

© 2005 г.

Н. Н. НОХРИНА

ТЕСТ КАК ОБЩЕНАУЧНЫЙ ДИАГНОСТИЧЕСКИЙ МЕТОД

НОХРИНА Наталья Николаевна - кандидат педагогических наук, доцент Южно-Уральского государственного университета (г. Златоуст).

Временем возникновения теста как общенаучного диагностического метода считается конец XIX века, когда для изучения индивидуальных различий начинают применяться различные способы измерения. Английский психолог Ф. Гальтон, воспользовавшись работами бельгийского статистика А. Кетле, разработал метод статического сравнения двух рядов переменных и ввел индекс совместного отношения, названный коэффициентом корреляции. Важным вкладом Гальтона в развитие теории тестов было определение трех основных принципов: применение серии одинаковых испытаний к большому количеству испытуемых; статистическая обработка результатов; выделение эталонов оценки. Все современные тесты построены на основе статистической теории измерений, а идея эталона оценки лежит в основе определения теста как стандартизированного инструмента.

Значителен вклад американского психолога Дж. Кеттелла в развитие идеи статистического анализа при определении индивидуальных различий. Считая тест средством для проведения научного эксперимента, он выделил ряд требований к его чистоте [1, с. 19]. Эти идеи в настоящее время составляют основу современной тестологии. Так, требование "одинаковости условий для всех испытуемых" положено в основу принципа стандартизации процедуры проведения тестирования; "ограничение времени" используется в зависимости от целей тестирования; "отсутствие зрителей в лаборатории" реализуется как учет влияния внешних факторов; идея "статистической обработки результатов" отражается в методах статистического анализа.

С начала XX века наметилось педагогическое направление в развитии тестологии. Американец В.А. Макколл разделил тесты на педагогические (Educational Test) и психологические (Intelligence Test). Основной задачей педагогических тестов являлось измерение успешности учащихся по тем или иным школьным дисциплинам за определенный период обучения. Макколл обосновал цель использования педагогических тестов - объединение в группы учащихся, усваивающих равный по объему материал с одинаковой скоростью [2, с. 4]. Однако основоположником педагогических измерений считается американский психолог Э. Торндайк. Ему принадлежит разработка первого педагогического теста и книга "Введение в теорию психологии и социальных измерений" (1904). С 1915 по 1930 гг. в Америке социальные измерения нашли широкое распространение, и этот период характеризуется как настоящий бум в развитии тестологии.

Распространение психолого-педагогической диагностики в России связано с возникновением в 20-е годы XX века новой науки - педологии, которую Л.С. Выготский [3] определял как науку о целостном развитии ребенка. Психологические службы в школах того периода, в силу особенностей понимания предмета педологии, "принимали комплексный характер, соединяя в себе черты психологической, социологической и медико-биологической служб одновременно" [4, с. 101].

Начало 30-х годов прошлого века стало временем массового использования тестов и, к глубокому сожалению, их неконтролируемого применения в народном образовании. Мнимая простота определения уровня интеллекта с помощью тестирования, легкость получения результатов и их толкование обусловили широкое использование тестов непрофессионалами. Массо-

вые тестовые обследования не подкреплялись серьезной проверкой качества инструментария [5, с. 29], тесты умственной одаренности были слабо адаптированы к местным условиям [6, с. 23]. Принятое в 1936 г. постановление ЦК ВКП(б) "О педологических извращениях в системе Наркомпроса" [7] охарактеризовало педологию как "враждебную марксизму лженауку" и наложило запрет на применение тестов и анкет. Метод тестирования был признан буржуазным орудием дискриминации учащихся и "изгнан" из советской школы. В дальнейшем в течение многих лет тестирование официально не признавалось объективным методом и не находило практического применения. По мнению ведущих советских психологов А.Н. Леонтьева, А.Р. Лурии, А.А. Смирнова [8], это постановление получило неправомерно расширенное толкование и послужило основой для свертывания исследований в области разработки диагностических методов и отказа от систематической психолого-педагогической диагностики.

