

КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ

ФАКУЛЬТЕТ ПСИХОЛОГИИ

КАФЕДРА ПРАКТИЧЕСКОЙ ПСИХОЛОГИИ

А.Ф. Корниенко

ПСИХОДИАГНОСТИКА

УЧЕБНОЕ ПОСОБИЕ



Казань 2003

УДК 159.9 (075.8)

ББК 88в7

К67

Печатается по рекомендации Редакционно-издательского совета Казанского государственного педагогического университета.

РЕЦЕНЗЕНТЫ:

Аболин Л.М. – докт. психол. наук, проф., декан фак-та педагогики и психологии КСЮИ

Попов Л.М. – докт. психол. наук, проф., зав. кафедры психологии КГУ

Корниенко А.Ф.

К67 Психодиагностика: Учебное пособие /Казанский пед. ун-т, Казань, 2003. –148 с.: ил.

ISBN 5-87730-034-2

Рассматриваются основные процедуры и проблемы психодиагностики, с которыми приходится сталкиваться при проведении психологических исследований. Особое внимание уделяется вопросам психометрии, возможностям использования в психодиагностике методов математической статистики, корреляционного, факторного и кластерного анализов. Раскрываются понятия объективности, валидности, надежности и дискриминативности психодиагностических процедур и методы их определения.

Книга предназначена для научных сотрудников, студентов и преподавателей психологических факультетов, а также аспирантов и практических психологов, интересующихся вопросами психодиагностики.

УДК 159.9 (075.8)

ББК 88в7

ISBN 5-87730-034-2

© КОРНИЕНКО А.Ф., 2003

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ	5
ГЛАВА 1. ПРЕДМЕТ И МЕТОДЫ ПСИХОДИАГНОСТИКИ	7
1.1. Понятие о психодиагностике	7
1.2. Истоки психодиагностики	11
1.3. Классификация психодиагностических методик	15
1.4. Этические аспекты психодиагностики	23
ГЛАВА 2. ИЗМЕРЕНИЕ В ПСИХОДИАГНОСТИКЕ	31
2.1. Понятие об измерении	31
2.2. Измерительные шкалы и их характеристика	32
□ шкала наименований	32
□ шкала порядка	34
□ шкала интервалов	36
□ шкала отношений	38
2.3. Виды измерений	40
2.4. Основные понятия и процедуры статистического анализа	41
□ понятие выборки	41
□ статистическая совокупность значений показателя	45
□ меры уровня (средние)	51
□ меры рассеяния	54
□ стандартизация значений показателя	56
□ понятие нормального распределения	60
□ нормализация значений показателя	64
□ понятие доверительного интервала	67
□ определение объема выборки	70
□ выявление различий в значениях показателей	71
□ определение достоверности различий	75
2.5. Методы корреляционного анализа	77
□ корреляция как мера статистической связи показателей	77
□ корреляция показателей в шкалах интервалов	80
- коэффициент корреляции Пирсона	80
- проверка значимости коэффициента корреляции	87
- множественный коэффициент корреляции	88

❑ корреляция показателей в шкалах порядка	89
- коэффициент ранговой корреляции Спирмена	90
- множественный коэффициент ранговой корреляции	94
❑ корреляция номинальных признаков	97
- коэффициенты ассоциации и сопряженности	98
- множественный коэффициент квадратичной сопряженности	100
❑ корреляция показателей в смешанных шкалах	104
❑ построение корреляционных плеяд	107
2.6. Понятие о факторном анализе	110
2.7. Кластерный анализ	113
ГЛАВА 3. ПОКАЗАТЕЛИ ЭФФЕКТИВНОСТИ ПСИХОДИАГНОСТИЧЕСКИХ ПРОЦЕДУР	119
3.1. Объективность	119
3.2. Валидность	123
❑ понятие валидности	123
❑ виды валидности и методы их определения	125
3.3. Надежность	128
❑ понятие надежности	128
❑ методы определения надежности	130
– метод ретеста	130
– метод параллельного тестирования	131
– метод расщепления	132
3.4. Дискриминативность	137
ЛИТЕРАТУРА	140
ПРИЛОЖЕНИЕ	142

ПРЕДИСЛОВИЕ

Обычно при изложении курса психодиагностики большое внимание уделяется изложению особенностей применения конкретных психодиагностических методик (тестов), позволяющих определять (диагностировать) те или иные особенности психических процессов, психических состояний, свойств личности или межличностных отношений. Фактически, рассматриваются вопросы техники проведения психодиагностического обследования, порядка обработки и интерпретации получаемых результатов.

Поскольку методик, используемых в настоящее время в психодиагностике, очень много, их невозможно изложить в одном учебном пособии. Впрочем, в этом и нет особой необходимости – многие из них заслуживают отдельного рассмотрения. В связи с этим авторам учебных пособий приходится ограничиваться либо подробным изложением сравнительно небольшой группы конкретных психодиагностических методик, либо обзором и изложением общих представлений о достаточно широком круге психодиагностических процедур.

В данном учебном пособии вовсе не ставилась цель рассмотреть какие-то отдельные психодиагностические процедуры. Основное внимание было уделено тому, что объединяет все психодиагностические процедуры, что есть в каждой из них и без чего ни одна из них не обходится – это измерение. Измерение присуще любой психодиагностической процедуре. По существу, психодиагностика – это и есть измерение, при котором возникают как минимум три вопроса: а) что измерять? б) зачем измерять? и в) как измерять? Если первые два вопроса определяют стратегию психодиагностики, то третий – ее тактику, от которой, как показывает практика, во многом зависит успех или эффективность диагностики.

К сожалению, многие психологи, применяя те или иные психодиагностические методики, порой не задумываются о том, какова погрешность

измерения, насколько можно доверять результатам, полученным с помощью данных методик, насколько объективны и надежны используемые психодиагностические процедуры, действительно ли они измеряют то, для диагностики чего их обычно используют. Освещению именно этих вопросов и посвящено данное учебное пособие.

Учебное пособие рассчитано, в первую очередь, на студентов и преподавателей, изучающих и читающих курс «Психодиагностика» на психологических факультетах. Содержание и логика построения учебного пособия соответствует требованиям Государственного образовательного стандарта высшего профессионального образования по специальности «Психология».

Август 2003 г.

ГЛАВА 1

ПРЕДМЕТ И МЕТОДЫ ПСИХОДИАГНОСТИКИ

1.1. Понятие о психодиагностике

Психологическая диагностика – одно из важнейших направлений в профессиональной деятельности психолога. Термин "психодиагностика" появился сначала в психиатрии с появлением работы Г. Роршаха "Психодиагностика" (1921). В основе этого термина лежит слово "диагностика", которое происходит от латинского "dia-gnosticos", что в буквальном смысле означает "различительное познание". Первоначальное использование этого термина в психологии предполагало определение различий между одними и другими людьми, между больными и здоровыми. В настоящее время термин "диагностика" широко используется в различных областях науки: и в медицине, и в психологии, и в педагогике, и в технике – есть медицинская диагностика, техническая диагностика, педагогическая диагностика и т.д. Во всех перечисленных случаях под "диагностикой" понимается некоторая процедура определения текущего состояния диагностируемого объекта (или системы) путем достаточно быстрой регистрации его существенных параметров (свойств, признаков) с целью оценки его работоспособности, прогноза его поведения и принятия решения о возможностях воздействия на это поведение в желательном направлении. Конечным результатом любого вида диагностики является постановка соответствующего диагноза.

Очевидно, что результатом психодиагностики является постановка психологического диагноза. Можно сказать, что психодиагностика – это наука и практика постановки психологического диагноза.

