## Лекция 4. Психометрические основы современной психодиагностики. Проблема и способы измерения в психологии.

1.
2.

## Психометрические основы психодиагностики

Краткий конспект

***Показатель в психологии*** – наблюдаемый, измеряемый или вычисляемый признак, позволяющий выявить наличие или величину какого-либо психического явления (состояния, представления, отношения, способности и т.п.)

Виды показателей по содержанию:

1. Физические

время реакции, пороги чувствительности, сила или амплитуда движений и т.п.

1. Демографические

пол, возраст, образование, род занятий и т.п.

1. Антропометрические

рост, вес, объем легких, цвет глаз или волос и т.п.

1. Физиологические

частота сердечных сокращений, дыхания, артериальное давление, электрическое сопротивление кожи, выраженность альфа-ритма ЭЭГ и т.п.

1. Биохимические

количество адреналина или сахара в крови или моче т.п.

1. Поведенческие

скорость и точность решения различных задач, частота определенных ответов на вопросы, частота тех или иных действий и т.п.

Виды показателей по источникам данных (Р.Кеттелл):

1. ***L-данные*** (life data) – данные реальной жизни (демографические, биографические) и наблюдения
2. ***T-данные*** (test data) – данные объективных тестов (интеллектуальных, проективных), антропометрические и физиологические данные
3. ***Q-данные*** (questionnaire data) – данные опросников, анкет, беседы

Подробный конспект

В результате наблюдения, опроса, тестирования и других методов психодиагностики мы получаем значения различных показателей. *Показатель* – наблюдаемый, измеряемый или вычисляемый признак, позволяющий выявить наличие или величину какого-либо явления. Показатели могут характеризовать различные психические явления: состояния, представления, отношения, способности и т.п. В психологии могут использоваться различные по содержанию показатели:

* физические (время реакции, пороги чувствительности, сила или амплитуда движений и т.п.),
* демографические (пол, возраст, образование, род занятий, семейное положение, национальная принадлежность и т.п.),
* антропометрические (рост, вес, тип телосложения и т.п.),
* физиологические (частота сердечных сокращений, частота дыхания, артериальное давление, электрическое сопротивление кожи, выраженность альфа-ритма ЭЭГ и т.п.),
* биохимические (количество адреналина или сахара в крови или моче и т.п.),
* поведенческие (скорость и точность решения различных задач, частота определенных ответов на вопросы, частота тех или иных действий и т.п.).

Основным способом получения значений показателей является измерение. *Измерение* – способ определения значения показателя с помощью установления его отношения к некоторому образцу или эталону. Для измерения показателей используются шкалы. *Шкала* – система измерения, устанавливающая соответствие между значениями показателей и наличием или величиной признаков.

Психодиагностическое тестирование предполагает использование стандартных шкал. *Стандартизация тестовых шкал* – это способ перевода первичных значений показателей или сырых баллов тестов в производные стандартизированные шкальные оценки. В результате тестирования мы получаем такие первичные показатели, как время выполнения задания, количество правильно решенных задач в заданное время или число определенных ответов на вопросы. Первичные показатели преобразуют в стандартные шкалы в целях:

* приведения различных показателей к единой шкале для их сопоставимости,
* определения величины показателей относительно статистической нормы.

Наиболее распространенным способом стандартизации тестовых шкал является нормирование – представление значений показателей в долях среднеквадратичного (стандартного) отклонения от среднего арифметического значения:



где:

*St* – стандартное шкальное значение,

*x* – первичное значение теста (сырой балл),

*M* – среднее арифметическое значение показателя теста по данным нормативной выборки,

*σ* – среднеквадратичное (стандартное) отклонение показателя теста по данным нормативной выборки,

*Mas* – коэффициент масштабирования,

*Cen* – коэффициент центрирования.