Примечательна история происхождения понятия "тест". Оно произошло из старофранцузского языка и имеет синоним "чашка" (лат. *testa* - ваза из глины). Этим термином обозначали небольшие сосуды из обожженной глины, используемые алхимиками для проведения опытов (см.: [9, с. 11]). В русском языке в XIX веке термин существовал в двух значениях: испытательная присяга, английская религиозная клятва, которую каждый вступающий в общественную должность должен давать, чтобы доказать, что он не тайный католик; плоский плавильный сосуд для выделения олова из золота и серебра (согласно Русскому энциклопедическому словарю 1872 года, ссылка по: [10, с. 158]). Большинство современных исследователей рассматривают происхождение этого слова от английского (*test* - испытание, исследование). Ряд авторов называют тест специальным или особым видом экспериментального метода [11, с. 21]. Иногда тест трактуют как испытательный эксперимент, используемый наряду с лабораторным и естественным экспериментами [12, с. 21]. Подчеркнем, что все методы исследования представляют собой систему, в которой взаимодействуют экспериментальные и неэкспериментальные методы. Экспериментальность метода определяется сущностью и целью исследования, а не его названием.

Анализ содержания определений теста позволяет сделать вывод, что при всем внешнем разнообразии многочисленных определений, понятие "тест" рассматривается в трех аспектах: во-первых, как инструмент теоретического исследования; во-вторых, как метод эмпирического исследования; в-третьих, как средство, предназначенное не только для распознавания качеств личности, но и для получения объективной количественной оценки.

В литературе, изданной после 1990-го года, встречается понятие "неформальный тест", под которым понимается метод диагностики, конструируемый исследователем с целью фиксации результатов планируемого им эксперимента. Отличает неформальные тесты от методик, разработанных профессиональными тестологами, отсутствие точности определения надежности и валидности теста. Не владея элементарными знаниями и методиками классической теории тестов, непрофессионалы неизбежно допускают серьезные ошибки при их конструировании. Недостаточное знание понятийно-терминологического аппарата приводит к тому, что одно понятие употребляется в различных смыслах. Так, понятие "тест" довольно часто не дифференцируется по отношению к понятию "тестовое задание", а последнее, в свою очередь, употребляется в смысле "вопрос".

Тест - это инструмент, краткое стандартизованное испытание, в основе которого лежит специально подготовленный набор заданий, позволяющих объективно и надежно оценить исследуемые качества на основе использования статистических методов. Структурными компонентами теста являются: *тестовое задание* - задача в тестовой форме, предназначенная для выполнения, к которой помимо содержания предъявляются требования тестовой формы, и *эталон* - образец полного и правильного выполнения действия, служащий для сравнения достигнутого уровня с планируемым.

Поскольку тестовое задание - это некоторая задача, которую приходится решать испытуемому, то необходимо усвоить трактовку последней. Под *задачей* понимается сформулированное для испытуемого задание по достижению определенной цели в известных условиях ранее изученными методами деятельности [13, с. 78].

При разработке и использовании уже готовых тестов важно знать, насколько они соответствуют запроектированным целям. Ответ на этот вопрос дают критерии качества теста, на основании которых оценивается эффективность любого теста. Важнейшими из таких показателей являются *валидность* и *надежность*.

Валидность теста выступает в качестве важнейшего критерия его доброкачественности и определяется как его *адекватность* и *действенность*. В наиболее общей формулировке валидность рассматривается как характеристика теста, указывающая на то, что тест измеряет

и насколько хорошо он это делает. Исходя из этого, тест считается валидным, если он измеряет то, для чего он предназначен.

Параллельно имеет место и другое определение валидности. Под последней понимается комплексная характеристика теста, включающая сведения об области исследуемых явлений и репрезентативности диагностической процедуры по отношению к ним [14, с. 31]. В.С. Аванесов [15, с. 187] употребляет термин "валидность" в смысле пригодности тестовых результатов для той цели, ради которой проводилось тестирование. Однако толкует его достаточно широко - это и качество заданий, и их число в тесте, и глубина охвата содержания, и баланс распределения заданий в тесте по степени трудности.