Психологический диагноз, являясь результатом психодиагностики, включает в себя описание индивидуально-психологических особенностей человека, характеризующих его актуальное состояние, прогноз дальнейшего развития его психики и личности, а также характеристику методов психологического воздействия, которые могут быть использованы в целях психокоррекции, психотерапии или психологического консультирования.

Другими словами, постановка психологического диагноза и, соответственно, психодиагностика включают в себя три основных момента:

а) установление наличия, индивидуальной степени выраженности, а также характера взаимосвязи наиболее существенных свойств и качеств

психики и личности человека, определяющих особенности его поведения и деятельности в сложившихся условиях (обстоятельствах);

б) прогноз дальнейшего развития индивидуально-психологических особенностей человека в существующих или других возможных условиях;

в) оценка методов психолого-педагогического воздействия на человека, которые могут быть использованы в целях нормализации функционирования его психики и личности.

Без психодиагностики невозможно ни проводить психокоррекцию и психотерапию, ни осуществлять на должном профессиональном уровне психологическое консультирование.

Основные функции психодиагностики можно представить в виде следующей схемы (см. Рис.1):

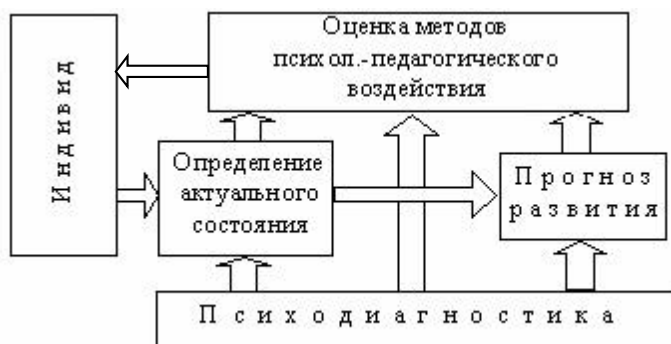


Рис.1.1. Функции психодиагностики.

Очевидно, что для проведения психодиагностики необходимы определенные условия, определенные методы и средства психодиагностики, которые обеспечивают оказание на индивида определенного психодиагностического воздействия. Данное воздействие необходимо для того, чтобы диагностируемое свойство или качество, присущее индивиду, могло как-то проявиться.

Психодиагностическое воздействия – обязательный атрибут психодиагностики, без которого, собственно, нет психодиагностики. Вместе с тем, наличие психодиагностического воздействия обуславливает возникновение в психодиагностике весьма существенной проблемы, суть которой состоит в том, что диагностируемое свойство в процессе психодиагностики может не просто проявляться, но и изменяться вследствие самого психодиагностического воздействия. В связи с этим по отношению к индивиду и его индивидуально-психологическим свойствам психодиагностика начинает выступать не только как психодиагностическая процедура, но и

как процедура психокоррекционная. Данное обстоятельство необходимо учитывать при интерпретации результатов психодиагностики, а также при разработке программы психодиагностики, чтобы избежать нежелательных последствий, которые могут возникать вследствие психокорректирующего эффекта психодиагностических воздействий.

Наличие в психодиагностике психокорректирующих эффектов можно отразить добавлением в схему, представленную на рис. 1.1, процессов психологического воздействия, направленных на индивида. В результате схема основных функций психодиагностики будет выглядеть следующим образом:

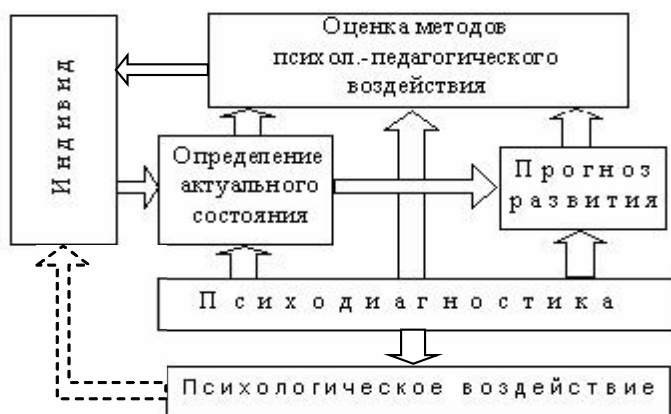


Рис. 1.2. Функции психодиагностики с учетом психокорректирующих эффектов.

Психодиагностика – не только одно из направлений практической деятельности психолога, но и область психологической науки, которая имеет дело с выделением, анализом и обобщением совокупности феноменологических признаков, а также качественных и количественных показателей, характеризующих диагностируемые свойства психики и личности.

Например, для диагностики такого свойства личности, как самооценка требуется выяснить, что это за свойство, в чем оно проявляется, что может служить показателем выраженности данного свойства и какова процедура его определения. В связи с этим можно говорить не только о практической, но и о теоретической психодиагностике, которая опирается на достижения различных отраслей психологии, прежде всего, общей пси-

хологии, психологии личности, возрастной и педагогической психологии, социальной психологии, психометрии и т. д.

Результатом работы психолога в области теоретической психодиагностики является психодиагностическая процедура (тест, методика), пригодная для профессионального использования широким кругом практических психологов. Очень часто тест или методика, разработанная для целей психодиагностики, носит имя автора-разработчика. Всем известны психологические тесты и методики психодиагностики Роршаха, Айзенка, Кеттелла, Люшера, Спилбергера, Розенцвейга, Томаса, Равена, Векслера, Торренса, Дембо-Рубинштейна, Лири, Морено, Рокича, Эйдемиллера и многих других. Тем самым подчеркивается тот факт, что в основе любой психодиагностической процедуры лежат представления и теоретические концепции его автора.

Эмпирические данные, получаемые в результате психодиагностики, могут быть использованы не только самим психодиагностом, но и другими специалистами и, вообще, другими лицами, в связи с чем в психодиагностике возникает ряд проблем, как профессионального, так и этического характера.

Обычно выделяют четыре варианта использования данных психодиагностики¹:

1. Данные, получаемые психодиагностом, сразу передаются другим специалистам для постановки ими не психологического диагноза. Например, в медицине, когда врач использует результаты психодиагностики для постановки медицинского диагноза и последующего медицинского воздействия, или в педагогике, когда педагог (учитель) на основе данных психодиагностики делает заключение о развитии учащихся, уровне их обученности и воспитанности.

2. Данные используются психодиагностом для постановки психологического диагноза, который затем передается другому специалисту для оказания определенного воздействия на субъект. Например, когда учитель (педагог) на основании заключения психодиагноста корректирует учебно-воспитательный процесс, оказывая соответствующие педагогические и психологические воздействия на учащихся, или в управлении персоналом на предприятии, когда на основании заключения психолога руко-

¹ Общая психодиагностика /Под ред. А.А. Бодалева, В.В. Столина.- М.: Изд-во МГУ, 1987.

водитель предприятия принимает определенные административные решения относительно того или иного работника.

3. Данные используются самим психодиагностом не только для постановки психологического диагноза, но и для разработки методов и средств последующего психологического воздействия на субъекта. Например, в ситуации психологического консультирования или психологической коррекции.

4. Данные, получаемые психодиагностом, передаются тому, кто подвергается диагностике, для принятия им самостоятельного решения относительно результатов диагностики, особенностей своего состояния и возможности оказания на него какого-либо психологического воздействия.

1.2. Истоки психодиагностики

Психологическая диагностика как специальная дисциплина психологии начала складываться на рубеже XX в. под воздействием требований практики. С другой стороны, ее возникновение было подготовлено достижениями и успехами в развитии экспериментальной психологии, дифференциальной психологии, прикладной психологии и тестологии².

