

Дисциплина

ОСНОВЫ ПСИХОДИАГНОСТИКИ

Лекция

ДОСТОВЕРНОСТЬ КАК КРИТЕРИЙ ОЦЕНКИ КАЧЕСТВА ТЕСТА

Показатели личностных опросников скорее всего отражают своеобразную комбинацию самообмана, попыток обмануть других и реалистического изображения себя... Вес каждого из этих компонентов варьируется в зависимости от того, кто конкретно и при каких обстоятельствах тестируется.

А. Анастаси

План лекции

- 1 Понятие о достоверности измерений в тестировании.
- 2 Непреднамеренные искажения данных тестирования.
- 3 Сознательные искажения данных тестирования и способы защиты от них.

Литература

Шмелев, А.Г. Основы психодиагностики / А.Г. Шмелев и колл. – М., Р/нД : Феникс, 1996. – 544 с.

Анастаси, А. Психологическое тестирование / А. Анастаси, С. Урбина. – СПб. : Питер, 2001. – 688 с.

Бодалев, А.А. Общая психодиагностика / А.А. Бодалев, В.В. Столин. – СПб. : Речь, 2000. – 440 с.

Для эффективного применения психодиагностических методик в практике психологической работы сегодня требуется резкое повышение психометрической культуры психологов. Как известно, основными критериями оценки качества тестов являются надежность и валидность. По-нашему мнению,

следует говорить также о достоверности измерений как относительно особом требовании к измерительным средствам, коими являются многие психодиагностические методы.

В отличие от обычных бытовых измерений, где объектом является неживой предмет, человек, подвергающийся психодиагностическому обследованию, проявляет субъектные свойства. Он может быть обеспокоен тем, какие результаты получаются, кто и как будет их использовать. Несмотря на хорошую внешнюю валидность, тест в любом случае порождает у испытуемого встречные гипотезы о цели обследования. Кроме того, в зависимости от степени доверительности, которую психолог смог с ним установить, тот так или иначе оценивает ситуацию тестирования либо как ситуацию клиента, либо как ситуацию экспертизы.

В этой связи испытуемый может бессознательно или даже сознательно выбирать определенную тактику выполнения теста, которая дает искажение результатов. В первом случае говорят о мотивационных искажениях, во втором – о фальсификации ответов. Если тест не защищен от подобных искажений, встает вопрос о достоверности результатов. Без достоверности не может идти речь о валидности, так как, получая не заслуживающие доверия оценки по тесту, мы не можем судить о том психическом свойстве, которое хотим измерить.

По А.Г. Шмелеву, достоверность – критерий качества теста, обеспечивающий защиту его результатов от непреднамеренных мотивационных искажений или сознательных фальсификаций (Шмелев, А.Г. Основы психодиагностики. – М., Р/нД : Феникс, 1996. – С. 482).

Остановимся на примерах бессознательных искажений и сознательных подлогов, которые возникают в процессе тестирования, и на тех приемах, которые используют разработчики тестов для защиты результатов. Вместе с тем определенные меры для выявления недостоверных самоотчетов могут принять сами психологи. Речь будет идти также об этих приемах.

Непреднамеренные искажения могут возникать в обширном континууме от ситуаций клиента (когда человек обращается за психологической помощью) до менее регламентированных ситуаций (когда человек работает с тестом в целях самопознания). Приведем некоторые из этих случаев.

Ситуация клиента (или консультативная ситуация). В ней испытуемый участвует в тестировании на добровольной основе и сам заинтересован получить рекомендации по результатам обследования (как, например, в профориентационной консультации). Психодиагност может использовать опросники, не опасаясь по большому счету фальсифицированных ответов. Однако и в этом случае могут возникнуть искажения. Например, если какие-либо изменения обстановки тестирования насторожили школьника (скажем, в

кабинет к школьному психологу зачем-то вошел директор или завуч), то диагностическая ситуация может моментально превратиться в ситуацию экспертизы. Вот почему в профессиональном руководстве к использованию теста психологам часто настоятельно советуют соблюдать стандартные условия его проведения. Может случиться и так, что клиент, обратившийся за помощью, будет склонен невольно подчеркивать свои трудности и проблемы, чтобы вызвать к себе повышенное внимание психолога. Естественно, что эта тенденция будет отражаться и при работе с тестом. Вот почему проведение личностных методик требует высокой образовательной квалификации и качественного опыта психолога, дающих ему возможность внимательно проследить за поведением испытуемого в процессе консультации и тестирования.

