

**Глава 3. СОЗДАНИЕ НАДЕЖНЫХ ТЕСТОВ:
ЛИЧНОСТНЫЕ ОПРОСНИКИ. РАЗРАБОТКА ЗАДАНИЙ.**

Личностные опросники - широко используемые методики измерения личностных качеств, поскольку они могут разрабатываться так, что им будут присущи многие атрибуты эффективных тестов: надежность, дискриминативность и стандартизованность. Однако, в силу природы личностных переменных, труднее всего бывает установить валидность именно этих методик.

Проблемы конструирования личностных опросников

Как и в случае тестов способностей, качество отдельных заданий (вопросов, утверждений) является определяющим для всего опросника, и поэтому в данной главе будут рассматриваться трудности, возникающие при формулировании заданий, а также преимущества различных типов заданий.

При разработке заданий для личностных опросников необходимо учитывать следующие проблемы, которые, если их не обойти, неизбежно приведут к низкой валидности тестов.

- (1) Установка на согласие (response set of acquiescence). Это тенденция испытуемого соглашаться с утверждениями или отвечать на вопросы "да" независимо от их содержания. Чаще всего проявляется, согласно Guilford (1959), когда утверждения (вопросы) неоднозначны и неопределенны.
- (2) Установка на социально одобряемые ответы (response set of social desirability). Это тенденция испытуемых отвечать на вопросы теста так, чтобы выглядеть "социально положительным": если возможен "социально желательный" ответ, то весьма вероятно, что испытуемые будут его давать. Эта установка, как показал Edwards (1957), оказывает, в частности, влияние на результаты теста ММРІ (Minnesota Multiphasic Personality Inventory).
- (3) Установка на неопределенные или средние ответы (response set of using the uncertain or middle category). Если в опроснике представлена средняя категория ответов, отражающая нерешительность или неуверенность в ответе (напр., "не уверен", "не знаю", или "затрудняюсь ответить"), то многие испытуемые склонны к ней прибегать, как к безопасному компромиссу. Это приводит к снижению валидности заданий, поскольку большинство методов анализа вопросов основывается на крайних значениях показателей.
- (4) Установка на "крайние" (расположенные по краям шкалы) ответы (response set of using the extreme response). Эта установка может проявляться при использовании многоэлементной рейтинговой шкалы. Некоторые испытуемые, независимо от содержания вопросов, предпочитают выбирать крайние ответы (Vernon, 1964).
- (5) Очевидная валидность (face validity) вопросов (утверждений). Несомненно, в тестах личности необходима уверенность в том, что ответы на вопросы могут рассматриваться как правдивые. Cattell и Kline (1977) рассматривают данные опросника как данные Q и Qi. Данные из первого множества рассматриваются так, как если бы они отражали поведение испытуемого, данные из второго - как ответ на отдельный вопрос, ответ, который либо нагружен некоторым фактором, либо нет, независимо от того, отражает ли он поведение испытуемого.
- (6) Выборка из генеральной совокупности вопросов. В тестах интеллекта и специальных способностей, обсуждавшихся в предыдущей главе, относительно легко убедиться в том, что задания принадлежат или с высокой степенью вероятности выглядят принадлежащими к подразумеваемой для них генеральной совокупности. Лингвистические задания нелегко перепутать с математическими или какими-либо другими. Однако, в области особенностей личности и темперамента все гораздо сложнее.

В самом деле, иногда даже опытные и искусные разработчики вопросов (например, Cattell, 1957) бывают удивлены тем, что вопросы нагружены не теми факторами, для выявления которых они были созданы, а также тем, что вопросы вообще не нагружены ни одним из факторов. Эта проблема с точки зрения классической теории тестов состоит в трудности определения конкретной выборочной совокупности вопросов. В результате возникает проблема подбора адекватной выборки из выборочной совокупности вопросов или утверждений для теста, а без этого тест не может быть валидным.

(7) Выборка из генеральной совокупности испытуемых. Как указано выше, в личностных тестах труднее обеспечить адекватность выборки из выборочной совокупности, чем в тестах способностей. При использовании тестов способностей обычно известна вполне определенная категория лиц (популяция), для которой данный тест предназначен, и таким образом, по крайней мере в принципе, можно эффективно подбирать выборки. Однако, в личностных тестах, в отличие от тестов, разрабатываемых для испытуемых с той или иной патологией, в идеале нужны выборки из общей популяции (т.е. всего населения страны), в которой встречаются все возможные показатели. Такие выборки, как показано, должны быть большими, и их обычно трудно получить.

(8) Проблемы в установлении адекватного критерия валидности. Существует, как уже подробно обсуждалось в разделе о валидности, значительная трудность в нахождении адекватного критерия. Например, если мы пытаемся измерить авторитаризм, то будем вынуждены полагаться на рейтинги, поскольку нет никаких других внешних мерил (в отличие от, например, школьных экзаменов в случае диагностики способностей). Рейтинги - это неадекватный инструмент, а если бы они подходили, тест был бы не нужен. Аналогично, если существуют высокоэффективные тесты авторитаризма, которые могут быть использованы в качестве критерия валидности, вероятно, в новом тесте нет необходимости.

Таким образом, мы вынуждены ограничиваться исследованиями конструктивной валидности, обычно основывающимися на мультивариативном анализе разрабатываемого теста по сравнению с другими переменными, а также исследованиями специальных групп, имеющих по предположению контрольные показатели по рассматриваемой переменной.

Таковы основные сложности, возникающие при конструировании личностных опросников, и следует отчетливо их себе представлять, прежде чем вы приступите к формулированию вопросов или утверждений. Две последние проблемы, хотя они и являются определяющими, выходят на первый план уже после того, как подобраны задания теста.

Формулирование вопросов (утверждений) для личностных опросников

Практически нет необходимости особо говорить о том, что для тестов личности имеет большое значение то, чтобы ответы на вопросы могли быть объективно и с высокой надежностью оценены. Следовательно, мы должны сделать это возможным путем разработки формы вопросов. Ниже приведены формы вопросов, используемых в наиболее известных опросниках - тех, которые оказались наиболее эффективными на практике.

(1) Вопросы с ответом типа "да-нет". Такие вопросы использовались автором этой книги в разработанных им тестах личности, AI3Q (Kline, 1971), OIQ и OPQ (Kline and Storey, 1978). Их легко формулировать, они понятны испытуемым и ответы на них даются быстро. Такова форма вопросов личностных тестов Айзенка. Типичный дихотомический вопрос: "Любите ли вы загорать на пляже?"

(2) Вопросы с ответами типа "да -?- нет", "да - затрудняюсь ответить - нет". Это вариант вопросов типа "да-нет", описанных выше, с добавлением категории неопределенности, потому что некоторые испытуемые становятся раздражительными к неговорчивыми, если их заставляют отвечать либо "да", либо "нет" на вопросы, в ответах на которые они не очень уверены.

Трудность с такими вопросами состоит в том, что эта "средняя" категория очень привлекательна для испытуемых и редко бывает информативной. Bendig (1959) показал это на примере теста MPI (Maudsley Personality Inventory - ранней версии EPI, Eysenck Personality Inventory), в котором не делалось различие между дихотомической и трихотомической формой вопросов, и сделал заключение

о том, что дихотомические вопросы более предпочтительны, поскольку они заставляют преодолевать нежелание делать выбор. С моей точки зрения между вопросами типа "да-нет" и "да-?-нет" различие столь невелико, что вопрос о том, какую форму выберет разработчик теста, становится не более, чем делом личного предпочтения. Cattell и его коллега используют трихотомическую форму для некоторых из вопросов в своих факторных тестах личности.

(3) Альтернативные задания (с ответами типа "правда-ложь"). Эти задания состоят из утверждений (часто от первого лица), которые испытуемые должны отметить как верные или неверные для них. Примером альтернативного задания является: "Я ненавижу втискиваться в переполненный автобус". Это форма утверждений, используемая в MMPI. По существу, она ненамного отличается от дихотомических вопросов, хотя формулирование их как альтернативных утверждений, а не как дихотомических вопросов, может до некоторой степени повлиять на языковую представленность задания.

(4) Задания с ответами типа "нравится-не нравится" (одно слово или фраза). Это весьма оригинальная форма задания, которая в настоящее время используется у Grygier (1961) в динамическом тесте личности (DPI - Dynamic Personality Inventory), тесте, основанном на шкале личностных предпочтений Кроута (Kroout Personal Preference Scale) (Kroout and Tabin, 1954). Примерами таких заданий могут быть: (1) "фонарики"; (2) "воротники из бобрового меха"; (3) "бас-барабан". Испытуемые должны указать, нравится им предложенное или не нравится. Grygier и Grygier в своем руководстве по DPI утверждают, что эти задания являются сугубо проективными по своей природе, а DPI - это в сущности проективный опросник. Однако, так это или нет, такие задания не обязательно должны быть проективными, и может оказаться, что это чрезвычайно полезная, хотя и необычная, форма заданий.

(5) Задания с рейтинговыми шкалами. Эти задания состоят из предложений, к которым прилагаются рейтинговые шкалы. Cornrey (1970) - наиболее видный исследователь, использовавший эти задания - отдавал им предпочтение потому, что в них преодолеваются трудности, связанные с коррелирующими дихотомическими заданиями (см. обсуждение далее, в гл. 6), а также потому, что испытуемым они кажутся более разумными; нежели дихотомические задания, хотя и может проявиться установка на крайние ответы (см. выше). Cornrey использует две семибалльные шкалы, крайними значениями которых являются: "всегда-никогда" и "определенно да - определенно нет", в зависимости от формулировки задания. Типичным примером может быть: "Мне нравится бывать с друзьями в кафе: всегда, очень часто, часто, от случая к случаю, редко, очень редко, никогда". Очевидной проблемой для этой шкалы является различная интерпретация индивидуумами терминов частоты.

(6) Разнообразные трихотомические задания. В действительности это варианты трихотомической формы вопроса типа "да-нет" (тип 1 в данном списке). Cattell, например, в тесте 16PF использует эти задания, потому что они более подходят по смыслу в отношении некоторых утверждений, чем "да - затрудняюсь ответить - нет". Примерами трихотомий являются: "обычно - иногда - никогда", "верно - не знаю - неверно", "согласен - не уверен - не согласен". Эти варианты предоставляют большую гибкость при формулировании заданий, чем жесткие формы типа "да-нет" (дихотомическая) или альтернативные.

(7) Трихотомические задания с выбором. Это варианты типа (6), позволяющие вложить в лаконичную форму задания почти любую мысль. В них используются три завершающих предложение фразы, одну из которых испытуемый должен выбрать. Вот типичный пример: "Когда мне нечего делать, я могу: (а) позвонить другу, чтобы поболтать; (б) заняться разгадыванием трудного кроссворда; (в) пойти на джазовый концерт". Некоторые разработчики тестов использовали такие задания (например, Myers-Briggs, 1962) с двумя, тремя и более вариантами выбора. Они могут быть упомянуты как отдельная категория заданий.

(8) Задания с вынужденным выбором. В таких заданиях, как уже говорилось выше, испытуемые принуждаются к выбору того, какое из (обычно двух) утверждений более точно применимо к ним или более верно для них, хотя может быть и больше вариантов выбора. Однако, задания с вынужденным выбором включены в данный список как отдельная категория не потому, что они концептуально отличаются от трихотомических заданий с выбором, а потому, что Edwards (1959) разработал довольно много заданий с вынужденным выбором, которые были специально сконструированы в пары, уравновешенные по тенденции к социально одобряемым ответам, чем пытался устранить влияние этой установки.

(9) Другие формы. Основные типы заданий, описанные в пунктах (1) - (8), - это те, которые в основном используются в наиболее известных личностных опросниках. Все они, с точки зрения формы (в отличие от содержания), довольно легко формулируются, для них существует несколько правил, которые будут обсуждены, ниже, они лаконичны, наконец, просты в работе, а также для подсчетов.

Очевидно, возможны и другие типы заданий) и в Myers-Briggs Type Indicator (MBTI), например, находим тип заданий, которые более напоминают проективный тест, хотя они оцениваются объективно. Испытуемые должны указать, какое из пары слов более привлекательно для них. Вероятно, их лучше рассматривать как вариант заданий "нравится-не нравится" (тип 3 в данном списке).

Cattell и др. (1970) используют аналогичные задания в тесте MAT (Motivational Analysis Test - тест анализа мотивов), хотя можно было бы возразить, что это задания объективного теста, а не теста личности; разработка объективных тестов описана в главе 4.

Все рассматривавшиеся до сих пор виды заданий требуют от испытуемых самостоятельной работы. Испытуемые должны отвечать настолько правдиво, насколько возможно, хотя, как я уже указывал, есть такие методы конструирования тестов, в которых предполагается, что это не обязательно. Некоторые задания в них, хотя по форме и являются такими же, как в приведенных восьми пунктах, отличаются по сопровождающим их инструкциям. Так, в личностном опроснике Эдвардса (Edwards, 1967) испытуемые должны отвечать на вопросы заданий так, как будто они смотрят на себя со стороны (как они полагают, другие видят их). Это, как можно надеяться, устраняет, по крайней мере частично, тенденцию к социально положительным ответам.

Хотя есть и другие типы заданий, используемых в личностных опросниках, нет никакого сомнения, что среди приведенных восьми типов можно найти задания, хорошо апробированные в практике тестирования личности. Более того, фактически нет такого содержания, которое не могло бы быть эффективно сформулировано в виде задания одного из этих типов.

Правила для формулирования заданий

Итак, у нас есть перечень типов заданий, приведенных выше, и теперь следует обсудить некоторые приемы, принятые разработчиками тестов с целью преодолеть описанные выше проблемы.