В *экспериментальной психологии* в работе "Элементы психофизики" (1860) немецкого физика, философа и психолога Густава Фехнера, с именем которого, собственно, и связано возникновение экспериментальной психологии, впервые был обоснован психодиагностический метод исследования, ставший затем основой психодиагностики. В основу метода был положен методологический принцип, который состоит в утверждении, что "...психические реакции вызываются вариацией раздражителя". Варьируя интенсивность раздражителя, Фехнер фиксировал возникающие при этом ощущения и, в итоге сформулировал закон, согласно которому интенсивность ощущения (S), как психической реакции, возникающей в ответ на воздействие раздражителя, пропорциональна логарифму интенсивности (I) этого раздражителя:

$$S = k * \lg(I / I_{\text{пор}}) \quad (1)$$

где: k – коэффициент пропорциональности;

$I_{\text{пор}}$ – пороговое значение интенсивности раздражителя, при котором возникает ощущение.

² Витцлак Г. Психодиагностика: Теория и практика. –М.: Прогресс, 1986.

Экспериментальные исследования ощущений и связанных с ними двигательных актов интенсивно проводились в Германии в первой лаборатории экспериментальной психологии, созданной в 1878 г. Вильгельмом Вундтом. Широко известны проводимые в его лаборатории измерения интервала времени между восприятием прохождения звезды через заданную точку в объективе телескопа и простой моторной реакцией (нажатием на кнопку), фиксирующей момент ее прохождения. Как оказалось, у разных наблюдателей этот интервал времени значительно варьирует.

Дифференциальной психологии психодиагностика обязана направленностью на изучение индивидуальных различий в особенностях психических процессов, состояний и свойств личности. Основателем дифференциальной психологии считается В. Штерн, опубликовавший в 1900 г. работу "О психологии индивидуальных различий". В методологическом плане дифференциальная психология мало чем отличалась от общей экспериментальной психологии. Она изучала не особенности конкретной личности, а ее отдельные свойства, точнее разброс в поведении и свойствах различных личностей. Именно разброс в свойствах и качествах личности являлся предметом изучения дифференциальной психологии. Соответственно, одним из основных методов исследования в дифференциальной психологии стал дисперсионный анализ.

Возникновение психодиагностики характеризуется переходом от исследования индивидуальных различий к исследованию индивидуальностей. При этом исследование изменчивости свойств и степени их проявления у одного человека остается в психодиагностике как необходимая предпосылка для характеристики целостной индивидуальности отдельной личности, что обеспечивает неразрывную связь психодиагностики с дифференциальной психологией.

В *прикладной психологии*, возникновение которой также связано с именем В. Штерна, психодиагностика, понимаемая как психологическая оценка (в отличие от психологического воздействия или психотехники), рассматривалась как ее составная часть. Проводя различие между психологической оценкой (психодиагностикой) и психотехникой, В. Штерн подчеркивал, с одной стороны, специфичность диагностических проблем, с другой стороны, указывал на единство диагностики и психотехники: без диагностики психотехника не эффективна, без психотехники диагностика не имеет смысла.

Мощным стимулом становления психодиагностики как самостоятельной отрасли психологии послужили достижения в развитии *тестологии*. В 1883 г. в Англии Френсис Гальтон, разрабатывая новую отрасль в науке о человеке, которую он назвал "Антропометрия", предложил особые измерительные испытания зрения, слуха, времени моторной и словесной ассоциативной реакции, для обозначения которых им же было предложено использовать термин "тест". Заслуга Ф. Гальтона состояла в создании техники изучения индивидуальных различий и, прежде всего, во внедрении в психологию статистического метода. Им были разработаны приемы вариационной статистики, которые, вооружали психологию и психодиагностику, в частности, важным методическим средством. Среди этих приемов наиболее перспективным оказался метод исчисления коэффициента корреляции между переменными, который затем был усовершенствован английским математиком Пирсоном.

Однако, основателем тестовой диагностики считается американский психолог Джеймс Кеттелл, который работал в лаборатории Ф. Гальтона и с особым энтузиазмом воспринял его приемы определения индивидуально-психологических качеств и статистической обработки результатов. В 1890 году им была опубликована статья "Умственные тесты и измерения", в которой он пишет: "Психология не сможет стать прочной и точной, как физические науки, если не будет базироваться на эксперименте и измерении. Шаг в этом направлении может быть сделан путем применения серии умственных тестов к большому числу индивидов"³. Вместе с тем Дж. Кеттелл высказал мысль о том, что научная и практическая ценность тестов возрастет, если условия их проведения будут одинаковыми. Тем самым он сформулировал необходимость стандартизации тестов, для того чтобы стало возможным сравнение их результатов, полученных разными исследователями на разных испытуемых.

В 1891 г., вернувшись в США, Дж. Кеттелл основал в Колумбийском университете самостоятельную психологическую лабораторию, в которой ставилась задача изучения целостной личности. Однако индивидуальные свойства личности рассматривались и выявлялись им изолированно друг от друга. Изолированное рассмотрение отдельных психических функций, а также неадекватность в выборе индикаторов (показателей) диагностируемых свойств явления весьма характерные для тестологии. Эти же недостатки порой имеют место и в современной психодиагностике, в связи с чем

³ см. Ярошевский М.Г. История психологии. –М., 1996.

отношение к использованию тестов, особенно заимствованных из зарубежной психологической практики и требующих рестандартизации, неоднозначно.

Развитие тестологии тесно связано с именем французского исследователя Альфреда Бине (1857-1911), который является создателем одного из первых тестов умственного развития, предназначенного для того, чтобы отделить детей, способных к учению, но ленивых и не желающих учиться, от детей с нарушениями в развитии и не способных учиться в нормальной школе. В своих исследованиях А. Бине исходил из представлений о том, что умственное развитие или развитие интеллекта происходит независимо от обучения в результате биологического созревания.

Большая заслуга в развитии тестовой диагностики принадлежит крупнейшему русскому невропатологу и психиатру Г.И. Россолимо, который, работая в Московском университете, поставил своей целью найти метод количественного исследования психических процессов в нормальном и патологическом состояниях. Россолимо измерял три группы психических процессов: внимание и воля, точность и прочность восприятия, ассоциативная деятельность. Он предложил графическую форму представления измерений психических процессов — вычерчивание психологического профиля, который наглядно демонстрировал соотношение указанных процессов. Форма профиля оказалась надежным критерием для диагностики умственной отсталости. Положительным в исследованиях Россолимо было то, что в отличие от западного тестирования он стремился к целостной оценке личности, к синтетическому способу изображения ее сильных и слабых сторон, что, по существу, является одной из основных задач современной психодиагностики.

1.3. Классификация психодиагностических методик

Современная психодиагностика обладает множеством разнообразных психодиагностических методик, которые разные исследователи классифицируются по-разному. Например, в работе В.К. Гайда, В.П. Захарова⁴ одним из оснований классификации является степень стандартизации методики. В учебном пособии по психологической диагностике под редакци-

⁴ Гайда В.К., Захаров В.П. Психологическое тестирование: учебное пособие. - Л.: Изд-во ЛГУ, 1982.

ей К.М. Гуревича, Е.М. Борисовой⁵ в качестве критерия классификации рассматривается степень формализации методики. К высоко формализованным методикам относятся различного рода тесты, опросники, проективные техники и психофизиологические методики, для которых характерна жесткая регламентация процедуры обследования (точное соблюдение инструкций, строго определенные способы предъявления стимульного материала, невмешательство исследователя в деятельность испытуемого и др.), а также наличие норм или других регламентированных критериев оценки результатов. Стандартизованные и формализованные методики позволяют собрать диагностическую информацию в относительно короткие сроки и в таком виде, который дает возможность количественно и качественно сравнивать индивида с другими людьми.

