

ПСИХОМЕТРИЧЕСКИЕ  
ОСНОВЫ  
ПСИХОДИАГНОСТИКИ

**КОНСТРУИРОВАНИЕ ТЕСТОВ**

# Этапы конструирования методик

---

Разработка нового психологического инструмента начинается с того, что на основе наблюдений за поведением людей, сбора тех поведенческих жизненных явлений, которые требуют объяснения и диагностики, автор создаваемого теста должен выдвинуть гипотезу о том:

- **что** лежит в основе тех или иных особенностей человеческого поведения,
- **почему** существуют различия в поведении разных людей
- **какая** именно особенность личности определяет эти различия.

# 1 этап конструирования методик

---

1. **название вновь вводимого гипотетического конструкта - определение предмета диагностики вновь создаваемого теста**
  - ***Теоретическое определение*** - это включение предлагаемого понятия в уже существующую систему научных понятий в данной области психологии.
  - ***Дать теоретическое определение*** – это значит прочно связать предлагаемый гипотетический конструкт с уже существующими и доказавшими свое право на существование конструктами и понятиями.

# Этапы конструирования методик

---

- **Операциональное определение** – это описание тех поведенческих проявлений, реакций и действий, чувств и оценок посредством которых данная переменная реализуется.
- **Операциональное определение** выводит изучаемую переменную на уровень фиксируемых, наблюдаемых и измеряемых проявлений и фактов, что делает ее доступной изучению и анализу.
- **Операциональное определение** является непосредственной основой для построения первичной формы создаваемого теста.

# Для операционализации гипотетического конструкта обычно используют

- **1. Поведенческие реакции, действия, поступки.** Гипотетический конструкт черты личности «общительность» может проявляться в количестве знакомых людей, в активности человека в ситуации общения, в степени его разговорчивости. *Я обычно первым начинаю разговор с соседями по купе в поезде».*
- **2. Мотивационные компоненты, проявляющиеся в предпочтениях, желаниях и побуждениях.** Общительность может проявляться в желании человека бывать в компаниях людей, в предпочтениях людских собраний и т.д. *«Я предпочитаю проводить свой отпуск на шумных курортах, нежели в уединенной хижине в горах»*
- **3. Эмоциональные проявления,** выражающиеся в получении положительных эмоций в ситуациях, где появляется данная личностная черта или наоборот. *«Одиночество меня пугает».*
- **4. Оценочные проявления,** выражающиеся в самооценке своих личностных особенностей или в их оценке со стороны других людей. *«Я – общительный человек» или «Меня считают общительным человеком».*

## 2 этап – составление заданий теста

### Проблемы (трудности) при разработке (формулировании) заданий личностных опросников

- **Установка на согласие** – тенденция испытуемого соглашаться с утверждениями или отвечать на вопросы «да» независимо от их содержания. Чаще всего проявляется, когда утверждения неоднозначны, неопределенны.
- **Установка на социально одобряемые ответы** – тенденция испытуемых отвечать на вопросы так, чтобы выглядеть «социально положительным»: если возможен «социально желательный» ответ, то весьма вероятно, что испытуемые будут его давать.
- **Установка на неопределенные или средние ответы**. Если в опроснике представлена средняя категория ответов, определяющая неуверенность в ответе (не уверен. Не знаю, затрудняюсь ответить), то многие испытуемые склонны к ней прибегать.
- **Установка на крайние ответы (расположенные по краям шкалы)**. При использовании многоэлементной шкалы ответов, некоторые испытуемые независимо от содержания предпочитают выбирать крайние ответы. (крайние значения – низкая валидность).

