковой. – М.: ПЕР СЭ, 2006. – С. 322–378.

4. Ушаков, Д.В. *Тесты интеллекта, или горечь самопознания* / Д.В. Ушаков // *Психология. Журнал Высшей школы экономики*. – 2004. – Т. 1. – №2. – С. 76–93.

5. Шмелев, А.Г. *Тест как оружие* / А.Г. Шмелев // *Психология. Журнал Высшей школы экономики*. – 2004. – Т. 1. – №2. – С. 40–65.

*Поступила в редакцию 02.10.2008.*

**Ушаков Дмитрий Викторович.** Заведующий лабораторией Института психологии РАН, руководитель Центра исследования и развития одаренности МГППУ, Первый вице-президент Евроталанта, Европейского комитета по образованию одаренных детей и юношей при Совете Европы, доктор психологических наук: dv.usakov@gmail.com.

**Dmitriy V. Ushakov.** Head of the laboratory at the Institute of Psychology of RAS (Russian Academy of Science), head of the centre of endowments research and development at MGPPU, first Vice-President of Eurotalent, European Committee for the Education of Children and Adolescents of the Council of Europe, Doctor of Psychology: dv.usakov@gmail.com.