К малоформализованным методикам относят процедуры наблюдения, опроса, анализа продуктов деятельности, обеспечивающие получение таких сведений об испытуемом, которые плохо поддаются объективизации (например, особенности субъективных переживаний, личностные смыслы) или являются чрезвычайно изменчивыми (динамика целей, состояний, настроений и т. д.). Даная группа методик достаточно трудоемка в использовании, требует больших затрат времени и предъявляет более высокие требования к профессиональной квалификации психодиагноста.

В практике психодиагностического обследования необходимо гармоничное сочетание формализованных методик с малоформализованными, которые, как правило, взаимно дополняют друг друга.

Наиболее стандартизованными и высоко формализованными психодиагностическими методиками, несомненно, являются тесты, включая опросники, которые можно рассматривать как вербальные тесты.

Психологические тесты (в переводе с английского test - «испытание», «проверка», «проба») – это стандартизированные краткие и чаще всего ограниченные во времени испытания, предназначенные для установления количественных (и качественных) индивидуально-психологических особенностей человека.

⁵ Психологическая диагностика: Учебное пособие /Под ред. К.М. Гуревича, Е.М. Борисовой. – М.: Изд-во УРАО, 2000.

Классифицировать психодиагностические методики и, соответственно, психологические тесты можно по разным основаниям (критериям), среди которых основными являются:

- ☐ *целевое назначение;*
- ☐ *содержательная направленность;*
- ☐ *характер тестовых заданий;*
- ☐ *форма представления тестовых заданий;*
- ☐ *количество диагностируемых факторов (показателей);*
- ☐ *характер ответов(реакций) на тестовые задания.*

По целевому назначению выделяют:

- *индивидуальные тесты*, предназначенные для диагностики психических свойств и качеств личности отдельного индивида в режиме непосредственного индивидуального взаимодействия с ним психодиагноста. Индивидуальная диагностика необходима при работе с детьми младенческого и дошкольного возраста, в клинической психологии – для тестирования лиц с соматическими или нервно-психическими нарушениями, людей с физическими недостатками. Необходима она и в тех случаях, когда нужен более тесный контакт с испытуемым с целью снятия напряженности, оптимизации условий диагностики, повышения надежности результатов тестирования. К категории индивидуальных тестов относятся, например, тесты интеллекта Станфорд-Бине, Векслера, проективные личностные тесты Роршаха, Люшера, Розенцвейга, тематический апперцептивный тест (ТАТ), тесты слуховой памяти, индивидуально-ориентированные тесты типа репертуарных решеток Келли).

- *групповые тесты*, предназначенные для проведения диагностического обследования одновременно большой группы испытуемых. Первые групповые тесты появились в США в период первой мировой войны, когда возникла необходимость диагностировать большие массы людей с целью отбора наиболее подготовленных к тому или иному виду деятельности, а также распределения по разным видам деятельности людей в соответствии с их индивидуальными особенностями. В то время как индивидуальные тесты применялись, в основном, в клинике и при консультировании, групповые тесты использовались преимущественно в системе образования, в промышленности и в армии. К групповым тестам можно отнести, например, тесты интеллекта Айзенка, школьный тест умственного развития (ШТУР), прогрессивные матрицы Равена, метод корректурных проб для диагностики устойчивости внимания, социометрический тест для диагностики особенностей межличностных отношений.

- *тесты достижений* (в развитии общих и специальных способностей, приобретении знаний, умений и навыков);
- *тесты профессиональной пригодности*.

По критерию содержательной направленности выделяются тесты:

- *интеллекта*;
- *отдельных психических процессов* (познавательных, эмоциональных, волевых);
- *психических состояний* (в том числе особенностей и свойств внимания);
- *свойств и типов личности* (особенностей мотивов, установок, интересов, ценностных ориентаций, уровня притязаний, самооценки, локус-контроля, черт характера, типа темперамента, общих и специальных способностей);
- *межличностных отношений* (особенностей стиля отношений, социального статуса, уровня совместимости);
- *групповых характеристик* (психологического климата, сплоченности, единства ценностных ориентаций).

По характеру тестовых заданий выделяются:

- *вербальные тесты*, связанные со знанием слов и словосочетаний и умением ими оперировать (запоминание слов, понимание пословиц, сравнение понятий, продолжение фразы и т.п.).

Как уже было отмечено выше, особую категорию вербальных тестов составляют *опросники*, в которых задания представлены в виде вопросов и утверждений. Среди них выделяются опросники личности и межличностных отношений.

- *невербальные тесты*, связанные с умениями:
 - оперировать цифрами или числами (арифметические тесты и тесты на вычисления);
 - оперировать предметами или их образами (запоминание изображения предметов, фотографий, классификация картинок), а также осуществлять определенные предметные или игровые действия (например, доска Сегена, кубики Коса);

- разбираться в графических начертаниях и проводить различные линии (прогрессивные матрицы Равена, графический тест Мира-Лопеца);

- оперировать цветами и цветовыми сочетаниями (тест Люшера);
- рисовать отдельные объекты или их сочетания (например, рисуночные тесты "Семья", "Несуществующее животное" или субтест "Нарисуй мужскую фигуру" из теста "Готовность к школе" Керна-Йерасака).

Невербальные тесты уменьшают влияние языковых и культурных различий на результат обследования. Они также облегчают процедуру обследования испытуемых с нарушением речи, слуха или с низким уровнем образования. В качестве отдельных субтестов они включены во многие тесты интеллекта, общих и специальных способностей, тесты достижений.

По форме представления тестовых заданий выделяются тесты:

- *бланковые*, в которых тестовые задания и инструкции по их выполнению изложены на бумаге в графической или текстовой форме на отдельных бланках или в тестовых тетрадях и брошюрах. Преимуществом бланковых тестов является то, что после соответствующего размножения (ксерокопирования), не требующего больших финансовых затрат, их можно использовать для одновременного тестирования достаточно большой группы испытуемых.

- *компьютерные*, когда тестовые задания или стимульный материал предъявляются на экране компьютерного монитора и испытуемый выполняет их с использованием клавиатура или "мышки". Применение компьютерных тестов требует наличия дорогостоящего компьютерного оборудования и специальных компьютерных классов. Однако компьютерные тесты позволяют автоматизировать процесс тестирования и осуществлять обработку диагностируемых показателей, так сказать, "в реальном времени", т. е. непосредственно в процессе тестирования.

- *аппаратурные*, когда предъявление тестовых заданий (стимульного материала) и регистрация ответных действий испытуемого осуществляется с помощью специальной аппаратуры, например, рефлексометра, тахистоскопа. Аппаратурные тесты наиболее часто используются при проведении психофизических или психофизиологических исследований, а также в профессиональной психодиагностике.

По количеству определяемых факторов (показателей) тесты делятся на:

- *однофакторные*, когда диагностируется лишь один показатель, например, общительности, объема памяти, времени реакции, уровня интеллекта;

- *многофакторные*, когда тест позволяет диагностировать сразу несколько показателей (факторов). К таким тестам относятся, например, тесты интеллекта Айзенка, Векслера, Амтхауэра или такие тесты личности, как 16-ти факторный опросник Кеттелла, тест СМОЛ, тест интерперсональных отношений Т. Лири, тест стратегий поведения в конфликтной ситуации К. Томаса, тест профессиональных интересов и склонностей Е.А. Климова. Как правило, многофакторные тесты содержат в своей структуре соответствующие субтесты.