Эффект, обусловленный действием закона Йеркса-Додсона. Согласно этому закону (Годфруа Ж. Что такое психология: В 2-х т. Т.1. – М. : Мир, 1999. – С. 138), сформулированному еще в 19-м веке, эффективность тестовых показателей зависит от уровня мотивации испытуемого приблизительно так, как это показано на рисунке.

Из приведенного графика на рис. 1 видно, что по мере усиления активации испытуемого исполнительская точность повышается до определенного уровня. Но с какого-то момента может начаться ее ухудшение из-за усиления мотивации вплоть до дезорганизации деятельности. Таким образом, имеется оптимальный уровень мотивированности испытуемого на выполнение теста, при котором он покажет все, на что способен.

Недомотивированный испытуемый снизит результаты обследования. Особенно это проявится при выполнении интеллектуальных и проективных тестов. В результате низкой мотивации испытуемый окажется рассеянным, не

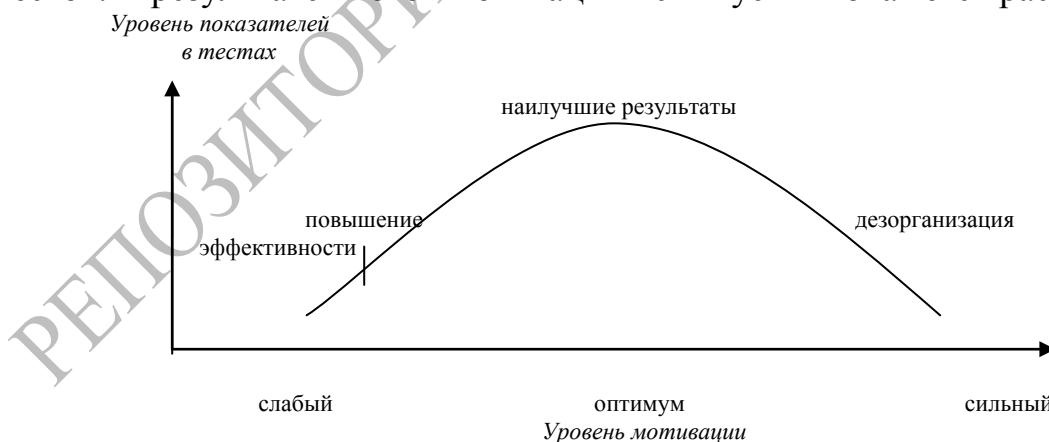


Рис. 1 - Иллюстрация закона Йеркса-Додсона, связывающего уровень активации организма с уровнем исполнительской точности любой деятельности, применительно к процессу тестирования

мобилизованным, он не будет активно искать те закономерности, которые позволяют успешно решить тестовое задание. При выполнении проективной методики такой испытуемый не выдаст обстоятельных результатов: не создаст оригинальный рисунок, не даст развернутой цепочки вербальных ассоциаций, не построит сюжетной линии по предложенному изображению. Например, в картинке теста тематической апперцепции GAT вместо связного сюжета он выдаст названия предметов и людей, которых видит. В проективный стимул, таким образом, он не привнесет от себя ничего, и эта бедная немногословная продукция не представит для психолога никакой диагностической ценности.

Сверхмотивированный испытуемый также снизит тестовые результаты. Избыточная активация приведет его в состояние тревоги. Ученик, который любой ценой хочет успешно пройти интеллектуальный тест, рискует потерять сосредоточенность, необходимую для понимания задаваемых вопросов. Особенно плохо чувствует себя испытуемый, уже приобретший опыт неудач и разочарований в предыдущих ситуациях тестирования. Для обозначения такого феномена в психодиагностике существует понятие тестовой тревожности (Анастаси, А. Психологическое тестирование / А. Анастаси, С. Урбина. – СПб. : Питер, 2001. – С. 37-38).