На наш взгляд, и по мнению ряда других исследователей, "пригодный" - наиболее адекватное в русском языке значение этого термина. Так, В.М. Кадневский отмечает: "Это, пожалуй, более близкий к сути аналог зарубежного термина" [16, с. 195]. Важно отметить, что термин "валидность" связан по многим параметрам с областью исследования и задачами диагностирования. Как справедливо замечает В.И. Тесленко: "Тест нельзя назвать валидным или невалидным без указания сферы его применения" [17, с. 192].

Рассматривая понятие "валидность", необходимо обратить внимание на еще один существенный момент - относительный характер валидности. По мнению А.Н. Майорова, в общем виде понятие валидности "раскрывает, насколько полученные результаты тестирования соответствуют объективной реальности, а поскольку отсутствует способ точного определения объективной реальности, то валидность - понятие относительное, так как тесты, как и любой измерительный инструмент, никогда не являются абсолютно эквивалентными реальному положению вещей. Соответственно, валидность устанавливается относительно характеристики, признака, величины, принимаемой (а не истинно являющейся) за объективную" [9, с. 72].

В современной тестологии выделяются несколько видов валидности, однако их толкование, как и количество в разных исследованиях неоднозначно. Не считая необходимым приводить характеристики всех видов валидности, остановимся более подробно на тех, которые имеют, на наш взгляд, наиболее существенное значение.

Очевидная валидность - внешнее представление о тесте с точки зрения испытуемого. Она определяет, насколько адекватно впечатление о предмете измерения, которое формируется у испытуемого при знакомстве с инструкцией и материалом теста, тому, что данный инструмент в действительности выявляет. Чем более тест выглядит соответствующим предлагаемой цели измерения, тем выше его валидность. Эффективным методом повышения очевидной валидности является привлечение испытуемых к анализу теста, что позволяет взглянуть на тест глазами тех, для кого он предназначен.

Такая валидность может быть присуща тесту, а может и не быть. Это зависит от целей исследования. Для большинства тестов очевидная валидность - позитивное либо нейтральное качество, поскольку, как правило, нет оснований скрывать от испытуемых реальные цели тестирования.

Содержательная валидность - основной показатель, характеризующий степень репрезентативности содержания заданий теста по отношению к измеряемой характеристике. Она определяется на основе экспертных методов. Наибольшее распространение нашли такие способы экспертной оценки, как методы коллективной оценки, средневзвешенной оценки, ранжирования и парного сравнения (см., [9; 17]).

Одним из удобных и широко применяемых на практике способов оценки содержательной валидности является метод расчета коэффициента корреляции без привлечения независимых экспертов. Для этой цели в экспериментальной группе проводят тестирование по разработанному тесту, а затем выявленные результаты сравнивают с оценками, полученными традиционным способом. После чего рассчитывается коэффициент корреляции. Если он составляет 0,9 и более, то валидность теста считается высокой, если же менее 0,7 - низкой.

Функциональная валидность - показатель теста, близкий к содержательной валидности, но имеющий свои особенности. Он характеризует степень соответствия содержания заданий теста по отношению к выявляемому уровню усвоения опыта деятельности. Задания, предлагаемые в тесте, должны быть решены соответствующими определенному уровню способами деятельности и не могут быть выполнены путем использования способов деятельности более низкого уровня. Функциональная валидность может быть присуща тесту только в том случае, когда при составлении теста заранее планируется уровень его сложности.

Прогностическая валидность - характеристика теста, отражающая степень точности и обоснованности суждения о диагностируемой переменной. Для установления прогностиче-

ской валидности изучаются корреляции между данными, полученными в ходе проведенного исследования и спустя определенное время.

Критериальная валидность - вид валидности, направленный на выбор критерия, по которому можно объективно оценивать результаты работы испытуемых с тестами. Вопрос о выборе критерия является в этом виде валидности основным. Критерий валидности теста - признак его обоснованности, правомерности применения как измерительного инструмента. Важнейшей характеристикой критерия валидности считается объективность, которая заключается в его независимости от субъективных мнений.