По характеру ответов (реакций) на тестовые задания выделяют тесты:

- с *альтернативными ответами* типа "да-нет", "верно-неверно" (например, тест Айзенка для диагностики темперамента или тест СМОЛ для диагностики свойств личности);

- с *множеством вариантов ответов* в виде:
 - перечня вариантов (тест Равена, школьный тест умственного развития (ШТУР), тест К. Томаса);

- оценочной шкалы, представленной либо числовыми значениями (например, от 0 до 5 или от -3 до +3), либо словесными формулировками (например: необходимо–желательно–нежелательно–недопустимо или никогда–редко–часто–всегда).

- с *свободным выбором ответов*, когда испытуемому не предлагаются какие-либо готовые варианты ответов, и он относительно свободен в формулировках своего ответа (например, тесты Торренса, Роршаха, Розенцвейга или "Неоконченные предложения").

Достоинством тестов, в которых предлагаются два или несколько вариантов ответов, является простота процедуры регистрации и обработки данных, четкая формализация оценивания. Однако при этом происходит определенное навязывание испытуемому готовых вариантов ответа, ограничение в ответах, которое нередко вызывает у испытуемого затруднения.

Тесты со свободным выбором ответов, в группу которых входят практически все *проективные тесты*, требуют, как правило, использова-

ния качественных методов анализа, соотнесения произвольных ответов со стандартными, нормативными или наиболее часто встречающимися у определенной категории испытуемых. При этом возникают сложности с формализацией ответов и их оценок, затруднения в интерпретации результатов. Процедура обработки и анализа результатов, полученных с помощью таких тестов, более громоздка, требует больших затрат времени и более высокой квалификации психодиагноста.

Для *проективных тестов* характерным является использование в качестве тестовых заданий неопределенных, неоднозначных (слабоструктурированных) стимулов, которые испытуемый должен структурировать, конструировать, развивать, дополнять, интерпретировать. В соответствии с исходной проективной гипотезой содержание и характер восприятия, эмоциональные проявления, чувства, высказывания, двигательные акты, следующие в ответ на предъявление стимульного материала, несут на себе отпечаток личности. Стимульный материал проективного теста побуждает испытуемого к определенным формам поведения, реагирования не столько в силу его объективного содержания, сколько в связи с тем личностным смыслом, значением, который испытуемый ему придает. Реакция испытуемого на предъявление тестовых заданий проективного теста определяется особенностями его личности, его личным опытом, содержанием его проблем, переживаний и прочими индивидуально-личностными свойствами, которые, по сути, "проецируются" на стимульный материал.

Различают следующие группы проективных тестов:

- *конститутивные*, направленные на структурирование, оформление стимульного материала, придание ему определенной оформленности и соответствующего смысла (тест Роршаха);
- *конструктивные*, связанные с созданием, конструированием некоторого осмысленного целого из отдельных уже оформленных частей, деталей (тест Мира);
- *интерпретативные*, направленные на получение определенного истолкования, интерпретации какого-либо события, какой-либо ситуации (тематический апперцептивный тест ТАТ);
- *экспрессивные*, связанные с выражением своих представлений, переживаний, проблем путем рисования на свободную или заданную тему (тесты "Дом-дерево-человек", "Несуществующее животное");

- *импрессивные*, связанные с предпочтением одних стимулов по отношению к другим, выделением их как наиболее предпочитаемых (тест Люшера, тест Сонди);

- *аддитивные*, направленные на завершение фраз, рассказов, историй, картин, которые допускают множественность вариантов завершения (тест "Неоконченные предложения").

В исследованиях личности проективные методики и проективные техники используются очень широко. Однако, в их применении есть серьезные проблемы и трудности, среди которых наиболее существенными являются:

- низкий уровень надежности и валидности диагностической процедуры;

- сильная зависимость результатов диагностики от ситуационных факторов: инструкции, состояния испытуемого, характера взаимодействия испытуемого и психодиагноста, а также особенностей личности самого психодиагноста;

- отсутствие нормативных данных, недостаточная стандартизованность и проблемы с формализацией ответов и поведенческих реакций испытуемых, необходимой для построения психодиагностических шкал;

- значительный субъективизм в оценках и интерпретации результатов диагностики.

Чтобы исключить или хотя бы снизить уровень субъективизма в психодиагностике, психодиагностические методики должны быть стандартизованы. Как отмечает известный американский тестолог А. Анастази⁶, "стандартизация – это единообразие процедуры проведения и оценки выполнения теста". Таким образом, стандартизация рассматривается в двух планах: как выработка единых требований к процедуре психодиагностики и как определение единого критерия оценки результатов диагностического обследования.

Стандартизация психодиагностической процедуры подразумевает унификацию инструкций, бланков обследования, способов регистрации результатов, условий проведения обследования.

⁶ Анастази А. Психологическое тестирование: Книга 1 /Пер. с англ. –М.: Педагогика, 1982.

К числу требований, которые необходимо соблюдать при проведении психодиагностики, можно, например, отнести такие:

1) инструкции следует сообщать испытуемым одинаковым образом, как правило, письменно; в случае устных указаний они даются в разных группах одними и теми же словами, понятными для всех, в одинаковой манере;

2) ни одному испытуемому не следует давать никаких преимуществ перед другими;

3) в процессе психодиагностики не следует давать отдельным испытуемым дополнительные пояснения;

4) психодиагностические процедуры с разными группами следует проводить в одинаковое по возможности время дня, в сходных условиях;

5) временные ограничения в выполнении тестовых заданий для всех испытуемых должны быть одинаковыми и т.д.

Обычно авторы психодиагностических методик в соответствующих руководствах приводят точные и подробные указания по процедуре их использования и проведению с их помощью психодиагностического обследования, включая порядок обработки и интерпретации результатов обследования.

1.4. Этические аспекты психодиагностики

Психодиагностика является одной из форм практической деятельности психолога, где пересекаются интересы многих заинтересованных лиц. Прежде всего, это интересы испытуемого, который подвергается обследованию, и психолога, который это обследование осуществляет. Однако, если испытуемым, например, является учащийся, в число заинтересованных лиц должны быть включены его родители, а также учителя, которые по своему статусу могут претендовать на получение сведений о результатах обследования. Кроме того, используя для целей психодиагностики определенные психодиагностические средства, психолог демонстрирует уровень своей квалификации и в какой-то степени олицетворяет и представляет ту организацию и тех преподавателей, где и у кого он получал профессиональную психологическую подготовку. В результате психолог-диагност прямо или косвенно оказывается вовлеченным в широкую систему межличностных отношений, реализация которых предполагает

знание и соблюдение определенных этических норм поведения. Обычно эти нормы формулируются и прописываются в специальных кодексах. В США, например, этический кодекс психолога был принят Американской психологической ассоциацией еще в 1963 году и затем неоднократно пересматривался и дополнялся. В нашей стране кодекс психолога-диагноста находится, по существу, на стадии разработки и обсуждения различных вариантов. Этическим стандартам должны соответствовать не только вопросы использования психодиагностических методик, но и их распространение.

В качестве основных проблем, отражающих этические аспекты психодиагностики, можно выделить следующие⁷:

1. *Уровень квалификации специалиста*, применяющего психодиагностические методики.

Требование, чтобы психодиагностическое обследование проводилось только достаточно квалифицированными специалистами в области психодиагностики, является первым шагом по защите испытуемого от неправильного использования психодиагностических тестов и опросников. Для правильного применения многих индивидуальных тестов и большинства личностных тестов и опросников требуется относительно длительный период интенсивного обучения и стажировки под руководством опытного специалиста. В то же время применение групповых тестов для тестирования, например, достижений в учебной и профессиональной деятельности, не требует наличие у психодиагноста высокой квалификации. Студенты, которые участвуют в учебном тестировании, обычно не готовы к самостоятельному проведению диагностического обследования других людей и к интерпретации тестовых оценок.