Большинство из того, что будет сказано, очевидно и исходит из здравого смысла. Несмотря на это, изучение многих опубликованных тестов, а также тестов, используемых для внутреннего подбора кадров в крупных организациях, убедило автора в том, что все это должно быть сказано. Слишком часто разработчики тестов, ослепленные блестящей методикой анализа заданий, забывают о том определяющем факте, что тест не может быть лучше (но может быть хуже), чем его задания. Guilford (1959), личностные опросники которого были среди первых, разрабатывавшихся при помощи факторного анализа (правда, на шкалах, уже хорошо изученных за сорок лет исследований), делает несколько предположений, полезных для разработчиков заданий в тестах личности. По нашему опыту, эти правила показали себя как весьма эффективные.

(1) Устраняйте возможность проникновения испытуемых в суть того, что изучается при помощи данных заданий. Это делается не ради того, чтобы ввести их в заблуждение, а потому, что если испытуемые догадаются, что некоторое задание предназначено для измерения черты X, то ответы будут отражать их точку зрения по выраженности у себя этой черты, а не реальное положение дел. Представления же некоторых испытуемых о своей личности могут быть значительно искаженными. Как утверждает Guilford (1959), идеальным было бы оценивать испытуемого по чертам, о которых он не знает, задавая ему вопросы о том, что он знает хорошо.

(2) Формулируйте понятные, недвусмысленные вопросы (утверждения). Это важно для того, чтобы уменьшить погрешность, возникающую из-за неверного понимания вопросов. Высокая надежность зависит, в известной степени, от этого качества теста.

(3) Задания должны отражать конкретные, а не общие аспекты изучаемой области поведения. Так, задание вида: "Нравится ли вам спорт?" - является слишком общим: термин "спорт" - неопределенный, как и термин "нравится". Необходимо задать более конкретный вопрос: "Вы играете регулярно в какую-либо спортивную игру?" или "Вы регулярно следите за игрой вашей любимой спортивной

команды?", "Следите ли вы за результатами скачек?" На эти вопросы, ответы на которые вряд ли будут фальсифицированы или относительно которых реальное поведение вряд ли изменится, испытуемые будут отвечать одинаково, когда бы вы их ни тестировали.

(4) В каждом задании следует задавать только один вопрос или делать только одно утверждение. Рассмотрим пример: "Думаю, следует заставить черных и представителей других рас эмигрировать". Если это задание предназначено для измерения расизма, то оно будет плохим. Есть расисты (такие, как южно-африканские), которые проводят резкие различия между черными и людьми других рас. Другие считают всех, кто не являются англосаксами, не относящимися к людям. Следовательно, некоторые расисты могут положительно ответить на этот вопрос, тогда как другие - нет (только черных следует принуждать к эмиграции). Более злостные расисты могут быть готовы оставить черных и представителей других рас, но видеть их исключительно в качестве рабов. А, например, немцы, проживающие в Англии, опять-таки ответят "нет", исходя из своей интерпретации понятия "другие расы" (см. п. 2 выше). Это задание явно бесполезно: у него слишком неопределенная формулировка и одновременно задаются два вопроса. Следовало бы дать утверждение более точно: "Думаю, следует заставить негров эмигрировать". Сейчас оно более понятно (выражение "черные и представители других рас" заменено на более точное), и в нем остался только один вопрос.

(5) Избегайте, насколько это возможно, слов, определяющих частоту действий. Они обычно настолько субъективны, что вносят большую неопределенность. Это положение проясняют примеры.

Пример А. "Часто ли вам снятся сны?" Здесь все зависит от значения слова "часто". Некоторые испытуемые могут испытывать такое чувство, что видеть сон раз в месяц - это часто, и ответят на вопрос утвердительно. Другие могут возразить, что один сон за ночь - это не часто, потому что исследования показали, что людям снятся три-четыре сна за ночь, и ответят "нет". Вопрос, таким образом, становится бессмысленным. Ответы указывают противоположное тому, что происходит на самом деле. Улучшенный вопрос будет звучать: "Снятся ли вам сны дважды в неделю или чаще?"

Пример Б. "Испытываете ли вы иногда беспричинное беспокойство?" Очевидно, что этот вопрос порождает ту же проблему, что и предыдущий. Однако, он еще хуже, поскольку "иногда" буквально означает "чаще, чем однажды". Следовательно, практически все испытуемые вынуждены, если они отвечают искренне, ответить на этот вопрос положительно. Однако, этот вопрос мог бы быть очень полезным, поскольку всем известно, что люди действительно различаются по частоте беспричинного беспокойства. Дело не в содержании, а в форме вопроса, которая не выдерживает критики. Как же можно более точно сформулировать вопрос, в котором частота столь существенна? В улучшенном виде это может выглядеть: "Испытывали ли вы беспокойство без особых причин в последние две (четыре) недели?"

(6) Насколько возможно, избегайте терминов, выражающих чувства. Вместо этого, попытайтесь представить задание в контексте поведения. Это правило было проиллюстрировано в пункте (3) выше, в котором задание было сформулировано более конкретно с целью повышения надежности при тестировании по истечении некоторого времени. Для этого слово "нравится" было заменено на "играете". Важным для нас здесь является то, что испытуемый без колебаний отвечает, играет он или нет, а со словом "нравится" возникает целая проблема. Сообразительные, высокообразованные, точно выражающие свои мысли испытуемые могут долго колебаться в выборе значения этого слова: "Да, им нравится это, но, вероятно, "наслаждаться" будет чуточку точнее; "находить удовольствие" - да, но...", -и т.д. Конечно, и для такого задания, как в вышестоящем пункте (5), в котором исследуется чувство беспокойства, эта проблема также существует. Однако, выражать это задание в поведенческой форме, например: "Бьется ли ваше сердце быстро и пересыхает ли во рту безо всяких причин?", кажется слишком искусственным. Особенность состоит в следующем. Если только задание может быть сформулировано так, чтобы неопределенные слова о чувствах могли быть заменены на описание поведения, это должно быть сделано. Когда это невозможно, конечно же, стоит попытаться составить задания, в которых описываются чувства. Их пригодность или непригодность будет выявлена на этапе анализа заданий.

(7) При помощи инструкций обеспечьте, чтобы испытуемые давали первые приходящие на ум ответы. Не позволяйте испытуемым долго размышлять над значением заданий. Эффективное задание личностного теста, которое действительно имеет отношение к поведению испытуемого, должно вызывать немедленную и точную реакцию. В противном случае вероятно, что задание не выявляет ту

область поведения, которая имеет существенное отношение к измерению личностных качеств. Некоторые примеры прояснят этот момент.

Пример А. "Нравится ли вам наблюдать за щеглами?" Этот вопрос у испытуемых, за исключением орнитологов, не вызовет, вероятно, ничего, кроме смеха над его очевидной эксцентричностью. Большинство людей это просто не интересует.

Пример Б. "Любите ли вы хлеб с маслом?" Это явно не эксцентричный вопрос. Многие испытуемые, по крайней мере на Западе, где нет недостатка в основных продуктах питания, не будут испытывать по этому поводу сколько-нибудь сильных чувств, воспринимая это как данность.

Пример В (взято из пробных заданий для разработанного автором теста А130; Kline, 1971). "Позволяете ли вы вашей собаке лизать вам лицо?" Это задание было предназначено для выявления чистоплотности, и ключевым ответом для него является "нет". Очевидно, те люди, у которых нет собаки, ответят "нет", но можно возразить, что часто чистоплотность не является причиной отсутствия домашних животных. К моему удивлению, этот вопрос вызвал исключительно значимые ответы. Некоторые индивидуумы, дойдя до этого вопроса, отказывались вообще отвечать на вопросы теста, говоря, что все это вообще отвратительно; один испытуемый даже говорил мне, что он никогда не был так оскорблен, и никогда бы не допустил подобной мерзости, и тому подобное в том же духе.

Нет никакого сомнения, что этот вопрос затрагивает нечто скрытое и подавляемое. Уместность вышеупомянутой инструкции к тесту теперь понятна. Именно первый ответ, по всей вероятности, является индикатором определенного личностного качества. Как только испытуемые начинают раздумывать над вопросами, включаются не только защитные процессы, но и сознательные искажения, такие как желание произвести впечатление, угодить экспериментатору, - все это отрицательно влияет на результаты теста.

И вот что еще более важно: многие задания не могут выдержать критической оценки в основном из-за вполне оправданного желания исследователей вложить слишком много информации в краткую форму задания. Вот почему особое значение при создании заданий тестов придается понятности и конкретности.

Можно было бы поразмышлять над приведенным выше примером В, и тогда не было бы никаких проблем. Некто либо позволяет, либо не позволяет собаке лизать его лицо. Нет никаких сложностей в решении того, что: (а) собака есть собака; (б) что она именно лижет; (в) и что именно она лижет - лицо или что-нибудь другое. Однако, большинство тестов личности не являются такими однозначными. Между прочим, стоит отметить, что этот вопрос оказался неудачным и его пришлось удалить.

Пример Г. "Бывают ли у вас яркие сны?" Этот пример хорошо иллюстрирует наше указание. Большинство испытуемых, вероятно, смогут дать немедленный ответ - либо "да", либо "нет". Однако, при внимательном рассмотрении этот вопрос оказывается неприемлемым. "В конце концов, насколько яркими могут быть сны? Мои сны кажутся мне яркими, но это мое субъективное мнение. Я действительно вижу яркие сны, но я также иногда вижу несколько однообразные и серые сны," - подобные мысли затрудняют ответы.

Пример Д. "У вас много друзей?" Как и в первом примере, на этот вопрос, по размышлению, чрезвычайно трудно дать точный ответ. Проблема опять заключается в значении определенных слов, в данном случае "много" и "друзей". С этими трудностями столкнутся внимательные, вдумчивые и добросовестные испытуемые. Они могут рассуждать так: "Думаю, у меня много друзей, но что значит "много"? Сколько на самом деле в среднем друзей имеют люди? Без этой информации, а также без знания о разбросе вокруг среднего значения, невозможно ответить, много их у меня, или нет."

Хотя все эти мысли несомненно правильны, на практике этот вопрос у многих испытуемых вызывает автоматический ответ. Слово "друзья" тоже может породить проблемы. Испытуемые могут рассуждать: "Что такое друг, чем он отличается от знакомого, есть ли вообще какие-либо различия? Возьмем А, является ли он мне другом? Некоторым образом, нам нравится бывать вместе, но если бы он мне был нужен, он был бы практически бесполезен. Теперь Б...", - и так далее.

Из этих примеров понятно, что в инструкциях необходимо просить испытуемых выполнять задания настолько быстро, насколько это возможно. Конечно, нельзя избежать того, что таким инструкциям не всегда следуют, и в исследовании навязчивости, выполненном автором книги, один из вопросов, касающийся того, быстро ли испытуемый решается на что-либо и придерживается этого решения, ясно показал возможность многих различных ответов, хотя на него и отвечали утвердительно "да".

Это последнее наблюдение подводит нас к той критике, которая направлена против личностных опросников, что мы сейчас и обсудим.

Возражение: задания личностных опросников по существу бессмысленны.

Это становится очевидным из приведенных выше примеров Г и Д. Рассчитывать на то, что испытуемые-интеллектуалы будут отвечать быстро, - это не только оскорбительно для испытуемых (так утверждает Alice Heim в связи с обсуждением некоторых тестов интересов; Heim and Watts, 1966), но более того, у таких испытуемых появляются серьезные сомнения в том, что такие задания могут вообще что-либо измерять.

На эти критические замечания получен эмпирический ответ: анализ заданий любого рода показывает, что наше множество заданий измеряет однородный фактор, в то время как исследования валидности демонстрируют, что же это за фактор. Тот факт, что задания имеют сомнительное значение, становится несущественным, если ответы на эти задания фактически являются эффективно дискриминативными или коррелируют с внешним критерием. Вышеприведенная критика, хотя и может показаться серьезной, не имеет, таким образом, большого значения.

Таковы правила формулирования заданий, описанные Guilford (1959), и они способствуют обеспечению того, что подготовленные вопросы будут обеспечивать решение тех задач, которые стоят перед ними. Автору этой книги при конструировании собственных заданий для личностных тестов всегда удавалось создать надежные и, как я полагаю, валидные методики, когда он следовал этим правилам. Удалось ли бы ему это без уже известной нам методологии, неизвестно.

Устранение влияния установок на ответ

Кроме того, что при разработке заданий следует всегда помнить о вышеперечисленных правилах, необходимо, конечно, ослабить, насколько это возможно, влияние наиболее существенных установок на ответ. Ниже описано, как это следует делать.

Установка на согласие (acquiescence)

Установки на ответы (response sets) были определены Кронбахом (Cronbach, 1946) как стилистическая последовательность, стимулируемая формой ответов на задания личностного опросника. Как указывалось ранее, одной из наиболее существенных является установка на согласие (см. Messick, 1962), то есть тенденция соглашаться с утверждением независимо от его содержания.

СБАЛАНСИРОВАННЫЕ ШКАЛЫ

Messick (1962) утверждает, что сбалансированная шкала - это один из способов минимизации влияния установки на согласие. Сбалансированной называется такая шкала, для которой равное или примерно равное количество заданий имеют ключевые ответы "да" или "нет", "верно" или "неверно".

Однако, здесь заслуживают внимания два момента. Во-первых, сбалансированная шкала не устраняет полностью тенденцию : установка на согласие все же может проявиться, но, как указывает Knowles (1963), такая шкала не позволит спутать "соглашающегося" испытуемого с тем, у кого реально высокие показатели, а это важно. Конечно, можно возразить, что "соглашающийся" испытуемый с потенциально высокими показателями не будет выявлен при сбалансированной шкале. Если

это происходит чаще, чем в нескольких отдельных случаях, то последующее исследование валидности теста будет обречено на неудачу.

Во-вторых, искажающего влияния этой установки на показатели по тесту можно избежать только в том случае, если будут сформулированы в равной степени содержательные, недвусмысленные и побуждающие к определенному ответу задания, имеющие как положительные, так и отрицательные ключевые ответы. Если вы сможете сформулировать лишь очень небольшое количество заданий с ключевым ответом "нет", то эти задания будут способствовать низкой валидности шкалы. Несколько примеров пояснят это.