Тестовая тревожность включает в себя два важных компонента: эмоциональность и озабоченность. Эмоциональная составляющая охватывает чувства и физиологические реакции типа увеличения частоты сердечных сокращений. Содержанием озабоченности как когнитивной составляющей являются негативные мысли, связанные с ожиданием неудачи и ее последствий. Эти мысли отвлекают внимание испытуемого от заданий теста, что и приводит к искажению его результатов.

Очевидно, что два этих противоположных случая (слабая и сильная мотивации) должны быть сведены к мотивационному оптимуму. Чтобы повысить мотивацию испытуемого, необходимо вызвать у него интерес к предстоящей работе, придать ей увлекательный характер. Чтобы уменьшить напряжение, надо снять эмоциональный накал, с одной стороны, и снизить интеллектуальную озабоченность испытуемого в отношении самого себя, с другой.

Чтобы эффективно проводить тестовые обследования, следует гибко приспосабливаться к нуждам конкретных индивидов. Представим, что вам надо изучить интеллектуальные особенности ребенка. Сделать это возможно, либо задавая ему определенные вопросы, либо предлагая рисуночный тест. Решая задачи какого-либо теста, ребенок, как правило, осознает, что он поставлен в ситуацию экспертизы, и сталкивается с вероятностью неправильного решения. Все это создает повышенное нервно-психическое напряжение. Что касается

рисования, то оно носит для ребенка скорее игровой характер и не вызывает тревожных переживаний.

Однако неуверенный ребенок, опасаясь поколебать сложившуюся самооценку, может просто отказаться выполнить инструкцию, которая не кажется ему достаточно серьезной. Предлоги могут быть самыми незамысловатыми: «Я не умею» или «Больше не хочется рисовать». Но дело не в том, что ребенок не умеет или ему не хочется, а скорее в том, что намечающийся результат не вполне соответствует ожиданиям взрослого, как их представляет себе ребенок. Если попытаться заставить его выполнить рисунок, то он, вероятно, уступит нажиму. Но при этом его работа будет продиктована мотивом «отделаться» от требования взрослого. Диагностическая ценность полученного таким образом рисунка резко снижается. Рисунок может быть выполнен торопливо, неаккуратно, вследствие чего ряд необходимых для обсчета деталей окажется не проработанным или упущенным. В случае отказа надо постараться заинтересовать ребенка предстоящим делом. Важно ненавязчиво подчеркнуть, что вы не намерены оценивать творчество ребенка, сравнивать его рисунок с другими. Именно таким образом часто удается оптимизировать мотивацию. Если опасения ребенка смягчатся, он, скорее всего, уступит просьбе закончить работу. Если рисунок, несмотря на убеждения и просьбы, остался незавершенным, обработка такого материала нецелесообразна. В нем могут быть упущены многие детали, имеющие диагностическую ценность. А это исказит итоговую оценку. Когда ребенок бросает работу, явно не исчерпав возможностей рисунка, это свидетельствует лишь о его мотивации и настроении, но не об уровне интеллектуального развития. Если же вам рисунок кажется незавершенным, но ребенок не демонстрирует явного отказа, а, по всей видимости, считает работу законченной, нельзя требовать, чтобы изображение было дополнено. Как показывает практика, если поинтересоваться, почему отсутствуют те или иные детали, или тем более — посоветовать внести еще некоторые дополнения, ребенок охотно «усовершенствует» изображение. Однако это есть прямое искажение результата. Диагностическая ценность полученного таким образом показателя также сомнительна.

В некоторых случаях, особенно если есть параллельные формы теста, с помощью них недомотивированного испытуемого можно поупражнять, поощряя так или иначе его ответы. А. Анастаси указывает, что эмоциональный компонент тестовой тревожности имеет тенденцию убывать с каждым последующим тестированием (Анастаси, А. Психологическое тестирование / А. Анастаси, С. Урбина. – СПб. : Питер, 2001. – с. 38). Такой прием будет способствовать стимулированию интереса, оптимизации мотивации, и в целом поможет профилировать искажение результатов теста.