## 3 этап – анализ заданий теста

---

- первичная форма теста разработана и проведена на специально подобранной выборке
- определение **диагностической пригодности каждого из заданий методики.**
- каждое задание психологического теста должно быть проверено по нескольким психометрическим показателям

# Дискриминативность заданий теста

- Определение диагностической пригодности каждого из заданий методики.
- Коэффициент корреляции между ответами испытуемого на каждый отдельный пункт опросника и итоговым баллом по той шкале, в которую входит данный пункт должен быть статистически значимым, поскольку в этом случае данный коэффициент корреляции отражает тот факт, что *данное конкретное утверждение способно измерять тот же самый конструкт, что измеряет вся шкала в целом, в которую, естественно, входит данный пункт.*



# Коэффициент дискриминативности

- Как и любой коэффициент корреляции, может принимать значение от -1 до +1.
- *Значимый положительный коэффициент дискриминативности* свидетельствует о том, что данное задание направлено на измерение того же параметра личности, который измеряет вся шкала в целом.
- Значимое отрицательное значение данного коэффициента далеко не всегда говорит о том, что такое задание подлежит немедленному исключению из опросника, поскольку так называемые «обратные задания» должны иметь именно отрицательные и значимые коэффициенты дискриминативности
- Если коэффициент дискриминативности не является значимым, это говорит о том, что данное задание не выполняет свою диагностическую задачу, поскольку измеряемое им свойство личности никак не связано с тем свойством личности, которое измеряет вся шкала в целом.

# Индекс трудности

- используется в отношении тестов достижений и интеллектуальных тестов
- трудность задания отражает процентное соотношение лиц, решивших или не решивших данное задание
- *если задание выполняется 100 процентами испытуемых, трудность задания будет равна 0, если задание не выполняется ни одним испытуемым, то трудность данного задания будет равна 100.*
- **индекса трудности может колебаться от 0 до 1**
- если индекс трудности заданий слишком высок, то это означает, что наибольшая часть испытуемых не может правильно решить данное задание, поэтому оно не является диагностичным.
- если индекс трудности слишком низок, это означает, что подавляющее большинство испытуемых правильно решают данное задание, что в свою очередь, также говорит о его диагностической непригодности.

# Индекс трудности

- *В тестах, диагностирующих не интеллектуальные, а личностные особенности индекс трудности также отражает процентное соотношение лиц, ответивших в соответствии с ключом.*
- В личностных тестах ответы в соответствии с «ключом» определяют выраженность у испытуемого того параметра личности, который является диагностическим конструктом данного теста.
- Существует диапазон приемлемых значений индекса трудности, который находится в пределах от **0.16 до 0.84**, или в процентном соотношении, **в соответствии с ключом на данный пункт должны отвечать не менее 16-ти процентов и не более 84-х процентов испытуемых.**
- **Наиболее удачными с точки зрения трудности являются задания, индекс трудности которых составляет 0.5.**
- *Чем выше будет доля лиц, ответивших в соответствии с «ключом», тем ниже будет индекс трудности.*