Здесь первичное значение тестового показателя выражается в отклонении от среднего значения этого показателя в долях среднеквадратичного отклонения этого показателя. Среднее арифметическое и среднеквадратичное отклонение в качестве параметров распределения этого показателя получены по данным предварительного обследования нормативной выборки испытуемых и представлены в методическом руководстве по тесту. Однако отклонения от среднего в долях стандартного отклонения не вполне удобны для работы, поскольку могут иметь отрицательные и дробные значения. Для того чтобы привести их к положительным целым значениям, они умножаются на определенный коэффициент масштабирования, и к ним добавляется определенный коэффициент центрирования. В шкалах различных тестов используются разные коэффициенты масштабирования и центрирования. В таблице 1 представлены параметры наиболее распространенных стандартных тестовых шкал.

Таблица 1. Параметры распространенных тестовых шкал

|  |  |  |  |
| --- | --- | --- | --- |
| Шкалы | Коэффициент масштабирования | Коэффициент центрирования | Тесты |
| IQ | 15 | 100 | Интеллектуальные тесты Векслера и Амтхауэра |
| T-баллы | 10 | 50 | Личностные опросники MMPI и CPI |
| Стэны | 2 | 5,5 | 16-факторный личностный опросник Кеттелла |

Существует соответствие между параметрами нормально распределенных значений показателей (средним M и стандартным отклонением σ), процентилями этих значений, z-оценками, баллами IQ и T-баллами (Рис. 1).

*Процентиль* – это значение показателя, меньше которого в выборке встречается данное количество процентов значений. Так, например 10-й процентиль – это первичное значение, меньше которого в выборке встречается 10% значений, а 50-й процентиль – это значение, меньше которого встречается половина значений выборки. *Z- оценка* – это нормированное отклонение, то есть отклонение первичного значения от среднего в долях стандартного отклонения. Например, z-оценка, равная -1, означает, что соответствующее первичное значение меньше среднего на величину одного стандартного отклонения. Среднее арифметическое M любого нормально распределенного показателя равно 50-му процентилю и z-оценке, равной нулю. Среднее значение в шкале IQ равно 100, а в шкалеT-баллов равно 50. Диапазоном статистической нормы считается диапазон M ± 2σ, включающий примерно 95% всех значений. Таким образом, диапазон статистической нормы показателя в шкале IQ составляет от 70 до 130 баллов, а в шкале T-баллов – от 30 до 70 баллов.



Рис. 1. Стандартные тестовые шкалы

Американский психолог Стенли Смит Стивенс (1906 – 1973) разработал теорию уровней измерения и шкалирования, описывающую шкалы наименований, порядка, равных интервалов и равных отношений.

1. *Шкала наименований или номинальная шкала* характеризует наличие или отсутствие признака, устанавливает, равно или не равно одно значение показателя другому значению, но не говорит, какое значение больше или меньше; допускает любые преобразования, сохраняющие различие между значениями, в том числе, перемену мест значений. Примеры показателей в шкале наименований: пол, род занятий, название или тип профессии, название национальности, цвет кожи, тип телосложения, характер заболевания, тип темперамента, переживаемая эмоция, способ реагирования, стратегия поведения, правильность решения задачи, воспринимаемый цвет, воспринимаемая идентичность раздражителей.

2. *Шкала порядка или ранговая шкала* характеризует величину признака, но не имеет единицы измерения, устанавливает, больше или меньше одно значение показателя относительно другого значения, но не говорит, насколько; допускает любые монотонные преобразования, сохраняющие порядок значений, поэтому значения нельзя менять местами. Примеры показателей в шкале порядка: уровень образования, уровень должности, балльная оценка академической успеваемости, привлекательность чего-либо или кого-либо, сила эмоционального переживания, воспринимаемая яркость или громкость раздражителя, воспринимаемое неравенство раздражителей.

3. *Шкала равных интервалов или интервальная шкала* характеризует величину признака, имеет единицу измерения, но не имеет точки отсчета (нулевой точки, характеризующей отсутствие признака), устанавливает, на сколько единиц больше или меньше одно значение показателя относительно другого значения, но не говорит, во сколько раз; определяет, равна или не равна разность между одной парой значений разности между другой парой, допускает любые линейные преобразования, сохраняющие разности между значениями, в том числе, добавление ко всем значениям постоянной величины. Примеры показателей в шкале равных интервалов: уровень интеллекта в шкале IQ, степень выраженности черт личности в шкале T-баллов или шкале стэнов.