Валидизация теста по критерию состоит в сравнении результатов тестирования с данными по критерию и вычислении коэффициента корреляции тестового результата с внешним критерием. В качестве внешнего критерия обычно используется привычная пятибалльная шкала оценки и традиционные отметки, выставляемые экспертом без использования теста. После чего результаты тестирования и оценки экспертов коррелируются. При получении согласованности между оценками экспертов и результатами тестирования считается, что тест обладает валидностью. Степень согласованности указывает на меру валидности.

Вторым критерием качества теста является его *надежность*, показывающая, обеспечивает ли тест как измерительный инструмент объективность результатов. Использование ненадежных тестов, допуск большого количества ошибок в таком ответственном деле, каким является тестирование людей, может стать причиной педагогических и административных ошибок, последствия которых трудно исправить [9, с. 163].

Термин надежность можно рассматривать в двух значениях. В широком смысле он понимается как характеристика того, в какой степени обнаруженные у испытуемых различия по тестовым результатам являются отражением действительных различий в измеряемых свойствах и в какой мере они могут быть приписаны случайным ошибкам. В узком смысле, под ним подразумевается степень согласованности результатов теста, получаемых при первичном и повторном его применении, по отношению к тем же испытуемым в различные моменты времени с использованием разных (но сопоставимых по характеру) наборов тестовых заданий или при иных изменениях условий обследования.

В исследованиях, проводимых в сфере педагогики, под надежностью понимается характеристика теста, отражающая точность диагностических измерений, а также устойчивость результатов теста к действию посторонних случайных факторов. Степень надежности теста определяется с помощью коэффициента надежности (коэффициента корреляции - r), который характеризует устойчивость показателей при повторных измерениях с помощью того же теста или его равноценного варианта.

В зависимости от условий, влияющих на результаты теста, выделяют ряд типов характеристик надежности.

Ретестовая надежность. Этот показатель определяется путем повторного обследования испытуемых с помощью одного и того же теста через некоторый промежуток времени. Он вычисляется по соответствию результатов первого и второго обследований или по сохранению ранговых мест испытуемых в выборке при ретесте, т.е. выявляет характер корреляции между результатами двух проверок. Коэффициент надежности соответствует коэффициенту корреляции между полученными результатами.

На практике ретестовая надежность определяется следующим образом. Результаты первичного тестирования в порядке возрастания распределяют на шкале тестирования. Эти данные принимают за истинный показатель. Теоретически это место для каждого участника должно быть постоянным, т.е. повторное тестирование должно дать такое же распределение мест на шкале, как и в первый раз. В этом случае тест считается абсолютно надежным. Полученные отклонения являются основой для расчета коэффициента надежности. Величина коэффициента надежности определяется величиной изменения оценок испытуемых на шкале в результате вторичного тестирования по сравнению с первичным. Следовательно, чем выше точность измерения первичного и вторичного тестирования, тем выше надежность теста как измерительного инструмента.

Метод повторного тестирования обладает как достоинствами, так и недостатками. К числу достоинств относятся естественность и простота определения коэффициента надежности. Основным недостатком метода является неопределенность в выборе интервала между двумя тестированиями. Исследователи рекомендуют проводить повторное тестирование в интервале от двух недель до шести месяцев. Кроме того, при повторном тестировании зачастую очень сложно, а порой и вовсе невозможно создать идентичные с первым обследованием условия. Это связа-

но с тем, что при всех, казалось бы, равных условиях, повторное испытание всегда отличается от первого.

Надежность параллельных форм. Данный показатель определяется с помощью взаимозаменяемых форм теста. При этом одни и те же испытуемые обследуются вначале с использованием основного набора заданий, а затем с применением аналогичного варианта. Коэффициент корреляции между результатами первичного и повторного тестирования - основной показатель данной характеристики надежности. Параллельными считаются такие тесты, которые измеряют одно и то же свойство с одинаковыми ошибками.

Надежность частей теста. Суть этого метода сводится к выполнению испытуемым заданий двух равноценных частей теста. Он возник как результат развития метода параллельного тестирования и основывается на предположении о параллельности не только форм теста, но и отдельных заданий внутри теста. Метод расщепления применяется, как правило, в тех случаях, когда повторное тестирование нежелательно, невозможно или существует только одна форма теста.