Феномен «кривого зеркала». Он является следствием проявления половозрастных особенностей. Например, мальчики-подростки, отвечая на вопросы теста, непроизвольно стремятся к проявлению своей независимости, демонстрируют подростковый нонконформизм. Возникает бессознательное мотивационное искажение – испытуемый пытается предстать хуже, чем он есть на самом деле, таким «крутым пацаном», которому «закон не писан». В повседневном поведении, вне ситуации тестирования, подросток может быть вполне благопристойным, воспитанным человеком, но тестирование провоцирует его на асоциальную выходку. Так же реагируют на ситуацию экспертизы неуравновешенные личности, для которых характерны асоциальные психопатические реакции. Очевидно, что во всех этих случаях очень важна степень доверительности, которую психолог может установить с испытуемыми в процессе обследования.

Феномен неадекватной самокритичности. Важным источником непреднамеренных искажений в результатах теста может выступить заниженная или завышенная самооценка испытуемого. Неадекватно оценивая свое поведение и личностные особенности, он, конечно же, будет ошибаться в выборе «правильного» ответа, то есть ответа, соответствующего собственной психологической реальности. Одни испытуемые будут проявлять в этой связи чрезмерную самокритичность, как бы «бичуя» себя за свои недостатки, другие – наоборот, сниженную самокритичность, считая себя полными достоинств, названия которых возможно отыскать в наборе вариантов ответов на вопросы теста. В том и другом случаях как комплекс неполноценности, так и комплекс превосходства приведут к неосознаваемому искажению результатов теста. Вот почему например, в опроснике 16 PF формы С предусмотрен расчет показателей по шкале MD – шкале самооценки. С ее помощью психолог получает дополнительную информацию о тенденциях развития самокритичности испытуемого, а, значит, может вынести заключение о том, стоит ли доверять результатам, полученным по основным шкалам.

Мотивация самопознания. Если у человека доминирует стремление к самопознанию и совершенствованию себя, он имеет определенные гипотезы о своей личности. В процессе обследования такой испытуемый будет невольно стараться подтвердить эти гипотезы с помощью теста.

Желание познать самого себя является важным качеством человека. Но мода на способы решения этой соблазнительной задачи меняется. Если подросток при этом обращается к развлекательным тестам, которые в изобилии печатают в журналах и газетах, то это еще ничего. К ним нужно относиться примерно с той же степенью серьезности, как кроссвордам и анекдотам. Но часто подросток обращается и к самостоятельному использованию такого мощного инструмента как профессиональный психологический тест. Подобно

скальпелю он хорош и полезен в руках хирурга, но страшноват и неуместен в руках играющего ребенка.

Никому не придет в голову продавать рентгеновские установки в магазинах игрушек. А вот диски с записью психологических тестов запросто продают. Зачастую это искаженные копии, безграмотные неадекватные переводы, самопальные интерпретации. И продают их рядом и вместе с компьютерными играми, страшилками, виртуалками. При таком использовании психологические тесты превращаются в игры детей с инструментами специалистов.

В этой связи практически единственной защитой является психопросвещение школьников, родителей и учителей в целях профилактики неконтролируемого применения психологических методик. Их надо убеждать в том, что к специалистам надо обращаться с самого начала. Заниматься самодиагностикой, также как самолечением, можно лишь в ограниченных пределах. Ведь когда у нас что-то происходит с телевизором, мы вызываем мастера. А человеческая психика посложнее любого технического устройства.

Что касается **тенденциозных** (сознательных) **искажений**, которые вносятся испытуемым в своих интересах, то суть и защита от них сводятся к следующим случаям.

Осведомленность испытуемого о тесте и процедуре тестирования. К великому сожалению психологов, многие профессиональные и классические методики в нашей стране в том или ином виде широко публикуются в популярной литературе и периодической печати, превращаясь в развлекательные средства. В некоторых случаях сами разработчики способствуют их неконтролируемому распространению, как это произошло, например, со знаменитыми тестами интеллекта Айзенка. Вот почему иногда можно получить индивидуальные результаты тестирования, далеко зашкаливающие за разумные пределы. При ближайшем рассмотрении оказывается, что испытуемый знаком с методикой или имеет опыт работы с аналогичными методиками. В некоторых случаях, как это делается за рубежом, в силу запроса со стороны клиентов, пытающихся успешно пройти тестовый профотбор, возникают фирмы типа «антитест». Эти фирмы организуют специальные тренинги, на которых учат претендентов на те или иные должности распознавать истинное назначение тестов, выбирать стратегию поведения во время тестирования и угадывать наиболее выгодные ответы. Такой испытуемый в процессе тестирования пытается выяснить, что за личностное свойство скрывается в опроснике за тем или иным конкретным поведением, описанным в суждении, как это свойство соотносится с существующими моральными нормами и профессиональными требованиями к претенденту на данную должность. Не исключено, что такой опыт подготовки к

тестированию будет перенесен и в нашу страну. К автору этих строк уже обращались клиенты с запросом на то, как успешно преодолеть тест в военно-командное училище МЧС или как подготовить ребенка к процедуре отбора в подготовительный класс престижной школы.