Пример 1. Рассмотрим вопрос на экстравертированность: "Вам нравятся вечеринки?" Обратной формой этого вопроса будет: "Вам не нравятся вечеринки?" В этих заданиях стилистические обороты языка удовлетворительны, потому что принято говорить, что вечеринки нравятся или не нравятся. Они также удовлетворительны, потому что, как это и случается, испытуемому с высоким показателем по параметру экстраверсии нравятся вечеринки, а испытуемому с низким показателем - не нравятся. Таким образом, полярность этих заданий эффективна. Однако, часто попытка получить обратную форму вопроса становится источником ошибок.

Пример 2. Следующее задание было создано, чтобы выявить удовлетворение, получаемое от статистики - как утверждают, оно относится к чертам навязчивости: "Являетесь ли вы одним из тех людей, кто находит статистические данные и таблицы совершенно скучными?" Это обратная форма вопроса с ключевым ответом "нет". Однако, очевидно, (прежде чем первое испытание теста показало, что этот вопрос не работает), можно не считать статистику скучной, но и не восхищаться ею. Таким образом, обращение вопросов, когда они относятся к некоторым крайним проявлениям в поведении, вероятнее всего не даст результата в силу континуальности, лежащей в основе многих ответов на задания. Обращение вопросов обычно предполагает дихотомичность ответа. Другими словами, если ответ на вопрос является дихотомическим, то обращение вопроса для сбалансированной шкалы возможно; если же используются крайние (граничные) значения из континуума, обращение не будет удачным.

Пример 3. Аналогично, можем рассмотреть вопрос о том, нравится ли испытуемым отдавать распоряжения. Обратная форма вопроса, такая как: "Вы не испытываете удовольствия, когда вам вменяется в обязанность отдавать распоряжения?" - не может быть удачной, потому что возможны случаи, когда отрицание неприязни к распоряжениям, приказам вовсе не означает получение удовольствия от этого. В основе этого ответа лежит континуум, а не дихотомия. Этот вопрос не может быть обращен.

Пример 4. В данном задании рассматривается аккуратность: "В школах сильно преувеличивают значение аккуратности". Можно утверждать, что навязчиво аккуратный индивидуум будет считать, что значение этой особенности поведения невозможно преувеличить. Следовательно, данное задание было написано так, чтобы отрицательный ответ действительно выявлял исследуемую черту. Таким образом, с точки зрения разработки заданий, это жизнеспособное "негативное" задание.

В этих примерах проиллюстрированы два наиболее важных момента в формулировании заданий с отрицательными ключевыми ответами: они могут быть сформулированы, во-первых, если отрицательный ответ соответствует реальному поведению, и во-вторых, если в основе ответов лежит дихотомия, а не континуум, как в приведенном выше примере 1.

ПОНЯТНЫЕ, НЕДВУСМЫСЛЕННЫЕ УТВЕРЖДЕНИЯ

Имеющий значение в контексте нашей работы вопрос был поднят Guilford (1959) при попытке устранить установку на согласие из личностных опросников. Он утверждает, что согласие наименее вероятно тогда, когда задания понятны, недвусмысленны и относятся к конкретному поведению. Поскольку эти качества уже были указаны среди тех, которые, по моему предположению, должны быть присущи разрабатываемым заданиям, я не буду более говорить о них, а лишь проиллюстрирую это положение несколькими примерами.

Так, вопрос: "Играете ли вы на каком-либо музыкальном инструменте?", - является настолько конкретным, что индивидуум должен был бы иметь предельно выраженную установку на согласие, чтобы утвердительно ответить на этот вопрос, если это не соответствует истине. С другой стороны, вопрос: "Нравится ли вам музыка?", - настолько неопределенный (в том, где же критерий того, что музыка нравится?), что согласие, вероятно, будет фактором, влияющим на положительные ответы в этом вопросе. В последнем примере есть две особенности. Прежде всего, обратите внимание, что данный вопрос несравним с вопросом: "Нравятся ли вам вечеринки?" В этом вопросе критерий того, нравятся ли вечеринки, известен каждому: частые посещения с желанием и энтузиазмом. Следовательно, этот вопрос - практически сокращенная форма вопроса: "Посещаете ли вы много вечеринок?" Таким образом, для него вероятность подвергнуться влиянию установки на согласие меньше, чем для вопроса: "Любите ли вы музыку?"

К этой особенности относится и неопределенное утверждение "нравится музыка". Это выражение бесполезно в задании. Так, слово "нравится" может соответствовать разнообразным чувствам: от тех, которые испытывают великие композиторы и исполнители, когда слушают, пишут или исполняют музыку, до чувств человека, который любит слушать популярные песенки, совершая субботние покупки. Подобно этому, термин "музыка" означает огромное разнообразие различных направлений: например, джаз, поп, народная музыка, рок, барокко, доклассическая, классика, романтизм, модерн, неоклассицизм, - а, кроме того, может иметь отношение к слушанию и исполнению, так что идентичные ответы на вопрос могут отражать совершенно различное поведение. Например, "да" для испытуемого 1 может отражать наслаждение от пения контр-тенора в ранней полифонической венецианской церковной музыке. А для испытуемого 2 это может отражать удовольствие прослушивания мотивчика типа "Танцуй со мной", когда он сидит за рулем своей машины. Любое задание, в результате оценки которого оказывается, что столь разные области поведения будут отнесены к одному и тому же классу, с очевидностью является бесполезным.

ПРОВЕРКА ВЛИЯНИЯ УСТАНОВКИ НА СОГЛАСИЕ

До сих пор все наши усилия были направлены - при помощи сбалансированных шкал и формулирования предельно ясных и конкретных заданий - на уменьшение влияния установки на согласие или на то, чтобы сделать менее вероятным получение за счет тенденции к согласию высоких показателей по переменной, для измерения которой предназначен тест. Однако, как уже говорилось, действие этой установки возможно даже при сбалансированных шкалах, и далеко не всегда можно создать настолько понятные вопросы, чтобы быть уверенным, что влияние установки устранено. Поэтому разработаны методики, предназначенные для проверки того, была ли установка на согласие фактором, влияющим на ответы. Эти методики будут описаны в главе 6.

Социально одобряемые ответы

Edwards (1957) показал на материале с заданиями ММРІ, что наблюдается высокая положительная корреляция между степенью социальной желательности утверждений, оценивавшейся экспертами, и количеством реальных социально одобряемых ответов на них испытуемыми. Поэтому он утверждает, что данная установка должна детерминировать ответы на задания и, следовательно, является основным источником невалидности тестов. Хотя, очевидно, невозможно полностью устранить влияние установки на социально одобряемые ответы, есть разнообразные приемы, полезные для уменьшения ее влияния. Эти приемы описаны ниже.

ЗАДАНИЯ С ВЫНУЖДЕННЫМ ВЫБОРОМ, ПОДОБРАННЫЕ ПО ИХ СОЦИАЛЬНОЙ ЖЕЛАТЕЛЬНОСТИ

В списке личностных предпочтений (Personal Preference Schedule) Эдвардса (EPSS) (Edwards, 1959), как мы уже отмечали в обсуждении типов заданий (см. п. 8 на стр.95), задания состоят из двух утверждений, подобранных по социальной желательности, одно из которых испытуемый должен выбрать. Это радикальное решение для устранения влияния установки на социально одобряемые ответы. Однако, этот метод не получил поддержки по следующим причинам.

(1) Крайне трудно получить согласованные утверждения, релевантные также в терминах содержания, которое мы хотим измерить (Edwards, 1957). Действительно, серьезное возражение против EPPS - это то, что данный тест является немногим более, чем упражнением по конструированию тестов, и мало свидетельств того, что он измеряет переменные с какой-либо эффективностью (см. Kline, 1979).

(2) Любые незначительные различия в социальной желательности между частями заданий имеют тенденцию увеличиваться, когда они предъявляются вместе, как это предусмотрено в форме вынужденного выбора, что сводит на нет большинство из усилий по подбору утверждений.

(3) Экспертное оценивание социальной желательности - значительное упрощение данного явления. Это становится ясно из того, как получается средняя экспертная оценка социальной желательности для каждого задания. Для этого предполагается, что социальная желательность одномерна (хотя а priori это не так). В действительности было бы нужно многомерное шкалирование влияния социальной желательности по всем заданиям, и затем подсчет показателей по отдельным измерениям - процедура, которая могла бы быть, несомненно, рекомендована в свете приведенного выше пункта (2)! Кроме того, как указывает Messick (1960), социальная желательность сама по себе допускает значительные индивидуальные различия: то, что социально желательно для члена парламента, почти наверное не является таковым для лондонского докера.

Исходя из этих трех причин, не рекомендуется прибегать к такому радикальному и, по-видимому, недостаточно эффективному шагу, как конструирование спаренных заданий, подобранных по их социальной желательности.

ИЗБЕГАЙТЕ ЯВНО СОЦИАЛЬНО ЖЕЛАТЕЛЬНЫХ / НЕЖЕЛАТЕЛЬНЫХ ЗАДАНИЙ

Существует ряд характеристик и качеств, с которыми мало кто захочет согласиться, по крайней мере в среде образованных европейцев. Они настолько очевидны, что когда мы проиллюстрируем их ниже несколькими примерами утверждений, то они покажутся явно абсурдными. Вот несколько примеров утверждений, которые являются явно социально желательными или нежелательными:

- (1) Я с трудом переношу неудачи: (да, нет).
- (2) Я не понимаю юмор: (да, нет).
- (3) Я постоянно лгу: (да, нет).
- (4) Я мало сексуален: (да, нет).
- (5) Я сексуальный психопат: (да, нет).
- (6) Я не могу контролировать свои эмоции: (да, нет).
- (7) Я завистливый и ревнивый человек: (да, нет).
- (8) Я скупой: (да, нет).
- (9) Когда возможно, я отлыниваю от работы: (да, нет).
- (10) Я лгу, чтобы избавиться от хлопот: (да, нет).
- (11) Я ненавижу негров: (да, нет).
- (12) Я в основе своей антисемит: (да, нет).

По нашему опыту разработки тестов, мы были бы сильно удивлены, если бы любой из двенадцати приведенных примеров оказался приемлемым даже для исследовательского теста. Например, для тестов личности, используемых в профотборе, они будут совершенно бесполезными. Представьте

себе утверждение (11) в тесте, предназначенном для работников социального обеспечения, или утверждение (12) в тесте, предложенном человеку, надеющемуся получить работу в фирме, владельцем которой является еврей.

ИЗМЕРЕНИЕ СОЦИАЛЬНО ЖЕЛАТЕЛЬНЫХ ИЛИ НЕЖЕЛАТЕЛЬНЫХ ЧЕРТ

Если измерению подлежат социально желательные или нежелательные черты, при разработке заданий следует избегать прямолинейного подхода, как в перечисленных выше утверждениях. Два примера пояснят сказанное.

Пример 1. Задание для измерения скупости. Поскольку (см. утверждение 8 выше) прямой подход не годится, я утверждал при разработке теста АИЗQ, что скупой человек вполне мог бы думать, что содержащиеся в народной мудрости высказывания о бережливости очень разумны, тогда как на менее скупого они не произвели бы впечатления. Следовательно, я попытался сформулировать задание следующим образом: "Каждый ребенок должен знать, что бережливость лучше богатства". Похоже, что в этом задании ответы как "да", так и "нет", не имеют отношения к социальной желательности (не являются социально одобряемыми). Действительно, этот вопрос оказался удачным по всем результатам анализа заданий.

Пример 2. Задание для выявления мстительности. Мало кто подтвердит утверждение: "Я мстителен". Однако, как уже говорилось, мстительные индивидуумы обычно проецируют свою мстительность на других, т.е. действуют механизмы защиты. Таким образом, я сконструировал утверждение: "Поступками большинства революционеров руководит мстительная жестокость". Я подумал, что мало кого из испытуемых можно было бы действительно причислить к революционерам, так что мнения обследуемых будут отражать их собственные защитные реакции и стремления. Это задание оказалось удачным, и я полагаю, что проекция - это механизм, действие которого может быть использовано в формулировании вопросов, которые слишком социально нежелательны, чтобы задавать их прямо.

Пример 3. Задание для выявления лени. Если опираться на механизм проекции, то можно сформулировать утверждение так: "Промышленность сегодня находится в плачевном состоянии, потому что рабочие в основном ленивы". Следует заметить, что если это утверждение рассматривается некоторыми испытуемыми как факт, то есть не "включает" проективный механизм, то анализ заданий полностью выявит это: вопрос не будет работать.

ИСПОЛЬЗОВАНИЕ ШКАЛЫ ЛЖИ

Некоторые авторы (например, Айзенк в тестах ЕРІ и ЕРQ) вводят особую группу заданий для выявления тех лиц, которые склонны давать социально одобряемые ответы. Они состоят из вопросов, касающихся незначительных проступков, которые, увы, иногда совершаются большинством людей. Набранный наибольшее количество баллов по такой шкале рассматривается как дающий социально одобряемые ответы, а поэтому его показатели по тесту могут быть проигнорированы. Хотя Eysenck и Eysenck (1976), исходя из факторного анализа заданий теста ЕРQ утверждают, что шкала лжи измеряет определенную личностную переменную, для практического тестирования со значительным потоком испытуемых этот прием для обнаружения установки на социально одобряемые ответы весьма полезен.

Вот некоторые типичные утверждения шкалы лжи: "Я никогда не лгу"; "Я всегда прихожу вовремя на встречи и свидания"; "Я всегда плачу за проезд на транспорте".

АДЕКВАТНЫЙ АНАЛИЗ ЗАДАНИЙ И ВАЛИДИЗАЦИЯ ТЕСТА

Несмотря на все сказанное, я по-прежнему утверждал бы, что при соответствующем анализе и отборе заданий во время разработки теста, при тщательной валидации теста, как и в случае с установкой на согласие, влиянием тенденции к социально одобряемым ответам можно пренебречь.