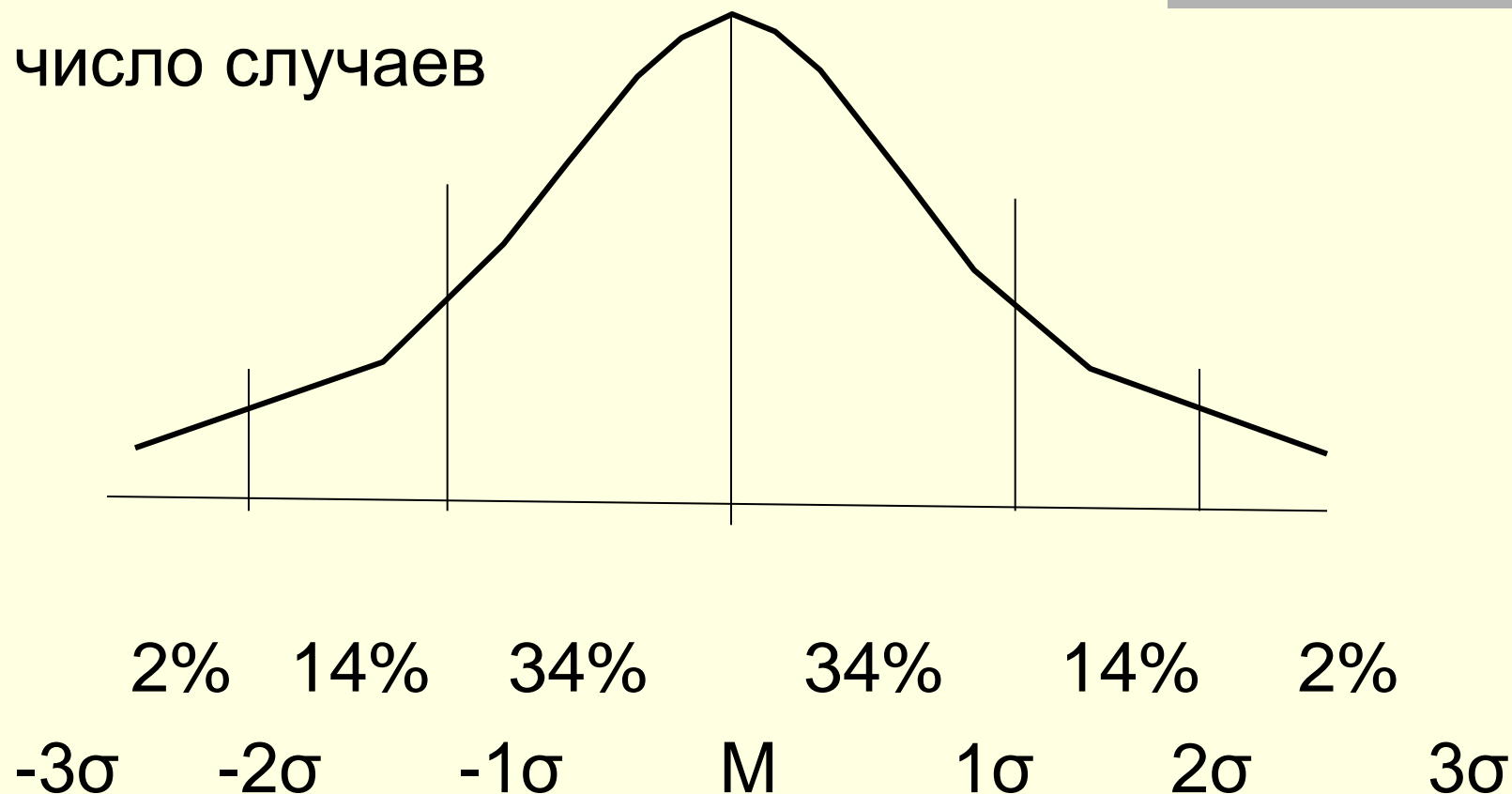
# Характер распределения тестовых оценок

---

- В окончательную форму теста должны попасть только те задания, тестовые оценки по которым подчиняются закону нормального распределения, т.е. *наибольшее количество испытуемых выбирает средние варианты ответа, в то время, как самые крайние варианты ответов выбираются наименьшим количеством испытуемых.*
- если характер распределения тестовых оценок не соответствует закону нормального распределения, можно говорить о том, что данное задание теста не является достаточно диагностичным и провоцирует большинство испытуемых давать варианты ответов, относящиеся к крайним полюсам шкалы ответов.

# Нормальность распределения

число случаев



## 4 этап - Проверка теста на надежность

---

- Одним из важных аспектов проблемы измерения в психологии является вопрос о степени доверия к его точности и стабильности.
- ***Надежность теста - характеристика того, в какой степени выявленные у испытуемых различия по тестовым результатам являются отражением действительных различий в измеряемых свойствах, а в каком они могут быть приписаны действию случайных факторов.***

# Надежность как устойчивость измерений

- демонстрирует **стабильность результатов во времени**, т.е. измерение, произведенное одним и тем же измерительным инструментом в разное время, дает одинаковые результаты. При этом результаты измерения зависят как от самой измерительной процедуры, так и от условий измерения.
- представляет собой согласованность результатов теста, получаемых при первичном и вторичном его применении по отношению к одним тем же испытуемым в разные моменты времени.
- **Коэффициентом надежности** в этом случае будет коэффициент корреляции между результатами первого и повторного тестирования, определяемый по формуле К. Пирсона.
- В связи с использованием процедуры повторного тестирования одной и той же выборки данный вид надежности получил в психодиагностике название **ретестовой надежности**.
- **оптимальный интервал для ретестовой надежности – 6 месяцев для взрослых и 3 месяца для детей.**

# Надежность как устойчивость измерений

---

- **Надежность параллельных форм** - использование двух взаимозаменяемых форм теста, обследование выборки сначала основным набором заданий, а, затем, - дополнительным, аналогичным первому.
- **ИЛИ разбиение всей выборки на две примерно равные группы.** Первичное тестирование заключается в тестировании первой группы одной формой теста (А), а второй группы – аналогичной формой (В). После небольшого интервала времени (примерно около недели) проводится повторное тестирование в обратном порядке.