Очевидно, что разработчики тестов вынуждены придумывать специальные «ловушки», направленные на измерение достоверности. Эти ловушки входят как компонент в структуру теста, а иногда даже не подлежат разглашению как элемент «ноу-хау» (информационного изобретения) и профессиональной тайны, разделяемой конструктором теста с лицензированными пользователями, подписавшими особое соглашение при приобретении методики. Важным средством для сохранения достоверности результатов теста является также соблюдение специалистами требований этического кодекса поведения, который неизбежно возникает в организуемых сообществах психологов. При наличии корпоративной культуры профессиональная методика должна распространяться только среди лиц, имеющих сертификат на право ее самостоятельного использования.

Тенденция испытуемого давать о себе социально одобряемую информацию. Этот эффект опосредован, с одной стороны, действием фактора осведомленности, а с другой – стремлением испытуемого к высокой морализации «Я-образа для других» или доказательства своей психической «нормальности». Особенно актуализирует установку на фальсификацию психологической реальности сама ситуация тестирования. Происходит это тогда, когда испытуемый знает, что тестирование производится по инициативе администрации, педагога или родителей, и видит, что эти другие лица заинтересованы в результатах больше, чем он сам.

Чтобы нейтрализовать эффект социальной желательности или хотя бы его зарегистрировать, в профессионально разработанных опросниках используются так называемые «шкалы лжи». Шкалы лжи состояются из вопросов-ловушек (или утверждений-ловушек). Для успешного использования этой шкалы пункты-ловушки перемешиваются в стимульном материале с основными пунктами, отражающими измеряемые психические свойства. Они улавливают ответы испытуемого, явно predeterminedенные установкой на социальную одобряемость. Например, если он соглашается с высказываниями типа «я никогда не хвастаюсь», «я всегда выполняю требования родителей», «я никогда не нарушаю правил поведения в школе», «я всегда перехожу улицу только на зеленый свет», то по шкале лжи появляется высокий балл. В таком случае все результаты по тесту признаются недостоверными, и никаких суждений об испытуемом не выносятся.

Вместе с тем в работе со школьниками надо учитывать, что дети еще не вполне усвоили социальные нормы, и для них шкалы лжи работают плохо. Они

могут не совсем правильно воспринимать значение того, что от них требуется в методике. Речевое развитие детей до 12-13 лет не позволяет им однозначно понимать смысл инструкций и тестовых вопросов. Чем младше школьник, тем в меньшей степени для него пригодны опросники. Поэтому психологу при обследовании младших школьников и подростков не следует увлекаться вербальными методиками, основанными на речевом стимульном материале. Однако, учитывая те факты, что девочки перегоняют своих сверстников в психическом развитии и что они более конформны, чем мальчики, шкалы лжи для них вполне могут пригодиться.

Ситуативная тактика испытуемого. Работая с тестом, испытуемый может четко дифференцировать, каков он не только для себя самого, но и для других. Возникновение этого явления связано со сложным содержанием процессов самосознания. В них с точки зрения характера субъективных представлений могут выделяться оппозиции «Я-образ для себя» и «Я-образ для других». Искажения в результатах тестирования возникают тогда, когда испытуемый при ответах осознанно предпочитает тактику раскрытия своего масочного образа, тщательно скрывая реальное Я. Маска полностью доминирует в ситуациях высокого социального риска. Например, преступник при судебно-психологической экспертизе стремится прежде всего предстать больным, неприспособленным к жизни, слабым, хотя в действительности ему приятно думать о себе как о социально адаптированном и психически вполне здоровом человеке. Ответы такого испытуемого так же, как и в предыдущем случае, характеризуются социальной желательностью.