(1) Анализ заданий. По определению, наличие в тесте задания, провоцирующего социально одобряемые реакции, должно приводить к тому, что распределение ответов на него будет смещено, то есть не будет соответствовать нормальному распределению. Таким образом, устранением из теста заданий с распределением ответов, отличным от нормального, мы избавимся от действия установки на социально одобряемые ответы. Более того, если большинство заданий дают нормальное распределение ответов и они нагружены некоторым общим фактором, тогда установка на социально одобряемые ответы не может оказывать сильное влияние и на другие задания, относящиеся к тому же фактору. Подобные аргументы приводятся при анализе заданий с использованием бисериальной корреляции результатов выполнения каждого задания с общим показателем по тесту. Таким образом, при адекватном анализе те задания, результат выполнения которых подвержен влиянию установки на социально одобряемые ответы, должны быть устранены, если только, по нелепой случайности, все задания, выбранные нами для теста, не оказались измеряющими эту черту. Возможность возникновения этой ситуации должна проверяться при валидации теста.

(2) Валидизация теста. Более важной процедурой, нежели проверка на тенденцию к социально одобряемым ответам, является валидизация теста. Если мы покажем, что тест валиден, то тогда не имеет значения, влияет ли на его выполнение установка на социально одобряемые ответы, или нет. Один из путей доказательства того, что тест свободен от действия установки на социально одобряемые ответы, - показать при исследовании валидности, что он не связан со шкалой социальной желательности Кроуна-Марлоу (Crowne-Marlowe Social Desirability Scale; Crowne и Marlowe, 1964), шкалой, состоящей из явно социально желательных (нежелательных) утверждений.

ВЫВОДЫ ОТНОСИТЕЛЬНО УСТАНОВКИ НА СОЦИАЛЬНО ОДОБРЯЕМЫЕ ОТВЕТЫ

Нет сомнения, что установка на социально одобряемые ответы может влиять на результаты выполнения заданий, и, следовательно, на показатели тестов личности. Несмотря на это, ее влияние может быть минимизировано. Вот некоторые выводы:

- (1) Следует всячески избегать заданий с явно социально желательной или нежелательной окраской.
- (2) Процедура анализа заданий позволит устранить те задания, ответы на которые явно смещены в одном направлении ("да" или "нет").
- (3) Должна быть всегда продемонстрирована валидность теста, что позволит окончательно устранить все задания, в ответах на которые сказывается установка на социальную желательность.
- (4) Нет необходимости предпринимать столь тщательно разрабатываемые меры предосторожности против установки на социальную желательность, как это делал Edwards (1957).

Другие установки на ответы

ТЕНДЕНЦИЯ ПОДТВЕРЖДАТЬ КРАЙНИЕ ОТВЕТЫ

Это тенденция, сказывающаяся в рейтинговых шкалах и таких тестах, как, например, тест личностных шкал Комрея (Comrey Personality Scales; Comrey, 1970), то есть тех, в форму заданий которых входят рейтинговые шкалы. С моей точки зрения, преимущества рейтинговых шкал (в плане чувства удовлетворения у тех испытуемых, которые полагают, что на вопросы лучше отвечать в такой форме) сводятся на нет проблемами, возникающими с теми испытуемыми, которые склонны подтверждать крайние ответы. Для того, чтобы не спутать их с лицами, имеющими высокие показатели по тесту, сбалансированная шкала уже не поможет, потому что при такой установке каждый из крайних пунктов шкалы одинаково привлекателен. Единственный способ избежать этой проблемы - не использовать задания данного типа. Если же мы по каким-либо причинам вынуждены использовать

такие шкалы, то тщательный анализ заданий и валидизация переменных теста, вероятно, смогут устранить те вопросы (утверждения), которые особенно побуждают испытуемых проявлять эту установку.

Как утверждает Guilford (1959), наиболее благоприятной почвой для проявления любых установок на ответы являются те задания, которые неопределенны, неоднозначны и неконкретны. Значение этого утверждения невозможно преувеличить. Формулирование заданий путем следования указанным выше правилам позволит минимизировать влияние установок. Более того, тщательная валидизация теста продемонстрирует, что влияние этих искажающих факторов незначительно. Однако, это влияние не будет столь несущественным, если не будут приняты меры предосторожности. Применительно к установкам на крайние ответы можем рекомендовать избегать использовать задания типа рейтинговых.

ТЕНДЕНЦИЯ К ПОДТВЕРЖДЕНИЮ СРЕДНИХ ЗНАЧЕНИЙ

Несомненно, лучший способ избежать проявления этой установки - это использовать дихотомические задания. Однако, у некоторых испытуемых возникают сложности при работе с дихотомическими заданиями. Они считают, что на них невозможно отвечать, а поэтому (для этих испытуемых) желательнее введение некоторой средней категории ответов. Исследования (Bendig, 1959) показали столь высокую корреляцию между дихотомическими и трихотомическими заданиями, что риск проявления побочных явлений, подобных обсуждаемому, не оправдывает использования последних.

Хорошо сформулированные задания могут быть написаны так, что выбор средней категории не будет притягательным для испытуемых. Эта категория ответов особенно привлекательна тогда, когда оба крайних значения в равной степени не затрагивают, безразличны для испытуемых. Например: "Что бы вы предпочли: (а) осмотреть консервный завод; (б) пойти в местный музей; (в) не знаю?" Как могут быть выявлены такие вопросы при проверке теста в том случае, когда мы используем дихотомические задания? Автор данной книги использовал такой прием: во время проверки теста он снабдил его специальной инструкцией, в которой указывалось, что все задания, при выполнении которых испытуемым потребовалась средняя категория ответов, должны быть помечены крестиком. От тех заданий, которые часто отмечаются таким образом, лучше отказаться, если только они не оказались особенно эффективными в ходе анализа заданий.

Задания, для которых необходима средняя категория ответов, не смогут пройти процедуру анализа и будут автоматически устранены, а исследование валидности может показать, что установка на средние ответы не является существенным источником ошибок.

Довольно об установках на ответы. Если будут использоваться те приемы, о которых речь шла выше, то мы склонны полагать, что влияние установок на ответы будет минимальным. Дальнейшая валидизация тестов покажет, действительно ли это так.

Содержание вопросов (утверждений)

До сих пор при обсуждении принципов конструирования заданий основное внимание уделялось их форме. При этом мы предполагали, что знаем, какие черты (особенности) пытаемся тестировать. Тем не менее, справедливо было бы спросить, каким же именно образом мы принимаем решения о содержании заданий. Трудно делать общие заключения об этом, поскольку многое зависит от особенностей конкретной области поведения, для диагностики которой мы хотим разрабатывать тест личности. Чтобы проиллюстрировать, как принимать решения о содержании тестов, я приведу несколько примеров.

Теоретический пример

Kline (1968) выполнил экспериментальное исследование особенностей анального характера (Freud, 1908), в ходе которого для конструируемого теста было поставлено большое количество эмпирических задач. Существует ли анальный характер? Если да, то измеряется ли он другими личностными

опросниками? Имеет ли он отношение к научению ребенка справлять большую нужду? Изменяется ли он от культуры к культуре, как следовало бы ожидать, исходя из психоаналитической теории? Чтобы ответить на эти вопросы, было необходимо сконструировать и валидизировать инструмент для измерения анального характера. Чтобы сделать это, я должен был решить, какие черты должны войти в тест.

Далее следует краткое описание того, как содержание вопросов (утверждений) было отшлифовано, прежде чем они превратились в задания теста. Получившийся в результате этого тест (A13Q), похоже, выдержал подробные критически проверки (см. Kline, 1978). Были подробно изучены психоаналитические описания анального характера (напр., Abraham, 1921; Jones, 1923; Menninger, 1943), и составлен список всех черт и специфических примеров поведения. Примеры из публикации Jones (1923) перечислены ниже.

- (1) Промедление в удовлетворении потребностей, доводимое до их интенсивной концентрации, монотонная настойчивость, своевольная независимость - уверенность, что никто не может сделать что-либо так хорошо, как он сам.
- (2) Неспособность перепоручать работу.
- (3) Мелочное внимание к деталям.
- (4) Нежность к детям.
- (5) Склонность к доминированию.

Подобным образом был собран длинный подробный список черт и особенностей поведения. Затем они были преобразованы в задания теста с использованием всех тех правил и приемов, которые обсуждались в данной главе.

Практический пример

Нет сомнений в том, что полезно иметь инструмент для измерения агрессивности, например, в центре по перевоспитанию правонарушителей или для малолетних заключенных. Основным подходом в этом случае было бы составление списка всех проявлений агрессивности, которые только можно вообразить, преобразование их в задания и апробация на соответствующей выборке. Следует отметить, что при таком подходе не стоит вопрос о том, что есть агрессия как личностное измерение. Для этого необходимо специальное исследование. Так, факторный анализ заданий продемонстрировал бы, есть ли только одно измерение, ряд коррелирующих между собой измерений, несколько некоррелирующих измерений или ни одного реального синдрома поведения, которые могли бы быть значимо описаны как агрессивность.

Получение содержания задания

Возвращаясь к нашему примеру, отметим, что есть набор признаков агрессивного поведения, которые бы могли быть использованы в заданиях: постоянно дерется (не в шутку); когда раздражен, вступает в драку по любому поводу; использует в драке разбитые бутылки; носит с собой бритвы, кастеты, велосипедные цепи, складные ножи, ремни с металлическими бляхами, имеет тяжелые ботинки; бьет упавших противников ногами; избивал кого-либо до потери сознания; страдал от несправедливости и оскорблений; кричит, если раздражен; разбрасывает вещи по комнате; лжет, чтобы раздражить конкурентов; придумывает страшные истории. Затем подобный список при помощи приемов, описанных в данной главе, может быть преобразован в задания теста.

Заключение

При тех сложных и неизбежно неточных инструкциях для конструирования заданий тестов личности, приведенных в данной главе, наиболее ценным может оказаться краткое пошаговое описание

интересующей нас процедуры. Оно и приводится ниже. Полное обоснование всех этих положений содержится в тексте главы.

- (1) Перечислите черты и особенности поведения из описаний в психологической литературе.
- (2) Преобразуйте описание каждой черты или особенности поведения в задание. Это лучше всего сделать при помощи обычных форм заданий опросников: (а) дихотомических заданий (типа "да-нет"); (б) трихотомических заданий (типа "да-?-нет"); (в) альтернативных заданий; (г) заданий типа "нравится-не нравится"; (д) заданий с вынужденным выбором; (е) заданий с рейтинговыми шкалами. Выбор формы задания зависит от особенностей используемого материала и оценки преимуществ и недостатков каждого типа заданий.
- (3) Независимо от используемого типа заданий старайтесь, чтобы назначение каждого задания не было слишком очевидным.
- (4) формулируйте каждое задание понятно и недвусмысленно.
- (5) Насколько возможно, обеспечьте, чтобы каждое задание относилось к какой-либо конкретной области поведения.
- (6) В каждом задании должен содержаться только один вопрос или одно утверждение.
- (7) Избегайте использовать термины частоты и другие субъективные выражения.
- (8) Когда это возможно, задания должны касаться поведения, а не чувств и отношений.
- (9) Обеспечьте, чтобы ответы на задания давались быстро.
- (10) Избегайте влияния основных установок на ответы, таких как установка на согласие и установка на социально одобряемые ответы.
- (11) Вероятность проявления установки на согласие лучше всего уменьшается формулированием понятных заданий и использованием сбалансированных шкал.
- (12) Проверьте при помощи процедур анализа заданий то, что тенденция к согласию - несущественный фактор.
- (13) Установки на социально одобряемые ответы легче всего избежать путем тщательного формулирования заданий.
- (14) Проверьте при помощи техники анализа заданий, что установка на социально одобряемые ответы - несущественный фактор.
- (15) Избегайте с помощью адекватной формы заданий влияния установок на выбор крайних или средних категорий ответов.
- (16) Проверьте при помощи процедуры валидизации, что влияние установок на ответы не является существенным. Если тесты валидны, то установки на ответы не могут оказывать влияние на их показатели.

Глава 4. СОЗДАНИЕ НАДЕЖНЫХ ТЕСТОВ: КОНСТРУИРОВАНИЕ ТЕСТОВ ДРУГИХ ТИПОВ.

В этой главе будет обсуждаться разработка заданий и подбор материала для других типов психологических тестов. Многое из того, что уже обсуждалось, особенно в предыдущей главе, применимо и здесь, а поэтому не будет упоминаться вновь. Основное внимание будет уделено приемам, специфичным для конструирования объективных тестов личности или темперамента, проективных тестов,

инструментов измерения настроения и интересов и, наконец, шкалам аттитюдов. Начнем с объективных тестов.

ОБЪЕКТИВНЫЕ ТЕСТЫ

Определение

Определение, используемое здесь, принадлежит Кэттеллу и его коллегам (напр., Cattell, 1957). Объективный тест – это тест, цель которого скрыта от испытуемого (а поэтому результаты не могут быть фальсифицированы), и данные, полученные с его помощью, могут быть оценены независимо от лица, проводящего тестирование и интерпретацию. Преимущество таких тестов состоит в их практическом значении. Когда испытуемые не могут произвольно манипулировать показателями, такой тест может быть использован в процедурах профотбора. Это означает, что на протяжении многих лет может быть накоплено огромное количество подлинно жизненных критериальных данных о соответствии между профессиональными успехами и переменными объективных тестов. Кроме того, отсутствие возможности произвольного искажения результатов полезно и в тех областях, в которых предъявляются менее жесткие требования, таких как профориентация и психиатрия. Однако и в случае объективных тестов мы не можем быть уверены, что испытуемые не будут искажать их результаты.

В объективных тестах произвольная фальсификация все же возможна. Так, в тесте "Slow Line-Drawing Test" испытуемые могут чертить линии не так плавно, как они в состоянии это делать. Однако, они не знают, ни как это скажется на их показателях, ни даже самих показателей. Подобное неповиновение инструкциям само по себе симптоматично для определенных черт личности, и на деле может быть использовано как переменная объективного теста.