В идеальном случае хорошо подготовленный психолог-диагност должен выбирать методики, которые подходят как для частной цели, с которой он проводит диагностирование, так и для конкретного человека, которого он обследует. Он также должен хорошо знать соответствующую научную литературу, касающуюся выбранной методики, и быть способен оценить технические параметры таких ее характеристик, как нормы, надежность и валидность.

⁷ см. Психологическая диагностика: учебное пособие /Под ред. К.М. Гуревича, Е.М. Борисовой. – М.: Изд-во УРАО, 2000.

Известно, что результаты диагностирования чувствительны к множеству условий его проведения. Поэтому психодиагност должен быть достаточно осведомлен в психологии, чтобы уберечься от неоправданных выводов в своей интерпретации полученных оценок. Если диагностирование проводится людьми других профессий, желательно, чтобы имелся квалифицированный психолог-консультант, который поможет обеспечить необходимые условия для правильной процедуры и последующей правильной интерпретации диагностических оценок.

Неверные представления о характере и цели обследования, а также неправильные интерпретации диагностических результатов лежат в основе многих распространенных ошибок и критических замечаний в адрес психологической диагностики.

Очевидно, в связи с разнообразием областей исследования и, следовательно, специализацией в профессиональной подготовке ни один психолог не является одинаково квалифицированным во всех областях психологии. Признавая этот факт, психолог должен знать границы компетентности и ограниченность своих методов и не предлагать свои услуги, а также не использовать технику, которые не удовлетворяют профессиональным стандартам, установленным в отдельных областях.

Следует признать различие, существующее между психологом, работающим в системе научных или государственных учреждений, например, в школе, университете, клинике или государственных органах, и психологом, занятым самостоятельной практической деятельностью. Поскольку независимый психолог-практик менее подчинен оценкам и суждениям квалифицированных коллег, чем психолог, работающий в учреждении, он должен отвечать более высоким требованиям к профессиональной квалификации. Это же справедливо и в отношении психологов, ответственных за контроль над другими учрежденческими психологами или консультирующих этих психологов. Для повышения профессиональных норм и качества психологического обследования во многих странах, и у нас в том числе, квалифицированным психологам предоставляются специальные лицензии, выдаются соответствующие удостоверения и сертификаты.

2. Порядок приобретения и использование психодиагностических методик.

Право приобретения и последующего использования методик психодиагностики должно предоставляться лицам, имеющим определенную квалификацию. Но сегодня приобретать их могут все. Могут быть, конечно, и индивидуальные покупатели тестов, например студенты, которым может потребоваться какой-то тест для учебных или исследовательских целей. Американские психологи считают, что они должны иметь заказ на приобретение методики, скрепленный подписью преподавателя психологии или другого лица, который берет на себя ответственность за ее правильное использование. В зарубежных этических кодексах психологов-диагностов предлагается ограничить распространение диагностических методик. Это ограничение имеет двоякую цель: неразглашение содержания методик и предупреждение их неправильного применения. Доступ к таким методам должен быть ограничен теми людьми, кто имеет профессиональную заинтересованность и гарантирует их правильное использование. Диагностические оценки, как и сами методики, передаются только лицам, которые способны их интерпретировать и использовать надлежащим образом.

Другие профессиональные обязательства должны быть связаны с продажей психологических методик, осуществляемой их авторами и разработчиками, а также издателями. Недопустимы какие-либо заявления относительно достоинств методики при отсутствии достаточных на то объективных оснований. Методику нельзя выпускать для общего применения преждевременно. Когда вначале ее распространяют только для исследовательских целей, это условие должно быть ясно указано и соответственно должно быть ограничено распространение методики. Руководство по ее использованию должно предоставлять достаточные данные, позволяющие оценить методику как таковую, а также давать полную информацию относительно ее проведения, способов оценки и норм. Руководство должно являться действительным изложением того, что известно о методике, а не средством рекламы, предназначенным для того, чтобы показать ее в выгодном свете. Обязанностью автора и издателя диагностических методов является их периодическая проверка и повторная стандартизация с целью предотвратить их устаревание. Быстрота, с которой методика устаревает, зависит от ее содержания.

Методики и их основные части нельзя печатать в газете, журнале или популярной брошюре ни в целях описания, ни для использования их при самооценке. А. Анастаси считает, что в этих условиях самооценивание не только приведет к таким крупным ошибкам, как признание бесполезности методики, но может также быть психологически вредным для индивида. Кроме того, любая гласность в отношении содержания методики сделает недействительным ее будущее применение другими людьми. Представление методических материалов таким способом приводит к созданию ошибочной, искаженной картины психологического обследования в целом. Такое разглашение может породить либо наивную доверчивость, либо огульное противодействие со стороны общественности всем направлениям психологической диагностики.

Не отвечающим профессиональным требованиям видом психодиагностики является заочная форма тестирования или тестирование по почте, которая достаточно широко применяется за рубежом и начинает распространяться в нашей стране. Такой способ не только не обеспечивает контроля за соблюдением условий тестирования, но предполагает также интерпретацию тестовых оценок в отсутствие другой относящейся к индивиду информации. При таких обстоятельствах тестовые результаты не только могут оказаться бесполезными, но и принесут вред.

3. *Обеспечение тайны результатов обследования*, особенно при использовании психодиагностических методик личностного характера.

Данная проблема и необходимость в соответствующих этических нормах возникают в связи с тем, что при тестировании невольно могут выявиться такие черты и особенности личности испытуемого, которые он не хотел раскрывать или, по крайней мере, не хотел бы акцентировать на них внимание психодиагноста. Поскольку о сущности некоторых тестов, раскрывающих эмоциональные и мотивационные особенности, а также установки личности, субъекту не сообщается, данная ситуация вполне вероятна. Существование таких не прямых методов психодиагностики налагает серьезную ответственность на использующих их психологов. Хотя беспокойство, касающееся посягательства на тайну личности, чаще всего проявляется по отношению к личностным тестам, его логично отнести к любому типу психодиагностических методик, поскольку и интеллектуальные тесты, и тесты способностей, и тесты достижений могут обнаружить такие недостатки в навыках и знаниях, которые индивид предпочитает не

раскрывать. Кроме того, любое наблюдение за поведением индивида во время интервью, беседы и других исследований личности может открыть такую информацию, которую он сам не знал или которую предпочитал скрывать. Таким образом, любые методы психологического исследования, а не только тесты тают в себе возможность проникновения в тайну личности. Конечно, возникающие конфликты ценности психологического исследования и посягательства на тайну личности могут быть разрешены в каждом конкретном случае, однако в действительности эта проблема не так проста, и ее решение требует большой осторожности со стороны психологов и других профессионалов. Чтобы оградить тайну личности, нельзя сформулировать универсальные правила, можно предусмотреть только общие пути. При применении этих общих способов к специфическим случаям необходимы этическая сознательность и профессиональная ответственность каждого психолога. Решения нужно вырабатывать в зависимости от конкретных обстоятельств.

Когда обследование проводится в интересах общества или какого-либо учреждения, обследуемый должен быть полностью проинформирован о том, как будут использоваться его оценки. Желательно также объяснить ему, что правильная оценка будет выгодна самому обследуемому, так как ему не принесет пользы, если он займет место, на которое в дальнейшем у него не хватит сил или для которого он потом будет признан непригодным. Диагностические результаты, полученные в клинике или консультации, не могут быть применены в интересах этих учреждений, если клиент не даст своего согласия на это.

Диагностические процедуры и экспериментальные планы, которые охраняют право индивида отказываться от участия в обследовании и, следовательно, защищают тайну его личности, осложняют работу психолога и повышают требования к его квалификации. При правильных взаимоотношениях и взаимном уважении психолога и обследуемого число отказов от участия в диагностировании может быть сведено к минимуму.