Этот показатель выявляют путем анализа устойчивости результатов отдельных совокупностей тестовых задач или единичных заданий теста. При этом распределение оценок за выполнение всего теста в целом должно быть близко распределению оценок при выполнении отдельных частей теста. Важнейшим условием объективности этого метода является деление теста на равноценные части (по числу, сложности задач, времени решения).

В целом при определении показателей качества теста следует помнить, что существует ряд объективных и субъективных факторов, влияющих на надежность теста. Рассмотрим их более подробно.

Содержательное разнообразие тестовых заданий — чем оно выше, тем ниже надежность теста.

Степень трудности тестовых заданий. Трудность задания определяется долями выборки, т.е. процентом испытуемых, давших правильный и неправильный ответы на вопрос при экспериментальном тестировании. Основным показателем трудности задач является *индекс трудности*. Для повышения надежности теста задания, которые имеют значение показателя индекса трудности больше 80 или меньше 20, в тест не включаются. Наибольшую практическую ценность имеют задания теста, на которые правильно отвечают 40-60% испытуемых. Однако следует иметь в виду, что трудность задачи - относительная характеристика, поскольку зависит от особенностей выборки и может меняться от выборки к выборке.

Величина теста. Надежность теста зависит от суммы существенных операций теста и повышается при увеличении их числа. Следовательно, чем длиннее тест, тем он надежнее.

Случайное угадывание правильных ответов - снижает надежность теста. Проблема угадывания ответов стоит особенно остро для альтернативных заданий. В случае с заданиями множественного выбора и достаточно большим количеством заданий в тесте, случайным угадыванием можно пренебречь.

Пропущенные задания. Наличие в тесте заданий, на которые испытуемые не дали ответов, в значительной степени искажает показатель надежности. Для повышения надежности перед подсчетом показателей по тесту, как советует П. Клайн [18], исследователю необходимо самостоятельно ответить на все пропущенные вопросы произвольным образом, что эквивалентно случайному угадыванию. В этом случае надежность теста будет искажена в меньшей степени. Проблема пропущенных заданий связана, как правило, с рядом факторов: временем, отвлекающим на тестирование; незнанием правильного ответа; невнимательностью испытуемых. Количество пропущенных заданий увеличивается в том случае, когда используются тесты с ограничением времени, называемые тестами скорости (методика, при которой практически ни одному из испытуемых не удается выполнить все задания). Если же 95% испытуемых успевают попробовать ответить на каждое задание, то такой тест относится к контролю без ограничения [19, с. 134]. С целью исключения пропусков по незнанию правильных ответов следует ввести в инструкцию требование "угадать ответ", а с целью привлечения внимания перед началом тестирования и в его конце озвучить требование "не пропускать ответы".

Источники внешнего и личностного характера. Тестирование, организованное с целью научного исследования, предполагает обязательный учет как внешних, так и личностно-психологических факторов, поскольку объективность результатов тестирования определяется не только технологией обработки информации, полученной в процессе диагностики, но и во многом зависит от оснащения, подготовки и процедуры проведения тестирования, а также от ряда субъективно-психологических факторов.

Главное требование к организации тестирования, как справедливо отмечает В.И. Тесеенко, обеспечение стандартизации работы с тестом испытуемых. "Только выполнение этого неперенного условия дает возможность для получения надежных результатов. При тестировании все должно быть одинаково для любого из испытуемых. Весь процесс измерений должен быть описан детально, и все предписания выполняться пунктуально, исключая любые факторы, которые могли бы существенно повлиять на результаты выполнения теста испытуемыми" [17, с. 242].

Решение вопроса оснащения тестирования, прежде всего, связано с целью диагностики, видом теста и количеством участников. В психологических, социологических и педагогических исследованиях, как при индивидуальном, так и групповом тестировании, чаще всего используются бланковые тесты, поскольку процедура тестирования в этом случае достаточно проста. Использование бланковых тестов, обеспечивая большую экономичность процедуры и массовость выборки, в то же время обеспечивает стандартизацию процедуры массовых обследований.