4. *Шкала равных отношений или абсолютная шкала* характеризует величину признака, имеет единицу измерения и точку отсчета (нулевую точку, обозначающую отсутствие признака), устанавливает, во сколько раз больше или меньше одно значение показателя относительно другого значения, определяет, равно или не равно отношение между одной парой значений отношению между другой парой. Примеры показателей в шкале равных отношений: время реакции, порог чувствительности, сила и амплитуда движений, возраст, рост, вес, частота сердечных сокращений, артериальное давление, количество сахара в крови, скорость решения различных задач, количество ошибок, частота определенных ответов на вопросы, частота тех или иных действий.

Можно установить соответствие между различными уровнями шкалирования, качественными и количественными показателями, неметрическими и метрическими показателями (Таблица 2).

Таблица 2. Виды показателей по типу шкал

|  |  |  |
| --- | --- | --- |
| По наличию или величине признака | По наличию единицы измерения | Тип шкалы по С.Стивенсу |
| Качественные показатели (наличие признака) | Неметрические показатели (отсутствие единицы измерения) | Шкала наименований |
| Количественные показатели (величина признака) | Шкала порядка |
| Метрические показатели (наличие единицы измерения) | Шкала равных интервалов |
| Шкала равных отношений |

Кроме того, показатели мы можем разделить на дискретные и континуальные, прямые и косвенные, объективные и субъективные.

Количественные показатели делятся на дискретные и континуальные:

* дискретные (прерывистые) показатели – число явлений, обладающих определенным признаком, выраженное натуральным числом, например, число лиц определенного пола, число определенных ответов на вопрос, число ошибочных двигательных реакций и т.п.,
* континуальные (непрерывные) показатели – число условных единиц определенного признака, выраженное действительным числом, например, возраст, рост, время двигательной реакции и т.п.

По степени непосредственности показатели делятся на прямые и косвенные:

* прямые показатели – основаны на использовании в качестве эталона признака той же природы, например, измерение расстояния в километрах, измерение роста, веса, силы, времени реакции в соответствующих единицах измерения,
* косвенные показатели – основаны на использовании в качестве эталона признака другой природы, например, измерение расстояния в днях пути, измерение скорости мышления в количестве правильно решенных задач в единицу времени, измерение черты личности в количестве определенных ответов на вопросы и т.п.

Наконец, показатели делятся на объективные и субъективные:

* объективные показатели – основаны на использовании в качестве эталонов объектов или единиц измерения, существующих независимо от сознания, например, расстояние, длительность, сила, порог чувствительности, время реакции и т.п.,
* субъективные показатели – основаны на использовании в качестве эталонов представлений, существующих только в сознании, например, субъективные оценки расстояния, длительности, яркости, громкости, цвета, запаха, вкуса, привлекательности, интереса, важности, трудности, уверенности и т.п.

В процессе любого измерения показателей неизбежны ошибки. *Ошибки измерения* – отклонение измеренных значений показателей от действительных. Можно выделить систематические и случайные ошибки измерения.

* Систематические ошибки измерения не изменяются от измерения к измерению или изменяются предсказуемым образом, проявляются в регулярном смещении, сдвиге результатов в какую-то сторону на какую-то величину, обусловлены неслучайным изменением условий измерения.
* Случайные ошибки измерения изменяются от измерения к измерению непредсказуемым образом, проявляются в рассеивании, разбросе результатов в разные стороны на разную величину, обусловлены случайными изменениями различных условий измерения.

Особой категорией ошибок в практике психологических измерений и психодиагностики являются ошибки субъективных оценок, которые могут проявляться в эффектах социальной желательности, ореола, уровня, центрации и поляризации.