4. Порядок использования и распространения результатов психодиагностики или принцип конфиденциальности.

Как и проблема сохранения тайны личности, проблема конфиденциальности предполагает наличие определенных этических норм. Основной вопрос ставится так: "Кто будет иметь доступ к диагностическим результатам?". С одной стороны, существует необходимость неразглашения со-

держания теста и опасность неверного понимания тестовых оценок, с другой – объективная необходимость для разных лиц знать результаты тестирования.

В настоящее время усилилось осознание индивидом собственного права иметь доступ к результатам своего обследования. Он также должен иметь возможность комментировать содержание своего ответа и в случае необходимости разяснять или исправлять фактическую информацию. Должны быть соблюдены соответствующие меры предосторожности, направленные против неправильного использования и неверной интерпретации диагностических результатов.

Один из этических вопросов конфиденциальности, возникающий при обследовании школьников, состоит в том, сообщать ли родителям ученика или учителям и администрации школы результаты тестирования. Однако все чаще исследователи приходят к мысли, что вопрос не в том надо ли сообщать, а в том – как это делать. Обычно родители и в какой-то степени учителя имеют законное право на получение информации о ребенке. Чаще всего они хотят получить такую информацию. Вместе с тем, в некоторых случаях школьная неуспеваемость ребенка или затруднения эмоционального характера могут возникать вследствие нарушений во взаимоотношениях ребенка с родителями или учителями. В таких случаях контакт психолога-диагноста с родителями и учителями имеет первостепенное значение как для того, чтобы понять причины результатов психодиагностики, так и для того, чтобы установить с ним сотрудничество в интересах ребенка.

Если обследование проводится в учреждении, например в школе, суде или при оформлении на работу, индивид должен быть заранее проинформирован о его целях, о том, как будут использоваться его результаты, и об их доступности тем лицам, кто в них заинтересован. Бывают ситуации, когда диагностические результаты запрашиваются другими людьми, например, когда будущий наниматель или администрация колледжа просят предоставить им данные тестового обследования индивида, проведенные в школе. В таких случаях требуется получить согласие индивида на передачу данных. Это же относится и к обследованию в клинике или консультации, а также к тестированию, осуществленному с исследовательскими целями.

Другая проблема относится к сохранению диагностических данных в учреждениях. С одной стороны, данные об индивидах, получаемые и сохраняемые в течение длительного времени, могут быть очень полезными не только для исследовательских целей, но и для правильного понимания и консультирования самого индивида. Однако их ценность предполагает правильное использование и правильную интерпретацию. В случаях, когда данные получены либо для длительного использования в интересах индивида, либо для научных целей, для предотвращения неправильного их применения необходимо, чтобы доступ к ним находился под чрезвычайно строгим контролем.

5. *Выбор формы сообщения результатов обследования при их передаче другим лицам.*

Психологи много размышляют о том, как сообщать результаты обследования в форме пригодной для использования. Ясно, что информацию нельзя передавать в том виде, в каком она получена. Ее нужно сопровождать объяснениями, которые должны включать объяснения цели и характера выводов, которые целесообразно сделать на основе полученных результатов, а также указания границ использования данных.

Сообщения о характере выполнения тестовых заданий и качественные описания, сделанные простым языком, предпочтительнее специфических числовых данных, за исключением тех случаев, когда результаты теста сообщаются опытному, хорошо подготовленному психологу-профессионалу.

При сообщении результатов любых диагностических методов желательно принимать во внимание характеристики, особенности того человека, кому передается информация. Это касается не только его образовательного уровня и знаний по психологии и тестологии, но также и его ожидаемой эмоциональной реакции на информацию. Если речь идет о родителях или учителях, например, то их эмоциональные конфликты с ребенком могут препятствовать спокойному и рассудительному восприятию фактической информации о ребенке.

Последняя, но не менее важная проблема касается сообщения результатов обследования самому индивиду, будь то ребенок или взрослый. В этом случае, как и при сообщении данных третьему лицу, применимы те же самые меры предосторожности против неправильной интерпретации. Однако здесь особенно важна индивидуальная эмоциональная реакция на информацию, если индивид занят изучением своих собственных досто-

инств и недостатков. Когда индивиду сообщают его диагностические результаты, следует не только сопровождать их интерпретацией, проводимой компетентными психологами, но нужно создать благоприятные возможности для индивидуальной консультации каждого, кто может быть эмоционально обеспокоен такой информацией. Например, студент колледжа может быть серьезно обеспокоен и озадачен, узнав о том, что он плохо выполнил тест школьных возможностей. Способный школьник может приучиться к лени и безынициативности или может стать непослушным и перестать действовать сообща с товарищами, если обнаружит, что по своим способностям намного превосходит своих сверстников. Развитие серьезных личностных нарушений может быть ускорено, если больному индивиду сообщить его оценку по личностному тесту. Такие негативные воздействия могут возникать независимо от того, правильной или неправильной является сама оценка. Даже в том случае, если обследование было тщательно проведено, а полученные оценки правильно интерпретированы, их знание без возможности обсудить их в дальнейшем может быть вредным для индивида.

ГЛАВА 2

ИЗМЕРЕНИЕ В ПСИХОДИАГНОСТИКЕ

2.1. Понятие об измерении

Проблемами, связанными с теорией и практикой измерений в психологии, занимается особая область психологической науки, называемая психометрией. По существу, психометрия является основой психодиагностики.

В самом широком смысле измерение – это способ установления отношений между объектами или явлениями. В более узком смысле, измерение есть определение наличия и степени выраженности какого-либо свойства (качества) объекта или его признака. В определении С.С. Стивенса⁸, одного из основоположников теории измерений в психологии, измерение – это приписывание числовых форм объектам или событиям в соответствии с определенными правилами.

Объективной основой для измерения является то, что всякое качество имеет определенную количественную меру. В философии качество рассматривается как существенная определенность объекта, в силу которой он является данным, а не другим объектом и отличается от других. Своеобразие качества обуславливает различия между объектами неподобными друг другу. Количество есть характеристика (мера) качества, которая определяет различия между подобными объектами. Поэтому, в принципе, любое качество и его количественная мера могут быть определены (измерены), если, конечно, на данном этапе развития науки для такого измерения существуют (уже разработаны) соответствующие методологические, методические и технические средства.

Сущность любого объекта или явления, которая заключается в его качественной определенности, проявляется в свойствах, отражаемых соответствующими *признаками*, которые по форме своего выражения подразделяются на *качественные (атрибутивные)* и *количественные*. *Атрибутивный признак* указывает на наличие определенного свойства (качества), *количественный признак* показывает меру этого свойства.

⁸ Стивенс С.С. Математика, измерение и психофизика // Экспериментальная психология / Под ред. С.С. Стивенса. -М.: Иностр. лит-ра, 1960.

Например, пол человека, семейное положение, профессия, тип темперамента - это атрибутивные или качественные признаки, а возраст, стаж семейной жизни, уровень профессиональной квалификации, выраженность темперамента - количественные признаки. Атрибутивные признаки могут быть простыми, характеризующими одноаспектные свойства объекта (например, мужской или женский пол; семейное положение: женат-холост), и сложными, характеризующими многоаспектные свойства (характер, интеллект, самооценка). Многоаспектность свойства выражается в наличии множества взаимосвязанных элементов, составляющих его структуру. В частности, можно говорить о структуре характера, структуре интеллекта, структуре самооценки.

2.2. Измерительные шкалы и их характеристика

Возможности и способы измерения качественных (атрибутивных) и количественных признаков различны как в силу различий в характере признаков, так и в силу различий в целях измерения. Тем не менее, можно выделить отдельные принципы, правила измерения, совокупность которых составляют, так называемую, *шкалу измерения*. По определению, *шкала измерения* - это правило, по которому рассматриваемому признаку (или объекту в соответствии с присущим ему признаком) присваивается определенное числовое значение.