Последнее предложение дает нам ключ к разрешению основной и имеющей глубокие корни проблемы конструирования объективных тестов. В соответствии с нашим определением, почти любое задание, которое может быть объективно оценено и не является непосредственной реакцией на вопрос (как в тестах личности), может быть объективным тестом. Например, данную страницу рукописи можно использовать для получения следующих объективно тестируемых переменных:

- (1) Продолжительность времени написания.
- (2) Количество слов на странице.
- (3) Количество существительных.
- (4) Количество глаголов.
- (5) Количество зачеркиваний.
- (6) Давление пера при письме.
- (7) Количество имен собственных.

Конечно же, психологическое значение таких переменных неизвестно – это вторая проблема, связанная с объективными тестами. Следовательно, в процессе создания объективных тестов нам предстоит решить и эту проблему. На каком основании мы можем отбирать объективные тесты (которых, по определению, может быть очень много), чтобы хоть в какой-то мере надеяться на то, что они валидны? Если тест является очевидно валидным, то он уже не является объективным, так как может быть угадана его цель. Следовательно, нам необходимы некие общие принципы для конструирования объективных тестов – таксономия объективных тестов.

Преимущества объективных тестов

При наличии упомянутых выше проблем, возникающих при разработке объективных тестов, вполне резонно спросить, стоит ли пытаться их разрабатывать, особенно если учесть то, что их единственное преимущество, – значительно меньшая вероятность фальсификации результатов, нежели в других методиках. В конце концов, лучше иметь валидный тест, результаты которого, вероятно, могут быть искажены испытуемым, чем невалидный, но противостоящий всем попыткам фальсификации.

Cattell и Kline (1977) утверждают, что существует теоретическое преимущество объективных тестов, и столь значительное по сравнению с опросниками, что усилия, затраченные на разработку валидных объективных тестов, вполне оправданы. Дело прежде всего в том, что значение (смысл) слов и содержание вопросов изменяются со временем, а это ставит под вопрос состоятельность лонгитюдных исследований личности. Значение наших вопросов (утверждений), конечно же, изменяется от культуры к культуре, поэтому проведение кросс-культурных исследований личности при помощи опросников чрезвычайно сложно, если вообще возможно. Неоспорим и тот факт, что с различиями в понимании наших заданий мы сталкиваемся при обследовании разных общественных классов в рамках одной и той же культуры. Kelly (1955) как-то даже заявил, что значения слов настолько различны для разных индивидуумов, что любая форма стандартизированного тестирования не имеет сколько-нибудь значительной диагностической ценности. Вследствие этого Келли и его последователи придавали особое значение методике репертуарных решеток, где каждая решетка была индивидуальна для испытуемого, чьи качества исследовались. Несколько примеров убедят в наличии этого обстоятельства.

(1) "Нравятся ли вам веселые вечеринки?" (англ.: "Are you like gay parties?"). По-русски этот вопрос звучит весьма прилично, и вполне приемлем как задание теста. По-другому дело обстоит с его английским вариантом. Примерно до 1960 года этот вопрос (в англоязычных странах) не имел гомосексуального подтекста. "Gay party" – это была вечеринка, отличавшаяся доброжелательным отношением ее участников друг к другу, живым весельем и юмором. Сегодня словосочетание "gay party" означает "встреча для гомосексуалистов".

(2) "Нравится ли вам игра "Урони платок"?" Этот вопрос имеет смысл только в условиях североамериканской культуры: ни в Англии, ни в других странах Европы в эту игру не играют и, следовательно, там ее скрытый смысл не известен. Поэтому этот вопрос нельзя использовать в кросс-культурных исследованиях.

(3) "Регулярно ли вы ходите в кино?" Сейчас положительный ответ на этот вопрос указал бы на неподдельный интерес к фильмам. А двадцать лет назад, до всеобщего распространения телевидения, посещение кинотеатра было принятой большинством людей формой развлечения, и ответ "нет" мог, вероятно, представлять интерес.

Объективные тесты, напротив, должны помочь выявить данные, которые не подвержены влиянию изменяющегося смысла заданий, а, следовательно, необходимы для изучения тех преобразований личности, которые происходят во времени, а также для сравнения различных групп.

Принципы конструирования объективных тестов

Поскольку возможно предположить существование огромного количества разнообразных объективных тестов, необходимо, чтобы их разработчики руководствовались некоторыми общими принципами. Cattell и Warburton (1967) в руководстве по объективным тестам личности и мотивов, а в нем перечислены 688 тестов для измерения более чем 2300 переменных, считают это количество лишь небольшой частью тех, которые могли бы быть разработаны. Однако, эти авторы отдают себе отчет в том, что прежде чем подумать об отказе от разработки объективных тестов из-за их ошеломляющего потенциального количества (ибо даже многие из приведенных в их руководстве переменных остались без экспериментального опробования и с неустановленной валидностью), необходима некоторая их таксономия, и это сделано в упомянутой книге. Приведем краткое изложение.

Психологические тесты могут отличаться один от другого по следующим признакам:

(1) *Инструкции.* Ясно, что они важны, поскольку определяют то, как испытуемый воспринимает тест (хотя испытуемые могут не всегда придавать им значение).

(2) *Материал теста.* Он может быть взят из социальной области или, например, из физики.

(3) *Особенности оценивания* полученных от испытуемого ответов.

Понятно, что эти три элемента не являются независимыми, так как инструкции *должны* отражать определенный стимульный материал. Кэттелл и Уорбертон предпочли объединить первые два элемента и обозначить это термином "*стимульно-инструктивная ситуация*" (stimulus-instruction situation).

СТИМУЛЬНО-ИНСТРУКТИВНАЯ СИТУАЦИЯ

Стимульно-инструктивная ситуация изменяется в зависимости от ряда параметров:

- (1) Реакция или ее отсутствие. В общем требуется реакция на стимул. Но это не всегда так, как, например, в тестах на чувствительность к боли.
 - (2) Ограниченность versus неограниченность способов реагирования. В тестировании не известны случаи неограниченного разнообразия ответов. Предполагается, что испытуемые (как аргументируют Кэттелл и Уорбертон) не будут жевать материал тестов или портить его. Однако ответы могут быть относительно свободными, как, например, в тесте Роршаха, или с накладываемыми на них ограничениями, как при измерении времени реакции.
 - (3) Ответы, придуманные (созданные самостоятельно) versus отобранные из набора вариантов. Значение этого параметра очевидно. Примером последних являются ответы типа "правда-неправда", а примером первых – описание рисунков Теста тематической апперцепции (Thematic Apperception Test – ТАТ).
 - (4) Единичные versus повторяющиеся ответы.
 - (5) Последовательность ответов: упорядоченная versus неупорядоченная. Это относится к тому, до какой степени от испытуемых требуется давать ответы в определенном порядке.
 - (6) Однородность versus разнородность ответов. В некоторых тестах требуются ответы одного и того же вида, а в других предусмотрены ответы различных типов.
 - (7) Естественность темпа versus ограниченность времени на ответы. Сравните такие инструкции, как, например: "Работайте со свойственной вам скоростью", "Не тратьте слишком много времени на каждый вопрос" и "Работайте настолько быстро, насколько это возможно."
 - (8) Завершенность (окончателность) реакции versus реагирование на реакцию. Испытуемый реагирует либо на материал теста, либо на свою собственную реакцию, например, тот или иной стимул как-либо оценивается, вызывает ассоциации или воспоминания.
- Cattell и Warburton (1967) считают, что это основные, не требующие дополнительной аргументации, параметры вариативности, связанные со стимульно-инструктивной ситуацией. Однако, ими представлены и три других, более субъективных параметра, которые при разработке объективных тестов могут приобрести особое значение.
- (9) Явный смысл versus неявный. Некоторые тесты не имеют никакого вне их лежащего значения, такие как измеряющие время реакции. Однако в определенных тестах (там, где интересуются мнением о чем-либо) это присутствует. Большинство видов искажений и сложностей, связанных с субъективным истолкованием слов и понятий возникают именно в тех тестах, значение которых неявно.
 - (10) Представление теста в виде отдельных заданий versus целостное представление. Этот параметр касается различий между тестами, состоящими из отдельных заданий или из одного единственного.
 - (11) Психологическая сущность действия, необходимого для выполнения задания. Этот параметр учитывает, требуется ли для ответа: (а) знания (напр., верное суждение); (б) мнение о чувствах, отношениях; или (в) опознание или распознавание.

Как утверждают Cattell и Warburton (1967), существуют еще два параметра:

- (12) Различия в мотивах приемлемости ситуации тестирования.
- (13) Различия в свободе выхода из ситуации тестирования.

Вышесказанное обуславливает существование 2 типов стимульно-инструктивных ситуаций.

Значение приведенных классификационных принципов огромно: если бы мы конструировали такой материал для объективных тестов, в целом содержащий бы в себе все эти параметры, то использовали бы все возможные виды материала. Это особенно важно, поскольку при применении этих те-

стов всегда есть опасность, что для данного теста специфично проявляется больше вариативности, чем это желательно.

Как уже говорилось, еще одним важным источником различий между тестами является то, каким образом обрабатываются показатели, полученные с их помощью. Cattell и Warburton (1967) в исследовании ответов на тесты вводят некоторые *параметры обработки показателей по тесту*. Хотя эти параметры и не являются полностью независимыми от стимульно-инструктивных параметров, приведенных выше, они существенно отличны.

Прежде чем перейти к описанию этих параметров, следует подчеркнуть, что ответы на тесты образуют весьма значительное число переменных. Например, если бы мы рассматривали в качестве объективного теста личностный опросник, то, в дополнение к стандартной шкале подсчета показателей могли бы быть вычислены следующие переменные: время выполнения теста, время выполнения половины теста, количество подтвержденных ответов "да", количество подтвержденных ответов "нет", количество альтернативных ответов, количество пропущенных заданий, количество неопределенных ответов ("не знаю", "не уверен"). Следует заметить, что все эти показатели, в отличие от показателя по стандартной шкале, являются объективными, поскольку ни один испытуемый не может знать, что они измеряют, и они могут быть измерены с высокой надежностью, не зависящей от лица, проводящего тестирование.

Параметры оценивания ответов

(1) Объективность versus субъективность оценивания. Этот момент касается в основном тех тестов, в которых используются задания на самооценку. Будет ли ответ испытуемого оцениваться так, как это предполагается инструкцией или каким-либо неизвестным для него образом? В этом нет никакого обмана для испытуемых. Например, объективным тестом является *тест критических оценок* (Critical Evaluations Test, T8, Cattell и Warburton (1967)). В нем испытуемых просят указать, является ли определенная производительность труда, – например, если у официантки уходит десять минут на то, чтобы принести шесть порций на столик, – очень хорошей, хорошей, плохой, очень плохой. А в этом тесте *между прочим* определяется количество критических оценок. Субъективное содержание заданий не имеет отношения к оценке. Таким образом, этот тест, внешне напоминающий опросник, является, фактически, полностью объективным, как это и определено в начале данной главы. Разумеется, все те тесты, которые будут обсуждаться в этой главе, являются объективными в смысле этого первого параметра.

(2) Целостное поведение (весь организм) versus физиологическая реакция (часть организма). Типичной физиологической реакцией может быть покраснение от смущения или тремор.

(3) Измерение одного возможного ответа versus классификация разнообразных ответов (параметрическое оценивание versus непараметрическое). Параметрически измеряются такие особенности ответа: время, количество ошибок, повторений. Непараметрически раскрывается количество и разнообразие классов ответов. В этом смысле многие тесты на креативность оцениваются непараметрически.

(4) Общее количество ответов versus часть, соответствующая некоторому критерию. Это пересекается с описанным выше третьим параметром, поскольку, как указывают Cattell и Warburton, различные оценки могут быть любой из указанных здесь категорий.

(5) Единый однородный показатель versus разнотипные относительные показатели. Определить единый показатель можно только тогда, когда тест оценивается в целом, как в случае личностных опросников. Разнотипные относительные показатели могут определяться в следующих формах: различия во времени выполнения первой и второй частей теста; запоминание материала при обычных условиях и при отвлекающих обстоятельствах. И, наконец, Cattell и Warburton добавили шестой параметр: (6) Нормативное оценивание versus ипсативное (личностное). Однако этот параметр, по-видимому, представляет совершенно иной тип классификации, чем та, о которой шла речь выше, так как он оказывает влияние на все тесты и более правильно было бы отнести его к области стандартизации, а не конструирования тестов. Если мы примем в рассмотрение и этот параметр, у нас получится 2⁶ возможных типов оценивания ответов.

Таким образом, выделенные параметры, связанные со стимульно-инструктивной ситуацией и оценением ответов, позволят нам конструировать тесты любого типа. Однако, такой подход, хотя и является полезным, все же не может помочь нам в том, что касается содержания объективных тестов. Кроме того, показано, что потенциальное количество типов объективных тестов действительно огромно: $2^{13} \times 2^6$, что значительно больше 50000. Многие из них, однако, нежизнеспособны, но даже при этом их количество все еще слишком велико. Выделенные параметры указывают, до какой степени могут различаться задания каждого возможного типа объективных тестов, которые проходили испытания. Мы никогда не будем испытывать недостатка в заданиях!

Для использования в практике конструирования объективных тестов эта классификационная схема должна быть сокращена. Кэттелл и Уорбертон подчеркивают, что особую важность имеют шестьдесят четыре варианта, полученные из трех наиболее важных ситуаций и параметров оценивания, генерирующих по восемь возможных вариантов. Огромное количество тестов в их сборнике не покрыло все шестьдесят четыре типа, так что такая краткая версия таксономии тестов также еще оставляет пространство деятельности для творческого разработчика тестов. Такая таксономия, основывающаяся на характеристиках тестов, может помочь разработчику создавать разнообразные тесты, но есть одна основная трудность: как мы узнаем, что созданные средства измерения будут измерять, например, переменные темперамента, а не другие? Иными словами, как нам получить тест для измерения темперамента, а не, например, способностей? Данная таксономия в этом помочь не может, и разработчику тестов нужна дальнейшая информация.