С целью получения достоверных результатов обследования исследователю следует особое внимание уделить комплексу бланковых тестов, в который в зависимости от поставленных задач могут входить следующие бланки: *бланк теста*, состоящий из тестовых заданий; *рабочая таблица результатов*, содержащая ответы испытуемых и по сути являющаяся протоколом; *эталон к тесту*, представляющий собой набор правильных ответов к тестовым заданиям; *шкала оценки результатов обследования*, позволяющая полученные результаты соотнести с нормами или критериями; *регистрационный индивидуальный бланк результатов*, в который последовательно заносятся сведения, полученные в результате обработки; *аспектная таблица оценки результатов*, отражающая сведения об успехах и неудачах испытуемого в определенных видах деятельности; *регистрационный групповой бланк*, содержащий все индивидуальные результаты группы в целом; *таблица отклонений индивидуальных показателей*, содержание которой позволяет осуществить ранжирование обследуемых по результатам тестирования.

Следует отметить, что все бланки должны создаваться с учетом основного требования - форма бланка не должна затруднять работу с ними. Обязательным условием здесь является соблюдение точности его формы - для ряда методик незначительные изменения внешнего вида бланка могут существенно изменить результаты работы. Следовательно, требования по заполнению бланков должны быть отражены в инструкции - обязательной составляющей всех стандартизованных тестов.

Вне зависимости от особенностей создания теста (создан заново, представляет собой модификацию, адаптацию имеющегося) *инструкция* составляется к каждому из тестов. При составлении инструкции рекомендуется соблюдать следующий порядок: написание первого варианта инструкции; первое опытное проведение теста; устранение грубых ошибок в инструкции; апробация инструкции при проведении проверочного тестирования; окончательная редакция инструкции [9, с. 230].

Надо учитывать тот факт, что практика проведения тестирования в вашей стране небогата, и у потенциальных испытуемых нет опыта работы с тестами. Это в свою очередь влечет за собой случайные ошибки в их работе. Поэтому основным этапом процедуры проведения тестирования является *подготовка испытуемых* к работе с тестом. Речь идет об инструктаже. Среди его целей выделяются следующие; формирование у испытуемых представлений о целях тестирования; формирование определенных мотивов, отношения к предстоящей деятельности; ознакомление со специфическими особенностями конкретного теста, с которым предстоит иметь дело [17, с. 243].

Процесс ознакомления с заданиями теста при групповом обследовании может быть организован разными способами: в форме как коллективного обсуждения заданий, так и индивидуальной работы с печатным пособием. От выбора способа ознакомления испытуемых с правилами работы в процессе тестирования зависит уровень их самостоятельности, количество случайных ошибок и темп работы.

На результаты теста во многом влияют *внешние условия* - физические, психологические и технологические. При определении этого влияния учитываются характеристики помещения, его оснащение, наличие отвлекающих факторов, состояние тестовых материалов, субъективно-психологические аспекты.

Непосредственная подготовка к проведению тестирования заключается в проверке состояния помещения, его пригодности для размещения определенного числа испытуемых, а также устранении или уменьшении ситуативных отвлекающих факторов. В идеале следует учитывать такие характеристики помещения, как его удельный объем на одного человека, освещенность, вентилируемость, уровень шума, акустические характеристики. Однако на практике вы-

полнение таких требований чрезвычайно сложно, поскольку степень их соответствия без требуемого оборудования и специалистов оценена быть не может.

При проведении группового теста для формальной стороны процедуры обязательно выполнение ряда требований: обеспечение инвентарием в необходимых количествах; наличие достаточного количества столиков и стульев; их размещение таким образом, чтобы было удобно подойти к каждому испытуемому; обеспечение каждому участнику удобного места; оборудование места экспериментатора с максимальным обзором [9, с. 233].

Особое значение среди факторов, влияющих на работоспособность испытуемых, занимает время проведения теста и длительность его протекания. Считается, что наиболее благоприятное время - период с 9 до 12 или с 16 до 18 часов.