Психометрист с помощью специальных методов (см.: Бодалев, А.А. Общая психодиагностика / А.А. Бодалев, В.В. Столин. – СПб. : Речь, 2000. – С. 130) еще на этапе конструирования теста может оставить в нем только такие пункты (вопросы, утверждения), которые обеспечивают устранение социальной желательности. Делается это с помощью факторного анализа, когда психодиагност-исследователь задается целью обеспечить дискриминантную валидность основных шкал теста относительно «шкалы социальной желательности». Однако в исследованиях личности во многих случаях приходится диагностировать свойства, связанные как раз с социальной желательностью. Поэтому устранять из теста пункты, коррелирующие с социальной желательностью, не представляется возможным. К сожалению, формализованных приемов для регистрации всех возможных вариантов фальсификации данных при выборе испытуемым тактики «Я-образ для других» просто не существует. А.Г.Шмелев считает, что «диагностика свойств, сопряженных с социальной желательностью, в ситуации экспертизы всегда рискованна» (см.: Бодалев, А.А. Общая психодиагностика / А.А. Бодалев, В.В. Столин. – СПб. : Речь, 2000. – С. 130).

В процессе практической диагностики, чтобы убедиться в достоверности полученных результатов, психолог после первого тестирования может провести второе. Повторное обследование проводится с иной, «намеренно фальсифицирующей» инструкцией: «А теперь снова выполните тест, но старайтесь ответить так, чтобы выглядеть максимально благоприятно в глазах большинства других людей». Очевидно, что высокая корреляция между тестом и ретестом может свидетельствовать о серьезном искажении в оценке психологической реальности, которой на самом деле обладает испытуемый. Прием такой «репертуарной модификации» опросника может быть видоизменен предложением испытуемому выполнить тест не только от своего имени, но и от имени определенного репертуара ролей – «самый удачливый человек из тех, кого вы знаете», «самый счастливый человек», «самый нравственный человек» и т. п.

Конформистский стиль поведения испытуемого во время тестирования. Внешне испытуемый согласен протестироваться, но внутренне он выбирает либо тактику позиционного стиля ответа (или «да», или «нет», или нейтральный ответ) либо случайного стиля ответа (например, на клавиатуре компьютера, использующегося при опросе, клавиши нажимает случайным образом). Явно, что в этих случаях также возникают сильные искажения результатов тестирования. В целях выявления подобных отказов применяется ряд мер. Для исключения влияния «соглашательства» или «отрицания» применяются «прямые» (ответ «верно» в пользу выявляемого свойства) и «обратные» (ответ «неверно» в пользу выявляемого свойства) пункты. Если баланс подтверждающих и отвергающих ответов нарушается, то результаты признаются ненадежными. Для исключения тестов со случайными ответами в стимульный материал вводятся вопросы-дубли, представляющие собой синонимичные перифразы. Если испытуемый слишком часто по-разному отвечает на одинаковые вопросы, значит, он применяет случайную тактику. Например, современная версия опросника 16 PF дает три индекса стиля ответов, оценивающих уступчивость, выбор ответов наугад и попытки представить себя нереалистически, как обладающего социально желательными или, напротив, нежелательными качествами (Анастаси, А. Психологическое тестирование / А. Анастаси, С. Урбина. – СПб. : Питер, 2001. – С. 397).

Если не соблюдать указанные выше предостережения, тесты неизбежно будут давать систематические искажения результатов в сторону повышения или понижения баллов. Существует теоретическая возможность предусмотреть искажающие эффекты и заложить в конструкцию методики приемы обеспечения достоверности. Но на практике это оборачивается огромной эмпирико-методической работой психометриста, включающей возможности компьютерной обработки результатов. Вот почему психологи-практики,

критически понимая допущения, сделанные разработчиками тестов, должны обязательно дополнять результаты стандартизированных методик данными об испытуемых, поступающими из других источников. Только в этом случае можно обеспечивать достоверность индивидуальных тестовых показателей.

Лекция опубликована в виде статьи в журнале «Психологія»:

Лытко, А.А. Достоверность как критерий качества тестирования / А.А. Лытко. – Психологія. – 2004. – № 1. – С.27-33.

РЕПОЗИТОРИЙ ГГУ ИМЕНИ Ф.СКОРИНЫ