Различия между объективными тестами способностей, темперамента и динамики (мотивов)

Один из методов определения того, что же измеряют объективные тесты, – подвергнуть их факторному анализу вместе с хорошо известными отличительными переменными (маркерами) этих трех модальностей (способностей, темперамента, мотивов). С помощью факторного анализа будет выяснено, переменные какой модальности измеряет данный тест. Однако, хотя это и эффективная процедура, необходимая до реального использования любого объективного теста, она не указывает нам, как можно сконструировать тест нужного нам типа. Это лишь адекватное средство проверки, но не руководство для конструирования тестов.

Cattell и Warburton (1967) посвятили обсуждению этой проблемы довольно много места, а здесь мы лишь кратко подытожим их рассуждения, поскольку они обеспечивают некоторое рациональное основание для выбора содержания заданий (в дополнение к таксономии, призванной помочь в выборе формы тестов).

(1) Различаются два вида ситуаций (ситуационализм, фактически, не является противоположностью психологии черт): побуждающие (incentives) и связанные с осуществлением действия (complexities).

(2) При изменении побудителей (incentives) изменяются показатели динамических тестов. При изменении деятельностей (complexities) изменяются также показатели тестов способностей. Тесты темперамента включают все остальные ситуации.

(3) Определение побудителей (incentives): побудители порождают стремление к некоторой цели и являются символом цели или удовлетворения от достижения цели, что само по себе может быть выявлено только анализом этого процесса. Это статистический анализ последовательности поведения во времени. Данная последовательность является набором действий, ведущих к достижению цели. Побуждающая ситуация, таким образом, распознается по ее отношению к цели (то есть по общим колебаниям в силе стремления и постоянству доминирования). Kline и Grindley (1974) действительно показали именно такие флуктуации между динамическими измерениями и ситуациями.

(4) Определение деятельностей (complexities). Когда побудители в данных обстоятельствах уже распознаны, можно определить и деятельность: это все то, что не является побудителем.

(5) Конечно, на практике это означает, что измерение динамики и способностей тесно взаимосвязано. Однако могут быть получены относительно "чистые" средства измерения для каждой из этих областей. Например, если бы у нас были очень простые средства измерения способностей, то различия в показателях отражали бы не способности, а динамику (например, насколько испытуемые старательно выполняют задание). Таким образом, можно так управлять уровнями деятельностей и по-

будителей, что объективные тесты могут стать практически "чистыми" средствами измерения каждой модальности.

(6) Нет необходимости говорить, что все тесты, сконструированные таким образом, должны быть подвергнуты проверке при помощи факторного анализа. Нужно убедиться, что то, что они измеряют, соответствует тому, что от них ожидали. Все тесты, предположительно измеряющие способности, должны дать близкие результаты, то же касается измерения мотивационных, динамических переменных.

С этим логическим обоснованием, с различиями между побудителями и деятельностями, у разработчика объективных тестов уже есть некоторое руководство, по крайней мере относительно выявления того, что его тесты измеряют, хотя все они по-прежнему должны подвергаться последующей проверке факторным анализом.

Однако, как сейчас должно быть очевидно читателю, эти принципы и таксономии интересны и, возможно, могут явиться стимулом к работе, но, вероятно, слишком абстрактны для практического использования при конструировании реальных тестов. Cattell и Warburton (1967), осознавая это, в своей работе чистосердечно признают, что при конструировании объективных тестов им помогало лишь интуитивное предвидение, основанное на значительном опыте исследований в этой области. Это мы сейчас кратко и обсудим.

Практические советы для разработки объективных тестов личности и мотивов

Следует избегать следующих пяти ошибок, которые Cattell и Warburton (1967) упоминают как распространенные среди любителей, и о которых необходимо знать, прежде чем приступать к работе по конструированию объективных тестов.

- (1) Не останавливайтесь на очевидно валидных заданиях из опросников.
- (2) Не применяйте задач или заданий-головоломок (загадок). С их помощью вы скорее всего будете измерять факторы способностей.
- (3) Не возлагайте слишком большие надежды на "стрессовую ситуацию". Как утверждают Cattell и Warburton (1967), это может сработать для выявления агрессии или страха, но ведь есть еще и другие эмоции.
- (4) Эстетические и стилистические предпочтения могут эффективно выявить лишь некоторые аспекты личности. Однако, будет абсурдом ожидать, что такие тесты выявят все возможные аспекты. Нужно учитывать тот очевидный факт, что на ответы по этим тестам влияют уровень образования и культуры.
- (5) Избегайте упрощенных способов использования проективных тестов. Wenig (1952) показал, что такие тесты диагностируют весьма широкий комплекс параметров, требующий факторно-аналитического исследования.
- (6) Используйте задания-вопросы, точно соответствующие определенным типам поведения.

Вышесказанное является основой для создания эффективных объективных тестов.

Решение некоторых общих проблем, связанных с объективными тестами

Существует ряд проблем, из-за которых интерпретация показателей любого теста (не только объективного) может стать сомнительной, и их следует учитывать при конструировании объективных тестов. Подробное обсуждение этих проблем, к которому можно отослать читателей, представлено в книге Cattell и Warburton (1967). Кратко подытожим те положения, о которых идет речь в этой работе.

РАЗЛИЧИЯ В МОТИВАЦИИ ИСПЫТУЕМЫХ

Различия в мотивации испытуемых при разработке тестов имеют особое значение в научных исследованиях, тогда как при отборе и выдаче рекомендаций значение этого источника индивидуальных различий по-видимому минимально. В конце концов, если для выполнения объективного теста тре-

буется, например, значительная концентрация внимания, почему испытуемый должен делать усилия, чтобы выполнить все наилучшим образом? С другой стороны, встречаются испытуемые, которые вообще все делают наилучшим образом. Cattell и Warburton (1967) обсуждают пять приемов построения объективных тестов, призванных в некоторой мере минимизировать различия в мотивации испытуемых.

(1) Поделите тест на две части. Получаемым показателем является отношение или разность оценок по первой и второй части, и две части сравниваются, при том резонном предположении, что у каждого испытуемого мотивация на выполнение каждой части одинакова, и таким образом этот эффект устраняется. Примером такой методики является тест "силы Эго", основанный на эффективности запоминания при отвлекающих факторах. В этом тесте показателем "силы Эго" является разность между оценкой по запоминанию чисел в обычных условиях и оценкой по запоминанию чисел в том случае, когда они перемежаются шутками. Это явно полезный прием для проектирования объективных тестов.

(2) Используйте в качестве мотивов базовые эрги, а не чувства. В терминологии Кэттелла (см. Cattell и Child, 1975; Cattell и Kline, 1977) эрги – это базовые влечения, такие как голод, половое влечение или страх, тогда как чувства – это то, что прививается, формируется при обучении и воспитании в определенной культуре, например, религиозные чувства или чувства по отношению к своей семье. Поскольку, как обсуждалось у Cattell и Child (1975), базовые эрги менее подвержены изменчивости, чем чувства, то количество различий вследствие изменений в мотивации сокращается, если при выполнении теста в качестве мотивов широко вовлекаются базовые эрги. Проще всего при этом управлять страхом (удара электрическим током) или сексуальными желаниями (предъявляя картинки с обнаженными), хотя следует отдавать себе отчет, что эти влечения не могут проявиться так, как в реальности (по степени возбуждения), ибо существуют определенные этические нормы, реализующиеся и в ситуации тестирования.

(3) Ограничивайте подсчет показателей стилистическими или формальными аспектами выполнения теста. Эти переменные менее подвержены изменениям мотивации, чем большинство других. Кэттелл и Уорбертон выбрали почерк как пример того показателя, который обычно остается распознаваемым для каждого индивидуума в разнообразнейших ситуациях.

Характеризуя эти три приема преодоления мотивационных искажений, необходимо отметить один очевидный, но важный момент: они применимы в основном к тестам темперамента, а не динамики.

(4) Вполне возможно, что мотивационный уровень испытуемого при работе над всей батареей тестов может быть отделен факторным анализом (вероятно, в один или более факторов) от других независимых факторов. Если это так, то переменные, имеющие тенденцию быть нагруженными такими факторами, могут быть отброшены.

(5) И наконец, разработчик объективного теста может намеренно искать способ задействовать мотивы каждого испытуемого таким образом, чтобы они были активно вовлечены в тестовые процедуры, но при этом следует постоянно помнить о том, что нельзя выходить за границы той системы ценностей, которой привержены испытуемые. К сожалению, это тот принцип, который легче понять, чем воплотить в практику.

ЛИЧНОСТЬ И ТЕСТОВАЯ СИТУАЦИЯ

Проявления личности связаны с социумом, однако при тестировании социум представлен только в виде тестовых ситуаций. Это, конечно, еще одна существенная проблема тестирования, кстати, давшая повод ситуационалистам (напр., Mischel, 1968) подвергнуть сомнению эффективность личностных тестов. Так, Mischel утверждал, что факторы, полученные при помощи традиционных личностных опросников, являются по своему существу факторами ситуации тестирования.

Чтобы преодолеть эту трудность, Cattell и Warburton (1967) предложили индивидуальные тестовые ситуации, которые несомненно включают в себя социальные взаимодействия. Однако, эти же авторы, к сожалению, утверждают, что такие мини-ситуации трудно создавать и трудно использовать в практической психологии, таким образом значительно уменьшая их значение для достижения каких-либо иных целей, кроме теоретических. Правда, эти авторы надеются, что последующие исследования позволят измерять при помощи других объективных средств более простого вида те факторы,

которые определяются тестами мини-ситуаций. Если это произойдет, то от ситуационных тестов можно будет отказаться. Однако, без них разработчик объективных тестов не может испытывать уверенность в том, что он измерял социальное поведение (поведенческие проявления личности). Было разработано несколько групповых тестов, нагруженных факторами социального взаимодействия (они перечислены в Cattell и Warburton, 1967), но в этой области еще многое предстоит сделать.

ВЛИЯНИЕ СПОСОБНОСТЕЙ И ДОСТИЖЕНИЙ НА ПОКАЗАТЕЛИ ОБЪЕКТИВНЫХ ТЕСТОВ

Это основная трудность, которая должна быть преодолена при конструировании объективных тестов личности (тестов темперамента и динамики). Например, известно (Cattell и Child, 1975), что информация о релевантных цели обследования предметах и явлениях является эффективным средством измерения интереса. Элементарная рефлексия, однако, с очевидностью показывает, как может быть искажено такое измерение. Незначительный интерес у студента к конным бегам (а корпуса естественно-научных факультетов Киевского университета находится рядом с ипподромом) может привести к некоторому запасу знаний, который не может быть сравним со знаниями менее интеллектуального человека, для которого это было единственным видом отдыха. Информационный тест, следовательно, будет работать неправильно.

Таким образом, должны разрабатываться приемы проектирования тестов, которые минимизируют влияние способностей и достижений. Cattell и Warburton (1967) дают следующие рекомендации.

- (1) Устраняйте, насколько возможно, из объективных тестов переменные способностей – например, направленные на выявление отношений, словарный запас, общие знания.
- (2) Как было ранее показано на примере разных мотивов, разбивайте тест на две части и используйте отношение или разность показателей. Это устраняет влияние уровня способностей испытуемого, как это происходило и с уровнем мотивации.
- (3) Подвергните тесты факторному анализу и устраните те, которые нагружены факторами способностей.
- (4) Используйте в батарее объективных тестов настолько широкий набор проб содержательного материала, навыков и интересов, насколько это возможно.

Переменные поведения должны обеспечивать разнообразие ролевых ситуаций, так как личностные факторы лучше всего, вероятно, выражаются в различных ролях. Содержание также должно быть подобрано таким образом, чтобы затронуть интересы всего диапазона выборочной совокупности испытуемых. Это также важно, поскольку *a priori* маловероятно, что одна частная тестовая форма, пусть даже эффективная, сможет охватить полностью всю личностную сферу, то есть всю гамму факторов.

ГРУППОВЫЕ ИЛИ ИНДИВИДУАЛЬНЫЕ ТЕСТЫ

Ориентируйтесь на разработку простых тестов для группового тестирования, а не индивидуальных тестов. Это необходимо как для исследований больших выборок, так и в практической психологии, где групповая работа является необходимостью. Следует, однако указать, что создание групповой версии индивидуального теста требует значительной изобретательности и исследовательских усилий для демонстрации того, что каждая из этих версий действительно измеряют одну и ту же переменную. В то же время некоторые объективные тесты, например, те, основу которых составляют физиологические измерения, нельзя преобразовать в групповую форму.

В дополнение к этим полупормальным рекомендациям Cattell и Warburton (1967) обсуждают интуитивные основания некоторых из своих тестов, поскольку при разработке объективных тестов все еще необходимо определенное чутье, интуиция, ибо нет еще единого алгоритма. Будет достаточно лишь перечислить эти основания, так как разработчики тестов могут их либо использовать, либо нет, иначе говоря, некоторые разработчики могут обладать достаточным воображением, чтобы творить тесты с их помощью, тогда как другие не смогут делать этого, поскольку правила, вытекающие из этих оснований, не могут быть точно сформулированы. Наиболее важными основаниями являются: (1) клиническая интуиция; (2) наблюдение в повседневной жизни случаев, являющихся определяющими для характеристики личности; (3) повседневная устная речь – пословицы и поговорки; (4)

эмоциональные ситуации в играх, например, в карточных; (5) поведение, наблюдаемое в ходе беседы; и (6) литературные источники.

Cattell и Warburton (1967) также утверждают, что ряд положений, которые, по их мнению, являются нечетко сформулированными психологическими принципами, весьма помогали им при разработке тестов. Однако автор данной книги счел эти положения слишком расплывчатыми, чтобы их использовать (напр., "избирательное действие восприятия и памяти в соответствии с общей ориентацией") при практической разработке тестов, хотя они весьма полезны для понимания оснований некоторых объективных тестов Кэттелла.