Вопрос о длительности проведения тестирования, как правило, решается исходя из особенностей протекания познавательного процесса - внимания, или, точнее, одного из его свойств - устойчивости. Под устойчивостью внимания [20, с. 107] понимается продолжительность времени, в течение которого человек может намеренно поддерживать свое внимание на объекте. Психологи, благодаря проведенным опытам, пришли к заключению о том, что интенсивное внимание может сохраняться произвольно, без заметного ослабления и непроизвольного переключения, в течение сорокаминутного интервала (хотя это и вызывает утомление). В дальнейшем интенсивность внимания снижается тем быстрее, чем более однообразна работа. Следовательно, планируя тестирование, процедуру проведения целесообразно ограничить рамками одного часа.

При организации обстановки для проведения тестирования важно учитывать ситуативные отвлекающие факторы, такие как различного рода шумы, запахи, состояние помещения и т.д., поскольку устойчивость внимания связана не только с длительностью времени выполнения определенных видов деятельности, но и с внешними обстоятельствами. При обстановке, в которой отсутствуют внешние отвлекающие моменты, внимание, как отмечают психологи [21, с. 173], бывает достаточно устойчивым. При наличии множества сильных отвлекающих раздражителей оно колеблется, становится недостаточно устойчивым.

В ряду причин, определяющих устойчивость внимания, при организации тестирования следует учитывать факторы, связанные с *индивидуальными физиологическими особенностями* человека: свойствами нервной системы, общим и психическим состоянием испытуемого в данный момент. Люди со слабой нервной системой или перевозбужденные могут довольно быстро утомляться, становятся импульсивными. Человек, который не очень хорошо чувствует себя физически, также, как правило, характеризуется неустойчивостью внимания [21, с. 173]. Поэтому, планируя тестирование, важно предусмотреть не только время его проведения и продолжительность, но и характер деятельности обследуемых до тестирования. Перед его началом экспериментатору необходимо оценить состояние исследуемых. Практика показывает, что состояние возбуждения или заторможенности можно почти безошибочно определить визуально. Обследуемых, у которых наблюдается проявление таких состояний, не следует допускать к тестированию. Однако это возможно только при индивидуальной работе. При групповом тестировании принятие таких мер весьма затруднительно, а порой и совершенно невозможно. В подобной ситуации на первый план выходит деятельность экспериментатора, направленная на нормализацию состояния испытуемых.

Ослабить влияние подобного рода факторов на результаты тестирования позволяет умение исследователя сконцентрировать внимание обследуемых на предмете деятельности, показать значимость последней для личности каждого испытуемого. Речь идет о мотивации, ведущей к повышению устойчивости внимания. Однако следует помнить, что и слишком сильная и слишком слабая мотивировка в одинаковой степени негативно сказываются на результативности выполнения заданий.

При организации тестирования важна разработка *сценария*, отражающего последовательность его проведения (см. [9, с. 234]). Обобщив информацию и используя собственный многолетний опыт работы с тестами, приведем порядок процедур тестирования. 1. Обозначить цель тестирования, сообщить, какие результаты ожидаются; 2. Объяснить, почему обследуемые должны приложить максимум усилий для выполнения теста; 3. Медленно, громко, четко озвучить инструкцию по работе с тестом; 4. Разобрать задания-образцы с целью проверки правильности понимания инструкции; 5. Объяснить правила заполнения бланка, предназначенного для ответов, и правила исправления допущенных ошибок; 6. Вместе с испытуемыми заполнить в бланке рабочей таблицы необходимые общие сведения, проследив за правильностью заполнения; 7. Ответить на имеющиеся вопросы; 8. Сообщить о временном ресурсе; 9. Дать команду начать решение задач теста, попросив записать время начала на бланке рабочей таблицы;

10. В процессе решения тестовых задач следить за состоянием испытуемых; за тем, чтобы испытуемые своевременно получали ответы на вопросы в ситуациях, связанных с процедурой проведения; за временем, затраченным на выполнение отдельных заданий, и требуемым для оставшихся задач; за тем, чтобы обследуемые не писали на бланках теста, не мешали друг другу; чтобы их внимание не отвлекали внешние факторы; 11. После сигнала к окончанию выполнения теста попросить испытуемых записать время окончания на бланке рабочей таблицы; 12. Собрать бланки; 13. Пересчитать бланки, проверив, чтобы их количество совпадало с числом испытуемых; 14. При выявленном несоответствии обратиться к испытуемым, чтобы они проверили свои рабочие места и сдали недостающие бланки; 15. По окончании тестирования поблагодарить испытуемых за работу и сообщить о том, что все свободны.