# Надежность как точность измерений

- предполагает наличие внутренней согласованности между заданиями теста и их «суммарную работу на общую цель - точное измерение тестируемой характеристики». Влияние условий эксперимента в этом случае невелико.
- Процедура проверки предполагает вычисление *интеркорреляционных связей между заданиями, составляющими тест.*
- **метод расщепления** - тестирование испытуемого двумя половинками теста.
- коэффициентом надежности теста в данном случае будет являться коэффициент корреляции между частями теста.
- Данный способ определения надежности как внутренней согласованности часто называют **надежность по однородности.**

# Надежность относительно постоянства условий - надежность экспериментатора

Три функции: *организатор исследования, оценщик и интерпретатор.*

- **Постоянство экспериментатора как организатора исследования** предполагает, что его личностные особенности, социальное положение и другие характеристики не оказывают существенного влияния на обследуемых лиц и, соответственно, на результаты тестирования. **Достигается это за счет стандартизации тестового материала, инструкций, условий проведения тестирования.**
- **Постоянство подхода оценщика** касается ведения протоколов и подсчета результатов теста. **При наличии стандартных бланков ответов, ключей объективность оценки повышается, что приводит и к повышению надежности.**
- **Постоянство подхода экспериментатора как интерпретатора** обеспечивается, если одинаковые оценочные данные, полученные от испытуемых, приводят к одинаковым заключениям. **В психометрике такое положение обеспечивается наличием норм, стандартных профилей оценки или критериев.**

# Коэффициент надежности

- Наиболее популярными способами является **проверка ретестовой надежности и надежности по однородности.**
- Коэффициент надежности как коэффициент корреляции может изменяться в пределах от -1 до + 1, однако тест считается надежным, когда его коэффициент находится в пределах **0,6 -1.**
- для большинства известных методик редко удается получить коэффициент надежности выше 0,7- 0,8.
- **выборки для вычисления надежности должны быть достаточно большими.** Минимальный объем выборки для расчета надежности должен быть не меньше **150-200 человек.** Качественный состав выборки должен полностью отражать **генеральную совокупность.**

- **Генеральная совокупность** - множество элементов, объединенных общей характеристикой, указывающей на их принадлежность к определенной системе. Примером генеральной совокупности может служить все взрослое русскоязычное население нашей страны.
- Однако на практике охватить тестированием всю генеральную совокупность не возможно, поэтому из этой совокупности извлекается некая ее **репрезентативная (представительная) часть**, получившая название выборка.
- **Выборка** - это часть элементов генеральной совокупности, являющаяся репрезентативной по отношению к ней, но имеющая меньший объем. Качественный и количественный состав выборки определяется характером и природой изучаемого явления и определяется его универсальностью и однородностью генеральной совокупности.

# 5 этап – проверка теста на валидность

---

- это обязательная и наиболее важная часть информации о психологическом тесте
- указывает нам, на то, **что тест измеряет и насколько хорошо он это делает**

*Все вопросы, относящиеся к валидности, могут быть сведены к двум основным:*

- 1) какие выводы можно сделать о том, что измерялось тестом? – **внутренняя валидность**
- 2) какие выводы могут быть сделаны о реальном поведении человека на основе тестовых оценок? – **внешняя валидность.**

# Виды валидности

---

- **Критериальная валидность** - отражает соответствие диагноза или прогноза, определяемого на основе тестовых оценок определенному кругу жизненных критериев измеряемого качества.
- валидность измеряется путем корреляции баллов по тесту и показателей по внешнему параметру, избранному в качестве критерия, при этом это должен быть именно внешний, т.е. жизненный, независимый от тестирования критерий.
- делится на ***текущую и прогностическую***.