И наконец, Cattell и Warburton считают, что при разработке объективных тестов могут быть полезны некоторые из экспериментально-психологических открытий (таких как те, которые касаются, например, взаимосвязей между обусловливанием и научением, ЭЭГ и характеристиками личности; см. Eysenck, 1967).

Таковы советы относительно формы и содержания объективных тестов для пытающихся их создавать, тестов, которые будут коррелировать с некоторыми внешними критериями. Есть надежда, что все эти советы окажут свое воздействие, побуждая читателей к разработке тестов. Несмотря на это, следует еще раз подчеркнуть, что ни один тест не может использоваться, прежде чем не будет продемонстрировано, что он измеряет те переменные, для измерения которых предназначен.

Объективные тесты и факторы личности Cattell и Warburton обсуждают также вопрос о том, как можно проектировать объективные тесты по отношению к явным конструктам, под которыми они, естественно, понимают личностные факторы. Конечно, принципы, обсуждаемые ниже, могли бы использоваться и для разработки тестов, не основанных на факторной стратегии. Однако, факторно-аналитические конструкты по определению имеют очевидное подтверждение (их факторные нагрузки), следовательно они представляют прекрасный объект для измерений (в отличие от многих клинических конструктов, за которыми может не быть никакой реальности, кроме воображения их создателей).

Суть этого метода состоит в использовании переменных-маркеров для уже установленных факторов (существуют их разнообразные списки, см., напр., Howarth, 1976, для факторов личностных опросников), а затем разработки тестов, которые, вероятно, нагружены этими факторами. Последующий факторный анализ выявляет тесты, которые нагружены этими факторами. Этот метод идеально подходит для разработки объективно-тестовых средств измерения, эквивалентных факторам, известным по другим тестам – полезная процедура, поскольку, по крайней мере, с точки зрения профотбора, результаты выполнения объективных тестов не могут быть фальсифицированы. Однако, как показывают на примерах Cattell и Warburton, этот метод может привести к открытию новых факторов, не обнаруживаемых ни при каком другом типе тестирования. Это можно легко себе представить в случае, когда множество объективных тестов формируют фактор, размещенный между двумя факторами-маркерами.

При разработке объективных тестов по уже установленным факторам, если мы хотим получить точные результаты, следует помнить о следующих моментах:

- (1) Необходимо воспроизведение всех факторных структур на различных выборках. В идеале, как показывает Nunnally (1978), для таких исследований требуется в десять раз больше испытуемых, чем переменных, с которыми ведется работа. Однако, в случае повторения выделенной факторной структуры, такое строгое требование не является, вероятно, необходимым.
- (2) Изучение тестов, нагруженных некоторым фактором, помогает уточнить конструкт рассматриваемого фактора. Так иногда нагруженность объективных тестов помогает узнать что-то новое о тех факторах, которые считались хорошо изученными.
- (3) Гораздо более интересным является то, что факторные нагрузки могут помочь прояснить, что же измеряют объективные тесты (вечная проблема объективных тестов).
- (4) Исходя из факторной нагруженности тестов нередко можно создать новые тесты. Другими словами, непосредственное наблюдение за тем, как выявляются переменные (то есть как бы взгляд назад) чрезвычайно полезно как вспомогательное средство при разработке тестов. Таким образом, одним из наиболее важных путей создания объективных тестов является изучение тех реальных

факторов, которые из них выделяются. Их порой трудно идентифицировать без последующего использования факторного анализа и разнообразного экспериментального материала, но по крайней мере при попытке измерения таких факторов разработчик объективных тестов будет нацелен на статистически значимый конструкт.

ЗАКЛЮЧЕНИЕ

Было немало сказано о двух подходах к разработке объективных тестов, предназначенных для изучения темперамента. Вы увидели, что существуют некоторые полезные наставления, приемлемые для работающих в этой наиболее трудной области тестирования, правила ; касающиеся формы и содержания тестов. Несмотря на то, что хотя я и показал, как некоторые явные проблемы могут быть искусно обойдены, очевидно, что многое зависит от: (а) интуиции, основанной на знании общей психологии и ранее изученных факторов темперамента, и (б) реальных факторов, выявляющихся в исследованиях.

Я не предложил более подробного рассмотрения конструирования объективных тестов, потому что с моей точки зрения существует такое огромное количество уже созданных, но с неизвестной валидностью, что разработчику тестов лучше было бы посоветовать изучить на практике то, что уже сделано, прежде чем приниматься за эту трудную задачу самому. Многие из широко известных и эффективно валидизированных объективных тестов были разработаны Айзенком и его коллегами в Моудслейском госпитале, и большинство из этих средств измерения относятся к трем основным факторам: экстраверсии, нейротизму и психотизму. Подробности об этих тестах могут быть найдены в различных местах объемных публикаций Айзенка, но лучше всего воспользоваться работой, изданной в 1971 г. (Eysenck, 1971).

Однако, наиболее значительные усилия по конструированию объективных тестов, были предприняты в Иллинойсе Кэттеллом и его коллегами. Их исследования сформировали основу нашего обсуждения, как об этом и было сказано в начале главы. Cattell и Warburton (1967) приводят столь значительный перечень объективных тестов, что их психологическое значение еще предстоит осмыслить. Будущий разработчик объективных тестов должен, конечно, знать этот перечень и опробовать все подходящие тесты, содержащиеся в нем, прежде чем браться за их самостоятельную разработку.

ОБЪЕКТИВНЫЕ ТЕСТЫ МОТИВОВ ИЛИ ДИНАМИКИ

Все, что до сих пор было сказано о разработке объективных тестов, касается измерения темперамента. Очевидно, что таксономия типов тестов в равной степени применима и к объективным тестам мотивов. То же касается и второго подхода, основанного на изучении результатов факторного анализа. Тем не менее, Кэттеллом и его коллегами были предложены определенные принципы для разработки объективных тестов мотивов, которые не применимы к тестам темперамента, их мы сейчас и обсудим.

Принципы конструирования объективных тестов мотивов были подробно описаны в следующих работах: Cattell (1957), Cattell и Warburton (1967), Cattell и Child (1975), Cattell и Kline (1977). Здесь вы найдете краткое изложение этих принципов, вполне удовлетворяющее потребности разработчика тестов.

ТЕОРЕТИЧЕСКИЕ ОСНОВАНИЯ В теоретическом подходе Кэттелла к мотивации поведения предполагается, что мотивы проявляются в аттитюдах. Так, например, тот факт, что некоторый индивидум очень заинтересован в получении денег, должен отражаться в силе его побуждения (drive).

Cattell и Child (1976) использовали динамическую решетку, чтобы показать, как такие аттитюды могут соотноситься с побуждениями. В данном примере наличие денег желательно для высокой самооценки (самоощущение), для обеспечения своей семьи (супружеские чувства) и для гарантированного будущего своих детей (эрг защиты) . Обратите внимание, что здесь предполагается конечное количество побуждений у людей, как следует, например, из представлений McDougall (1932). Считается, что побуждения могут быть двух видов: (1) эрги, базовые для всех человеческих существ (например, половое побуждение); и (2) чувства, сформированные культурой и воспитанием (такие, как чувства к семье). Если мы в чем-либо заинтересованы, то наши аттитюды зависят существенным образом от того, как посредством активности они выражаются в наших побуждениях и чувствах.

Интерес к психоанализу, например, может оказаться средством выражения полового побуждения. Все это означает, что сущность объективных тестов, предназначенных для измерения мотивационных факторов, связана с аттитюдами и интересами.

Заслуживает упоминания еще один аспект теоретического подхода, принятого Кэттеллом и его коллегами. Это касается силы интереса. Два человека могут быть заинтересованы в одном и том же, но их интересы могут значительно различаться по силе. В работе, проведенной Cattell и Child (1975), четко показано, что из анализа аттитюдов с помощью объективных тестов, выделяются факторы силы и интересов, как и факторы, отражающие структуру интересов, то есть основные человеческие побуждения, эрги и чувства.

Принципы конструирования объективных тестов мотивов должны исходить из этих основополагающих теоретических положений. Это основа для измерения мотивов при помощи тестов аттитюдов и интересов. Cattell и Kline (1977) перечисляют шестьдесят восемь психологических принципов, на которых должно основываться измерение мотивов. Их можно также найти в Cattell и Kline (1975). Эти принципы приведены в табл. 4.1.

Как это подробно описано у Cattell и Child (1975), многие из этих принципов были реализованы в тестах и в воспроизводимых (повторно измеряемых) факторах силы интереса, были выделены и некоторые отчетливые эрги и чувства. Например, были опубликованы групповой тест MAT (Cattell и др., 1970) и его версия для подростков – SMAT.

Таблица 4.1. Некоторые принципы измерения мотивов, применяемые при конструировании тестов
С возрастом интереса в направлении действия ожидайте, что будут нарастать:

- (1) Предпочтения. Готовность признать предпочтение в направлении действия.
- (2) Аутизм: неправильное, искаженное восприятие объектов, звуков и т.п., в соответствии с интересным (примером является исследование Брунера по восприятию величины монет).
- (3) Аутизм: ошибочные убеждения. Необоснованная уверенность в том, что факты и происходящие события благоприятствуют направлению действий.
- (4) Искажение логических построений: средства – конечный результат. Готовность утверждать, что малоэффективные средства достижения цели в действительности являются эффективными.
- (5) Искажение логических построений: конечный результат – средства. Готовность утверждать, что конечного результата можно легко достичь при помощи несоответствующих для этого средств.
- (6) Искажение логических построений: индуктивных.
- (7) Искажение логических построений: дедуктивных.
- (8) Искажение логических построений: выявление отношений из воспринимаемого (напр., аналогий).
- (9) Выбор вспомогательных средств. Готовность использовать землю, труд и капиталовложения во имя интереса.
- (10) Макиавеллизм. Желание использовать предосудительные средства для достижения конечного результата во имя своего интереса. (Желание достичь цели любой ценой – все средства хороши).
- (11) Предпочтения в воображении. Готовность выбирать связанные с интересом темы для чтения, рассказа или объяснения.
- (12) Предпочтения в размышлениях. Проведение времени в размышлениях над тем, что связано с интересом.
- (13) Предпочтения в идентификации. Предпочитает быть похожим на индивидуумов, одобряющих направление его действий.

- (14) Защитная сдержанность (скрытность). Запинание ("спотыкание") при перечислении (назывании) негативных последствий избранного направления действий.
- (15) Защитная беглость речи. Беглость речи при перечислении позитивных следствий направления действий.
- (16) Защитная беглость речи. Беглость речи при перечислении оправдывающих обстоятельств в пользу действия.
- (17) Рационализация. Готовность интерпретировать информацию таким образом, чтобы представить свой интерес в более уважаемом виде, чем он есть, и т.п.
- (18) Наивная проекция. Ошибочное восприятие других людей как имеющих его собственные интересы.
- (19) Истинная проекция. Ошибочное восприятие других людей как осуществляющих столь же предосудительное поведение, связанное с преследованием интереса, как его собственное.
- (20) Проекция Id. Ошибочное восприятие других как имеющих его собственные примитивные желания, связанные с интересом.
- (21) Проекция Superego. Ошибочное восприятие других как имеющих его собственные убеждения в правоте, связанные с интересом.
- (22) Чувство вины. Выражение чувства вины от неучастия в деятельности, связанной с интересом.
- (23) Вовлеченность в конфликт. Трата времени на принятие решения при выборе путей для реализации интереса (оба альтернативных пути благоприятны для реализации интереса).
- (24) Вовлеченность в конфликт. Трата времени на принятие решения в конфликте, вызванном желанием избежать альтернативных решений (обе альтернативы противоположны интересу).
- (25) Реагирование на угрозу. При возникновении угрозы интересу падает сопротивление при изменении КГР.
- (26) Реагирование на угрозу. При возникновении угрозы интересу усиливается сердечно-сосудистая активность.
- (27) Физиологическая вовлеченность. При возрастании интереса (под угрозой или нет) усиливается сердечно-сосудистая активность.
- (28) Физиологическая вовлеченность. При возрастании интереса увеличивается температура пальцев.
- (29) Физиологическая вовлеченность. При возрастании интереса увеличивается ригидность мышц.
- (30) Интеграция при восприятии. Организация неструктурированного материала в соответствии с интересом.
- (31) Перцептивная замкнутость. Способность видеть незавершенные рисунки завершенными, если материал имеет отношение к интересу.
- (32) Избирательность восприятия. Легкость нахождения связанного с интересом материала, вкрапленного в сложно оформленные объемы информации.
- (33) Острота сенсорных ощущений. Тенденция ощущать свет ярче, звуки громче и т.п. при возросшем интересе.
- (34) Внимательность. Сопротивление отвлекающим факторам (световые сигналы, звуки и т.п.) при работе с относящимся к интересу материалом.
- (35) Спонтанное внимание. Непроизвольные движения, вызванные связанными с интересом раздражителями (напр., движения глаз).