* Эффект социальной желательности проявляется в улучшении оценок качеств собственной личности или личности значимых людей, вызванное желанием произвести благоприятное впечатление и продемонстрировать соблюдение социальных норм.
* Эффект ореола (гало-эффект) – изменение оценок одних качеств личности под влиянием оценок других качеств, например, повышение оценок уровня интеллекта лиц, которые кажутся более внешне привлекательными, активными или исполнительными.
* Эффект уровня – изменение оценок качеств личности одних людей под влиянием оценок этих качеств у других людей, например, повышение оценок интеллектуальных способностей человека на фоне других людей, уровень интеллектуального развития которых кажется невысоким.
* Эффект центрации или поляризации – смещение оценок к центру или крайним значениям шкалы вследствие привычного стиля оценивания, установки на осторожные или категоричные оценки.

Для предотвращения ошибок субъективных оценок рекомендуется ранжирование человеком разных людей отдельно по величине каждого качества, усреднение оценок, полученных группой людей, или независимое измерение установок людей на искажение результатов (шкалы «лжи»).

Основным требованием к измерению является максимально возможная точность, минимизация систематических и случайных ошибок. Точность измерения характеризуется валидностью и надежностью.

*Валидность* – степень соответствия измеряемых показателей изучаемым психическим явлением, соответствия между тем, что должно измеряться, и тем, что измеряется в действительности; отсутствие ошибок, связанных с непригодностью средства измерения. Выделяются различные виды валидности: теоретическая – содержательная, и эмпирическая – конструктная и критериальная.

*Содержательная валидность* – соответствие содержания заданий представлению о реальной учебной или профессиональной деятельности, соответствие содержания вопросов представлению об измеряемом свойстве. Например, корректурная проба позволяет измерять концентрацию и устойчивость внимания, а скорость первого ответа и количество ошибок в тесте опознания сходных изображений Дж. Кагана являются валидными показателями импульсивности личности.

*Конструктная валидность* – соответствие результатов измерения представлению об измеряемом свойстве, его структуре или динамике; соответствие результатов измерения гипотезам о наличии корреляций со сходными показателями и отсутствии корреляций с несходными показателями. Например, ответы на ряд вопросов личностного опросника Г. Айзенка коррелируют между собой, образуя фактор экстраверсии, и не коррелируют с фактором нейротизма, а ответы на определенные вопросы личностного опросника Р. Кеттелла коррелируют между собой и группируются в 16 независимых факторов, представляющих черты личности.

*Критериальная валидность* – соответствие результатов измерения показателя результатам измерения этого свойства с помощью других методов, которые считаются достоверными, таких как демографические или биографические данные, успешность учебной или трудовой деятельности, врачебный диагноз, экспертные оценки, данные другого теста, валидность которого установлена. Например, показатель интеллекта теста Р. Амтхауэра коррелирует с академической успеваемостью, показатели клинических шкал личностного опросника MMPI коррелируют с психиатрическим диагнозом.

*Надежность* – степень соответствия значений измеряемого показателя значениям того же показателя при других измерениях в одинаковых условиях; отсутствие ошибок измерения, возникающих вследствие различных случайных факторов. Выделяют различные виды надежности: устойчивость (ретестовая надежность) и однородность (внутренняя согласованность, надежность параллельных форм или частей).

*Устойчивость* – корреляция между результатами повторного измерения тем же методом. *Однородность* – корреляция между результатами измерения эквивалентными формами теста или расщепленными половинами теста, например, включающими четные и нечетные задания или вопросы. Часто однородность теста оценивают с помощью коэффициента альфа Кронбаха, характеризующего средний коэффициент корреляции между пунктами теста. Чем ближе α к 1, тем выше точность теста, если α=0, результаты теста совершенно случайны.

В качестве наглядной метафоры отношения валидности и надежности измерения можно представить стрельбу по мишеням. Валидность означает, что стрелок попадает именно по той мишени, по которой нужно, а не по другим мишеням. В свою очередь, надежность означает, что при повторной стрельбе по какой-либо мишени разброс небольшой.

*Автор конспекта: И.Л. Соломин*

## Уровни измерения и шкалирования

**Измерение** – способ определения значения показателя с помощью установления его отношения к некоторому образцу или эталону.