# Виды валидности

- **Конструктивная валидность** - показывает, насколько результаты теста могут рассматриваться в качестве меры некоего теоретического конструкта или свойства.

**Проверка конструктивной валидности может осуществляться следующими способами:**

- Сопоставление исследуемого теста с другими методиками, конструктивное содержание которых известно. Положительная корреляция указывает на то, что разрабатываемый тест измеряет примерно ту же сферу поведения, способность, психологическое качество, что и эталонная методика.
- В случаях когда трудно найти какой-либо апробированный тест на измеряемое свойство, валидность которого была бы известна, исследователь формулирует ряд предсказательных гипотез о том, как будет коррелировать его новый тест с другими тестами, измеряющими родственные характеристики.
- Внутренняя согласованность пунктов теста также является важным аспектом конструктивной валидности

# Виды валидности

---

- **Содержательная валидность** - означает систематическую проверку содержания теста с тем, чтобы установить, соответствует ли оно выборке измеряемой области поведения.
- Характеризует степень репрезентативности содержания заданий теста измеряемой области психических свойств.
- Такая валидность обычно применяется для тестов, исследующих деятельность, близкую или совпадающую с реальной, т.е. для тестов достижений, критериально-ориентированных тестов, а также методик, предназначенных для профотбора, анализа успешности овладения профессией и т.п.
- Для валидации личностных опросников и тестов интеллекта содержательная валидность используется ограниченно, лишь на начальных стадиях составления теста.



# 6 этап - СТАНДАРТИЗАЦИЯ

- процесс унификации, регламентации, приведения к единым нормативам процедуры психодиагностики и тестовых показателей.

*Две формы стандартизации.*

1. Обработка данных, регламентация процедуры проведения тестирования и его периодичности, унификация инструкции и бланков, способов регистрации результатов, стандартность условий проведения обследования, характеристик контингента обследуемых и т. д.
2. В результате обработки тестовых показателей получают сырые баллы, которые переводятся в стандартные оценки, составленные в соответствии с законом нормального распределения.

# Этапы стандартизации

## 1. создание единообразной процедуры

### *тестирования*

---

- 1) Условия тестирования (помещение, освещение и др. внешние факторы).
- 2) Содержание инструкции и особенности ее предъявления (тон голоса, паузы, скорость речи и т.д.). Например, в тесте "10 слов" каждое слово должно предъявляться через определенный интервал времени в секундах.
- 3) Наличие стандартного стимульного материала. Например, достоверность полученных результатов существенно зависит от того, предлагаются ли респонденту изготовленные самодельные карты Г. Роршаха или стандартные - с определенной цветовой гаммой и цветовыми оттенками.
- 4) Временные ограничения выполнения данного теста. Например, для выполнения теста Равена взрослому респонденту дается 20 минут.

# Этапы стандартизации

## 1. создание единообразной процедуры тестирования

---

- 5) Стандартный бланк для выполнения данного теста. Использование стандартного бланка облегчает процедуру обработки.
- 6) Учет влияния ситуационных переменных на процесс и результат тестирования. Под переменными подразумевается состояние испытуемого (усталость, перенапряжение и т.д.), нестандартные условия тестирования (плохое освещение, отсутствие вентиляции и др.), прерывание тестирования.
- 7) Учет влияния поведения диагноста на процесс и результат тестирования. Например, одобрительно-поощряющее поведение экспериментатора во время тестирования может восприниматься респондентом как подсказка "правильного ответа" и др.
- 8) Учет влияния опыта респондента в тестировании. Естественно, что респондент, который уже не в первый раз проходит процедуру тестирования, преодолел чувство неизвестности и выработал определенное отношение к тестовой ситуации.

# Этапы стандартизации

- **2. создание единообразной оценки выполнения теста** стандартной интерпретации полученных результатов и предварительной стандартной обработки. Этот этап предполагает также сравнение полученных показателей с нормой выполнения этого теста для данного возраста (например, в тестах интеллекта), пола и т.д.
- **3. определение норм выполнения теста.** Нормы необходимы при интерпретации тестовых результатов (первичных показателей) в качестве эталона, с которым сравниваются результаты тестирования.