- (36) Вовлеченность. Кажущаяся быстрота, с которой течет время при занятиях, связанных с интересом.
- (37) Настойчивость. Продолжение работы во имя интереса даже при встрече с трудностями.
- (38) Персеверативность. Упорство в неадаптивном поведении, связанном с интересом.
- (39) Рассеянность. Неспособность управлять вниманием при помехе со стороны связанного с интересом раздражителя (стимула).
- (40) Ретроактивное торможение при помехе выполнению связанной с интересом задачи.
- (41) Проактивное торможение заданием, связанным с интересом.
- (42) Рвение: усилия. Предвосхищение затраты еще больших усилий в направлении действия.
- (43) Деятельность: время. Трата времени в направлении действий.
- (44) Рвение: деньги. Предвосхищение еще больших затрат денег в направлении действия.
- (45) Деятельность: деньги. Трата денег в направлении действий.
- (46) Рвение: исследования. Готовность предпринять исследования для достижения связанных с интересом конечных результатов.
- (47) Импульсивность: решения. Быстрота принятия решений в пользу интереса, почти отсутствие конфликта.
- (48) Импульсивность: согласие. Быстрота согласия с мнениями, благоприятными для интереса.
- (49) Сила решения. Высшая степень уверенности в ситуациях, благоприятствующих направлению действия.
- (50) Быстрота воодушевления (раскачки): обучение. Быстрота воодушевления (раскачки) для изучения задач, связанных с интересом.
- (51) Обучение. Быстрота обучения связанному с интересом материалу.
- (52) Моторные навыки. Соответствующая производительность, когда затрагивается интерес.
- (53) Информация. Знания, имеющие отношение и влияющие на направление действия.
- (54) Соппротивление угасанию реакций, относящихся к интересу.
- (55) Контроль. Способность координировать деятельность, преследующую интерес.
- (56) Расположенность: легкость. Легкость, непринужденность в письменном изложении того, что связано с направлением действий.
- (57) Расположенность: свободные ассоциации. Готовность к ассоциациям на темы связанного с интересом материала без каких-либо ориентирующих стимулов.
- (58) Расположенность: скорость свободных ассоциаций. Увеличивающееся количество ассоциаций при возросшем интересе.
- (59) Расположенность: ориентированные ассоциации. Готовность к ассоциациям на темы связанного с интересом материала при стимуляции.
- (60) Расположенность: память. Свободное воспроизведение связанного с интересом материала.
- (61) Память на вознаграждения. Немедленное воспоминание о наградах, связанных с интересом.
- (62) Реминисценция. Эффект Уорда-Ховлэнда. Улучшение сохранения связанного с интересом материала при его воспроизведении через небольшие промежутки времени.

- (63) Реминисценция. Эффект Бэлларда-Уилльямса. Улучшение сохранения связанного с интересом материала при его воспроизведении через значительные промежутки времени.
- (64) Сохранение в памяти по Зейгарник. Тенденция вспоминать незавершенные задачи, связанные с интересом.
- (65) Персеверативность по Зейгарник. Готовность возвращаться к незавершенной задаче, связанной с интересом.
- (66) Защитная забывчивость. Неспособность вспомнить относящийся к интересу материал, если цель недостижима.
- (67) Фацилитативность рефлекса. Легкость, с которой возбуждаются определенные рефлексы при возросшем интересе.
- (68) Торможение рефлекса. Трудность возбуждения определенных рефлексов при возросшем интересе.

Источник: *R.B. Cattell and D. Child. Motivation and Dynamic Structure. London: Hoil, Rinehart and Winston, 1975.*

Однако, как и в случае объективных тестов темперамента, разработчику объективных тестов следовало бы посоветовать попытаться использовать тесты из "Сборника объективных тестов" (Cattell и Warburton, 1967), прежде чем приступать к разработке каких-либо собственных средств измерения. Действительно, конструирование объективных тестов требует значительного опыта и мастерства, а для психологов, которые хотят получить какие-либо тесты для практического использования, будь то решение исследовательских задач или профотбор, попытка конструировать такие тесты вряд ли будет удачной, несмотря на значительные затраты времени и работу по апробации вариантов теста и их валидации. Внимательный читатель мог заметить, что я опустил изложение специальных методов для измерения динамических конфликтов. Дело в том, что в настоящее время свидетельства исследователей в пользу валидности этих методик слишком умозрительны, чтобы в полной мере использовать их на практике, а не только для дальнейшей разработки этой несомненно важной области исследования мотивов. (Более подробные сведения читатели могут получить в работе Cattell и Child, 1975).

ДРУГИЕ АСПЕКТЫ КОНСТРУИРОВАНИЯ ОБЪЕКТИВНЫХ ТЕСТОВ МОТИВОВ

Хотя в таблице 4.1. приведены основополагающие принципы, в свете которых должны разрабатываться объективные тесты мотивов, заслуживают внимания и некоторые другие аспекты.

ПЕРЕХОД К ИПСАТИВНЫМ (ЛИЧНОСТНЫМ) ПОКАЗАТЕЛЯМ

Если мы рассмотрим принципы 42, 43, 44, 45 и 60 из табл.4.1., то становится очевидным, что добровольная трата времени и денег на некоторую деятельность и наличие соответствующей информации о ней являются мерами мотивации субъекта к этой деятельности. Однако, как уже указывалось ранее, такое измерение сводится на нет индивидуальными различиями, такими как способности, наличие свободного времени и денег. Для того, чтобы избавиться от этого очевидного источника ошибок, показатели ипсатизируются, то есть показателем испытуемого является отклонение от среднего значения, определенного у него тем же тестом. Так, если миллионер тратит 400 фунтов стерлингов на флейту, это явно указывает на меньший интерес, чем у студента консерватории, делающего то же самое. Ипсатизация устраняет этот не нужный нам источник различий.

ЧУВСТВА И ЭРГИ

Как уже было сказано в кратком изложении теоретических основ разработки объективных тестов мотивов, есть два аспекта в измерении мотивов: сила интереса и базовые мотивационные цели, побуждения (*drives*); в принятом нами подходе – эрги (*ergs*) и чувства (*sentiments*). Принципы конструирования тестов, представленные в табл.4.1., предназначены для решения первой задачи: измерения силы мотивов.

Следовательно, здесь имеет смысл рассмотреть конструирование тестов, направленных на измерение эргов и чувств. Фактически, как указывают Cattell и Warburton (1967), обсуждавшиеся принципы применимы к тестам побуждений потому что, например, тест аутизма (принципы 2 и 3 в табл. 4.1.) может быть использован для измерения желаемого, принимаемого за действительное применительно к большому разнообразию объектов: к пище (эрг голода), к женщинам (половой эрг) или к решению задач (эрг любопытства). Здесь все зависит, как и всегда при конструировании тестов, от изобретательности разработчика тестов. Однако, специфические свойства каждого конкретного теста оказывают влияние на то, какие аспекты побудительно-мотивируемого поведения будут измеряться. Из этого следует, что для того, чтобы измерить весь диапазон мотивов индивидуума необходим широкий выбор тестов (включающих столько принципов из табл.4.1., сколько возможно), измеряющих адекватный набор его аттитюдов, чем обеспечивается полный охват всего разнообразия чувств и эргов.

ФОРМИРОВАНИЕ НАБОРА АТТИТЮДОВ

Вы, надеюсь, заметили, что в предыдущем параграфе было использовано выражение "адекватный набор аттитюдов". Это было сделано потому, что явно невозможно достичь идеала – охватить все аттитюды испытуемого. При таком положении дел в процессе конструирования объективных тестов важно определить, какие аттитюды должны войти в набор. Cattell и Warburton (1967) утверждают, что это может быть сделано первоначально *a priori* продуманным отбором аттитюдов, о которых можно судить, что они, вероятно, относятся к наиболее важным побуждениям (напр., голод и половое влечение), а затем позже эмпирически дополнять другими аттитюдами (для которых побуждения лишь интуитивно угадываются либо даже неизвестны) в процессе исследований с переменными – маркерами повторяющихся побуждений. Эта проблема вновь указывает на значительные трудности, с которыми сталкивается разработчик тестов при изучении мотивов.

В исследованиях Кэттелла и его коллег (описанных в Cattell и Warburton, 1967; Cattell и Child, 1975) подвергнуты изучению около шестидесяти аттитюдов, каждый из которых, похоже, насыщен как эргом, так и чувством. Это поясняется нижеследующими примерами. (Тем не менее, прежде чем будут описаны эти примеры, следует осознать, что этот подход имеет ограничения: наши гипотезы могут быть неверными, и можно не заметить чрезвычайно важные переменные.) Например, (1) Я хочу быть таким человеком, которого каждому хотелось бы видеть рядом с собой. Это выявление эрга стадности (потребности в обществе, в общении) и чувства по отношению к самому себе, самоощущения. (2) Я хочу пойти домой и отдохнуть после трудного дня. Этим аттитюдом измеряется эрг стремления к отдыху и чувство к своему дому. (3) Я хочу видеть, что смертность из-за несчастных случаев и болезней сократилась. Здесь мы имеем дело с выявлением эрга страха.

Примеры объективных тестов

О конструировании объективных тестов сказано уже достаточно. Были приведены все принципы и методические указания, и единственное, что остается – это дать несколько примеров. Я выбрал их из "Сборника объективных тестов" Cattell и Warburton (1967). Они были отобраны, чтобы показать как изобретательность, с которой были разработаны тесты, так и разнообразие тех, что сконструированы к настоящему времени. Кроме того, выбор указывает и на трудности при разработке объективных тестов. Вот названия некоторых из 400 объективных тестов, разработанных к настоящему времени:

- Желание подшучивать действием
- Готовность принимать быстрые решения в процессе адаптации к темноте
- Амплитуда голоса при нормальных условиях по сравнению с условиями задержанной обратной связи (искусственно создаваемой отсрочкой в поступлении звука собственного голоса)
- Знание общественного этикета
- Базовая скорость метаболизма
- Эйдетическое воображение
- Вычеркивание букв (задание на внимательность), сравниваемое в двух ситуациях

- Готовность имитировать звуки, издаваемые животными
- Критическая частота слияния мигающих кадров
- Скорость появления негативных последовательных образов
- Предпочтение раскрашивать цветными карандашами свои собственные, а не предъявляемые рисунки
- Частота тремора в решающей ситуации
- Продолжительность смеха над шутками
- Расширение зрачка после испуга
- Большее количество беспокойных движений при ожидании, измеренное фиджетометром
- Скорость копирования изображения (эскиза, чертежа)
- Высота башни из кубиков (для детей 6 лет)
- Внимательность при следовании подробным инструкциям
- Точность в завершении гештальта
- Расстояние, которое проходит палец испытуемого по тесту-лабиринту, при электрошоке и без него

Эти названия дают прекрасное представление о богатейшем разнообразии разработанных тестов. Сейчас некоторые из них будут описаны более подробно. Это высветит существующие проблемы и сделает понятным, почему я не советую предпринимать легкомысленные попытки разработки таких тестов.

ЖЕЛАНИЕ ПОДШУЧИВАТЬ ДЕЙСТВИЕМ

Это объективный тест из разряда опросников. Испытуемый выражает свое желание подшучивать действием. Основанием для разработки этого теста послужило предположение о том, что застенчивые испытуемые должны быть расположены к таким невербальным шуткам. Факторные исследования подтвердили это предположение, а кроме того оказалось, что и уверенным в себе, не проявляющим застенчивости испытуемым также нравятся эти действия. Последнее – неожиданное открытие.

Этот тест показывает сложности проектирования объективных тестов, поскольку следовало бы ожидать, что этот тест будет насыщен фактором экстраверсии, так как экстравертам нравятся такие шутки, а интровертам нет.

ГОТОВНОСТЬ ПРИНИМАТЬ БЫСТРЫЕ РЕШЕНИЯ В ПРОЦЕССЕ АДАПТАЦИИ К ТЕМНОТЕ

При проведении этого теста испытуемый находится в темной комнате. Затем включается яркий свет, и ему предлагают смотреть на белый экран. Потом говорят, что когда свет выключится, он увидит букву. Испытуемый должен назвать букву, как только ее увидит. Используются три буквы. Измеряемой переменной является время, которое проходит до того, как буква будет названа. Основанием для разработки этого теста явилось предположение о том, что испытуемые с доминированием процесса торможения должны демонстрировать более быструю адаптацию к темноте.

Этот тест не сложен, не требует дорогостоящего оборудования, но он пригоден в основном для индивидуального применения, хотя возможна и групповая форма с использованием оборудования лингвистической лаборатории. Kline и Gale (1969) показали, что с помощью такой процедуры можно удачно использовать проективный тест "Картинки Блэки" (Blurn, 1949) для группового тестирования (без потери элемента неразглашения тайны, связанного с индивидуальным тестированием и существенного при проведении данного объективного теста).

БАЗОВАЯ СКОРОСТЬ МЕТАБОЛИЗМА

В этом тесте величина минимального потребления испытуемым кислорода за шесть минут преобразуется в калории/час на квадратный метр поверхности тела. Первоначально полагали, что этот тест будет иметь отношение к витальности (*exuberance*), реактивности, мобилизационной силеорганизма и экстраверсии. Действительно, он насыщен первыми тремя факторами, но также – фактором застенчивости и непрактичности.

Как вынужденно свидетельствуют Cattell и Warburton (1967), совсем не просто теоретизировать с какой-либо степенью точности и ответственности о психологических аспектах физиологических функций.

БЕСПОКОЙНЫЕ ДВИЖЕНИЯ, ИЗМЕРЕННЫЕ ФИДЖЕТОМЕТРОМ

Фиджетометр – это вращающееся кресло с электрическими контактами в различных местах, которые замыкаются при движениях. Показателем является общее количество движений, зарегистрированных в течение фиксированного промежутка времени. Конструкция этого кресла такова, что испытуемые не замечают в ней ничего необычного. Основанием для разработки этого теста явилось предположение о том, что тревожные (*anxious*) люди должны совершать больше суетливых, беспокойных движений, как и те, у кого высокая степень базальной тревоги. На деле же исследования так ни разу и не показали зависимости между показателями по этому тесту и фактором тревожности, хотя для базальной тревоги это имело место.

Фиджетометр – это хороший пример объективного теста с той точки зрения, что его показатели трудно фальсифицировать, и в то же время можно легко и точно обрабатывать. Он также с очевидностью показывает проблемы, существующие для объективных тестов, так как подходит только для индивидуального использования и, несмотря на изобретательность, с которой он был реализован (хотя в действительности это электрифицированная модель прибора, изобретенного еще Гальтоном), он лишь незначительно насыщен только одним фактором. Такой тест должен был бы быть лучше!

На этих примерах в полной мере показаны опасности и трудности, возникающие на пути объективного тестирования. Было уже сказано достаточно, чтобы позволить разработчику тестов продвигаться далее, но я должен опять повторить свое предупреждение: этот путь не для слабодушных. Разумеется, самые лучшие тесты личности – независимые от культурных особенностей испытуемых, объективные, те, получаемые с помощью которых показатели трудно исказить. Объективные тесты удовлетворяют этим требованиям. Однако, чтобы получить такой тест, необходимо приложить огромные исследовательские усилия.