**Шкала** – система измерения, устанавливающая соответствие между значениями показателей и наличием или величиной признаков

Стандартизация тестовых шкал

* Перевод первичных значений показателей или сырых баллов *(времени выполнения задания, количества правильно решенных задач в заданное время или числа определенных ответов на вопросы)* в производные с использованием параметров распределения значений показателей *(средние арифметические, среднеквадратичные отклонения, процентили)*, полученных в выборке стандартизации
* Используется в целях:
	+ Приведения различных показателей к единой шкале для их сопоставимости
	+ Определения величины показателей относительно статистической нормы

Нормирование тестовых шкал

* Представление значений показателей в долях среднеквадратичного отклонения от среднего арифметического значения
	+ St – стандартное значение
	+ x – первичное значение (сырой балл)
	+ M – среднее арифметическое значение показателя
	+ σ – среднеквадратичное отклонение показателя
	+ Mas – коэффициент масштабирования
	+ Cen – коэффициент центрирования



Параметры распространенных тестовых шкал

|  |  |  |  |
| --- | --- | --- | --- |
|  | Коэффициент масштабирования | Коэффициент центрирования | Тесты |
| IQ | 15 | 100 | WAIS, IST |
| T-баллы | 10 | 50 | MMPI, CPI |
| Стэны | 2 | 5,5 | 16PF |

Стандартные тестовые шкалы



Стэнли Смит Стивенс (1906-1973)

* Американский психолог
	+ Использовал прямые методы оценки ощущений в психофизике
	+ Разработал концепцию уровней измерения и шкалирования, включающую шкалы:
		- Наименований
		- Порядка
		- Равных интервалов
		- Равных отношений



Уровни измерения и шкал (С. Стивенс):

1. Номинальное измерение (шкала наименований)
2. Ординальнное измерение (шкала порядка или рангов)
3. Интервальное измерение (шкала равных интервалов)
4. Абсолютное измерение (шкала равных отношений)

1. Номинальное измерение и шкала наименований:

* Характеризует наличие или отсутствие признака
* Устанавливает, равно или не равно одно значение показателя другому значению, но не говорит, какое значение больше или меньше
* Допускает любые преобразования, сохраняющие различие между значениями, в том числе, замену значений местами
	+ *Примеры*: пол, род занятий, название или тип профессии, название национальности, цвет кожи, тип телосложения, характер заболевания, тип темперамента, переживаемая эмоция, способ реагирования, стратегия поведения, правильность решения задачи, воспринимаемый цвет, воспринимаемая идентичность раздражителей.

2. Ординальное измерение и шкала порядка:

* Характеризует величину признака, но не имеет единицы измерения
* Устанавливает, больше или меньше одно значение показателя относительно другого значения, но не говорит, насколько
* Допускает любые монотонные преобразования, сохраняющие порядок значений, поэтому значения нельзя менять местами
	+ *Примеры*: уровень образования, уровень должности, балльная оценка академической успеваемости, величина привлекательности, сила эмоционального переживания, воспринимаемое неравенство раздражителей.

3. Интервальное измерение и шкала равных интервалов:

* Характеризует величину признака, имеет единицу измерения, но не имеет точки отсчета (нулевой точки, характеризующей отсутствие признака)
* Устанавливает, на сколько больше или меньше одно значение показателя относительно другого значения, но не говорит, во сколько раз
* Определяет, равна или не равна разность между одной парой значений разности между другой парой
* Допускает любые линейные преобразования, сохраняющие разности между значениями, в том числе, добавление ко всем значениям постоянной величины
	+ *Примеры*: уровень интеллекта в шкале IQ, степень выраженности черт личности в шкале T-баллов или шкале стенов,

4. Абсолютное измерение и шкалы равных отношений:

* Характеризует величину признака, имеет единицу измерения и точку отсчета (нулевую точку, характеризующую отсутствие признака)
* Устанавливает, во сколько раз больше или меньше одно значение показателя относительно другого значения
* Определяет, равно или не равно отношение между одной парой значений отношению между другой парой
	+ *Примеры*: время реакции, порог чувствительности, сила и амплитуда движений, возраст, рост, вес, частота сердечных сокращений, артериальное давление, количество сахара в крови, скорость решения различных задач, количество ошибок, частота определенных ответов на вопросы, частота тех или иных действий