Заключительный этап в организации тестирования - вербализация либо визуализация результатов тестирования. *Обсуждение* полученных результатов в зависимости от его целей может существенно варьировать от краткого сообщения испытуемым общих результатов до всестороннего анализа процесса тестирования. Организуя обсуждение результатов тестирования, следует придерживаться ряда принципиальных положений: проводить анализ результатов тестирования оперативно; этот анализ должен носить диагностический характер, раскрывая главное - причины возникновения ошибок испытуемых; быть эффективным средством корректирования; оказывать положительное влияние на испытуемых, способствовать их совершенствованию.

Эксперимент может стать источником психологических травм испытуемых, однако совершенно очевидно и то, что главный источник травмирования - недостатки в организации эксперимента и обсуждении результатов. Важнейшие правила организации эксперимента - недопустимость риска для здоровья и развития испытуемых, гарантии от нанесения вреда их самочувствию, от ущерба для жизнедеятельности в настоящем и будущем (см.: [22, с. 43]).

СПИСОК ЛИТЕРАТУРЫ

1. *Аванесов В.С.* Тесты в социологическом исследовании. М., 1982.
2. *Цатурова И. А.* Из истории развития тестов в СССР и за рубежом. Таганрог, 1969.
3. *Выготский Л.С.* Педология и психотехника // Психотехника и психофизиология труда. 1931. № 2-3.
4. *Баранов В.Ф.* Педологическая служба в советской школе 20-30-х гг. // Вопросы психологии. 1991. №4.
5. *Нечаев АЛ.* Современная экспериментальная психология в ее отношении к вопросам школьного обучения. М., 1901.
6. *Бернштейн М.С.* Вопросы методологии на Первом Всесоюзном педагогическом съезде // На путях к новой школе. 1928. № I.
7. Постановление ЦК ВКП(б) "О педологических извращениях в системе наркомпросов от 4 июля 1936 г." // Народное образование в СССР. Общеобразовательная школа: Сб. док. 1917-1973. М., 1974.
8. *Леонтьев А.Н., Лурия А.Р., Смирнов АЛ.* О диагностических методах психологического исследования школьников // Советская педагогика. 1968. № 7.
9. *Майоров А.Н.* Тесты школьных достижений: конструирование, проведение, использование. СПб., 1997.
10. *Бурлачук Л.Ф., Морозов С.Н.* Словарь-справочник по психологической диагностике. Киев, 1989.
11. *Крутецкий В. А.* Психология. М., 1986.
12. *Возрастная и педагогическая психология / Под ред. М.В. Гамезо.* М., 1984.
13. *Беспалько В.П.* Педагогика и прогрессивные технологии обучения. М., 1995.
14. Энциклопедия психологических тестов. Личность, мотивация, потребность. М., 1997.
15. *Аванесов В.С.* Композиция тестовых заданий. М., 1996.
16. *Кадневский В.М.* К проблеме формирования основных понятий тестологии // Методология и методика формирования научных понятий у учащихся школ и студентов вузов: Сб. матер, конф. Челябинск, 2001. Ч. 1.
17. *Тесленко В. И.* Теоретико-методические основы диагностики и прогнозирования процесса обучения будущего учителя физики в педвузе: Дис.... докт. пед. наук. Красноярск, 1996.
18. *Клайн П.* Справочное руководство по конструированию тестов. Введение в психометрическое проектирование: Пер. с англ. Киев, 1994.
19. *Аванесов В.С.* Научные проблемы тестового контроля знаний. М., 1994.

20. *Платонов К.К., Голубев Г.Г* Психология. М., 1973.
21. *Немов Р.С.* Психология, Кн. 1. Общие основы психологии. М., 1995.
22. Педагогика / Под ред. *П.И. Пидкасистого*. М., 1996.