Виды показателей по уровню измерения и типу шкал:

|  |  |  |
| --- | --- | --- |
| По наличию или величине признака | По наличию единицы измерения | Название уровня измерения и типа шкалы по С.Стивенсу |
| Качественные показатели | Неметрические показатели | Номинальное измерениеШкала наименований |
| Количественные показатели | Ординальное измерениеШкала порядка или рангов |
| Метрические показатели | Интервальное измерениеШкала равных интервалов |
| Абсолютное измерениеШкала равных отношений |

Виды количественных показателей по непрерывности:

* **Дискретные показатели** – число явлений, обладающих определенным признаком, выраженное натуральным числом. Например:
	+ число лиц определенного пола, число определенных ответов на вопрос, число ошибочных двигательных реакций и т.п.
* **Континуальные показатели** – число условных единиц определенного признака, выраженное действительным числом. Например:
	+ возраст, рост, время двигательной реакции и т.п.

Виды показателей по степени непосредственности:

* **Прямые показатели** – основаны на использовании в качестве эталона признака той же природы. Примеры:
	+ Измерение расстояния в километрах
	+ Измерение роста, веса, силы, времени реакции в соответствующих единицах измерения
* **Косвенные показатели** – основаны на использовании в качестве эталона признака другой природы. Примеры:
	+ Измерение расстояния в днях пути
	+ Измерение скорости мышления количестве правильно решенных задач в единицу времени, черты личности в количестве определенных ответов на вопросы и т.п.

Виды показателей по степени объективности:

* **Объективные показатели** – основаны на использовании в качестве эталонов объектов или единиц измерения, существующих независимо от сознания. Например:
	+ расстояние, длительность, сила, порог чувствительности, время реакции и т.п.
* **Субъективные показатели** – основаны на использовании в качестве эталонов представлений, существующих только в сознании. Например:
	+ субъективные оценки расстояния, длительности, яркости, громкости, цвета, запаха, вкуса, привлекательности, интереса, важности, трудности, уверенности и т.п.

Ошибки измерения:

* Систематические ошибки
* Случайные ошибки
* Ошибки субъективных оценок

Систематические ошибки измерения:

* Не изменяются от измерения к измерению или изменяются предсказуемым образом
* Проявляются в смещении, сдвиге результатов
* Обусловлены неслучайным изменением условий измерения: состава испытуемых, их состояния, содержания заданий и т.п.

Случайные ошибки измерения:

* Изменяются от измерения к измерению непредсказуемым образом
* Проявляются в рассеивании результатов
* Обусловлены случайными изменениями различных условий измерения

Ошибки субъективных оценок:

1. **Эффект социальной желательности** – улучшение оценок качеств собственной личности или личности значимых людей, вызванное желанием произвести благоприятное впечатление и продемонстрировать соблюдение социальных норм
2. **Эффект ореола (гало-эффект)** - изменение оценок одних качеств личности под влиянием других качеств личности
* например, повышение оценок уровня интеллектуальных способностей лиц, которые кажутся более внешне привлекательными, активными или исполнительными
1. **Эффект уровня** – изменение оценок качеств личности одних людей под влиянием оценок этих качеств у других людей
* например, повышение оценок интеллектуальных способностей человека на фоне других людей, уровень интеллектуального развития которых кажется невысоким
1. **Эффект центрации или поляризации** – смещение оценок к центру или крайним значениям шкалы вследствие привычного стиля оценивания, установки на осторожные или категоричные оценки

Способы предотвращения ошибок субъективных оценок:

* Ранжирование человеком людей отдельно по величине каждого качества
* Усреднение оценок, полученных группой людей
* Независимое измерение установок людей на искажение результатов (шкалы «лжи»)
* Маскировка истинных целей исследования, неожиданная постановка задач, отвлечение внимания или повышение эмоциональной напряженности с целью провоцирования непроизвольных ответов

*Автор конспекта: И.Л. Соломин*