Шкала (в переводе с лат. *scala* - лестница), по существу, есть определенный измерительный инструмент. В настоящее время в теории измерений выделяются четыре типа шкал, наименования которых и описание основных особенностей были даны С.С. Стивенсом.

Шкала наименований

Шкала наименований по-другому называется номинальной или номинативной шкалой. Это самая простая шкала, измерение по которой состоит в классификации объектов по определенным признакам. Как правило, признаки, с которыми имеет дело шкала наименований, это атрибутивные признаки. Каждому классу в классификации объектов присваивается некоторый номер или дается некоторое имя, название (в переводе с лат. *poten* - имя, название). Соответствующее числовое значение или имя приписывается объекту в зависимости от того, к какому классу объектов он принадлежит

или относится. Числа в шкале наименований используются лишь как символы для определения (указания) принадлежности объекта к определенному классу. В связи с таким "символьным" назначением, число, приписываемое объекту, может быть заменено на любое другое и даже на любой другой символ, например, на какую-либо букву. Главное требование при таком преобразовании состоит в том, чтобы вновь вводимые числа (или другие символы) были не тождественны друг другу, т. е. позволяли различать обозначаемые ими классы объектов. Внутри класса объекты характеризуются тождеством, т. е. им приписываются одинаковые, тождественные числа (символы). Объектам нетождественным, относящимся к разным классам, приписываются, соответственно, нетождественные числовые значения (символы). Главное, как отмечал Стивенс, не приписывать один и тот же символ разным классам объектов и разные символы одному и тому же классу.

Простейший случай номинальной шкалы - дихотомическая шкала, состоящая всего лишь из двух классов, например: "мужской пол – женский пол", "с высшим образованием – без высшего образования", "успевающий – неуспевающий" и т. п. Признак, по которому объекты (в данном случае испытуемые) распределяются по классам дихотомической шкалы, называется альтернативным признаком.

Более сложным вариантом номинальной шкалы является шкала с тремя и более классами, например: "экстрапунитивные - интрапунитивные - импунитивные реакции", "холерик - флегматик - сангвиник - меланхолик".

Распределив все измеряемые объекты по классам, можно далее перейти от наименований (символов, обозначающих класс) к числам, характеризующим наполненность классов, т. е. указывающих, сколько объектов входит в соответствующий класс. Таким образом, шкала наименований позволяет нам измерять частоты встречаемости объектов, имеющих разные значения (наименования) признака.

Графически шкала наименований может быть представлена в виде линии, на которой обозначены наименования классов измеряемых объектов.

В качестве примера на рис. 2.1 представлена шкала наименований для измерения типа темперамента.

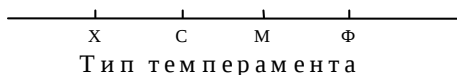


Рис. 2.1. Шкала наименований для измерения типа темперамента.

Наименования этой шкалы могут быть преобразованы (заменены) на любые другие отличающиеся наименования (обозначения). Например, вместо "Х – С – М – Ф" могут быть взяты цифры "1 – 2 – 3 – 4" или буквы "а – б – в – г". Обычно для удобства либо оставляют полное наименование признака, либо дают его в сокращенном виде.

Шкала порядка	<i>Шкала порядка</i> (или, по-другому, ранговая шкала) имеет больше измерительных возможностей по сравнению со шкалой наименований.
--------------------------	---

Принцип измерения по этой шкале состоит в *ранжировании* или *упорядоченном распределении* объектов по определенному признаку в зависимости от меры (степени выраженности) этого признака у каждого из объектов. Отсюда следует, что в шкале порядка мы имеем дело уже с количественными признаками.

В качестве *объектов*, подлежащих ранжированию, могут быть, например:

- *испытываемые или группы испытуемых* (ранжирование по росту (размеру), успеваемости, уровню интеллекта, воспитанности и т. д.),
- *отдельные качества, особенности психики и личности* (по предпочтению, по значимости, по степени выраженности и пр.).

Расположение i -го объекта на шкале порядка и соответствующее приписывание ему числового значения x_i определяются тем, в какой мере (в большей или в меньшей) присущ этому объекту измеряемый признак по сравнению со значением этого признака у другого объекта. Число x_i при измерении по шкале порядка обозначает место i -го объекта в упорядоченном ряду среди других объектов. Это место и это число называется рангом объекта, которое лучше обозначать символом r_i .

Ранжирование объектов может осуществляться как в восходящем, так и в нисходящем порядке, т. е. ранговая шкала, в отличие от шкалы наименований, имеет определенную направленность, которая в графическом представлении обозначается стрелкой, указывающей направление в изменении рангов. Обычно принято, и это удобно, присваивать ранги объектам

(испытуемым) с 1 до N (в восходящем варианте ранжирования) или с N до 1 (в нисходящем варианте), где N - количество ранжируемых объектов.

В качестве примера на рис. 2.2 представлена шкала порядка (шкала рангов) для измерения интеллекта в группе из N испытуемых.



Рис. 2.2. Шкала порядка для измерения уровня интеллекта в группе из N испытуемых.

В шкале порядка допустимо преобразование исходных значений r_i измеряемого показателя x в другие числа a_i в соответствии с формулой:

$$a_i = f(r_i) \quad (2.1)$$

где $f(r_i)$ - любая монотонно возрастающая функция.

Графически данная форма преобразования представлена на рис. 2.3.

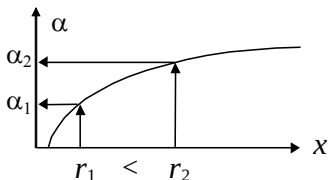


Рис. 2.3. Допустимое преобразование значений измеряемого показателя в шкале порядка.

В результате соответствующего преобразования соотношение $\alpha_2 > \alpha_1$ сохраняется эквивалентным соотношению $r_2 > r_1$.

Данные, полученные по шкале порядка, можно преобразовать в показатели соответствующей номинальной шкалы. Например, для преобразования значений рангов в показатели дихотомической шкалы достаточно множество рангов от 1 до N разбить на две группы. В первую группу включить ранги от 1 до $N/2$, во вторую группу – от $N/2$ до N . Если первой группе дать наименование "А", а второй – "Б", то все объекты, имеющие в шкале порядка ранги от 1 до $N/2$, в новой дихотомической шкале будут иметь значение А. Другие объекты, ранги которых более высокие, получают значения Б. Аналогичным образом можно преобразовать данные, полученные по шкале рангов, в значения номинальной шкалы с несколькими классами (наименованиями).

Шкала интервалов

Обладая всеми свойствами порядковой шкалы, шкала интервалов имеет дополнительные возможности. Она отличается тем, что определяет не только наличие различий объектов по измеряемому признаку, но и степень различия. Минимальная степень различия, определяемая в шкале интервалов, принимается за единицу измерения. Поскольку единица измерения, по существу, представляет собой атрибутивный признак, ей можно дать определенное наименование. Обычно, если есть такая возможность, стараются использовать известные, общепринятые наименования.

Измерение по шкале интервалов состоит в определении числа (количества) единиц измерения, составляющих величину измеряемого признака.

Один объект отличается от другого по измеряемому признаку x на определенное количество интервалов, составляющих единицу измерения:

$$x_2 - x_1 = k \cdot \delta x \quad (2.2)$$

где: δx - единица измерения признака x ;

k - число единиц измерения, характеризующее степень различия объектов;

x_1 и x_2 - значение измеряемого признака x у первого и второго объекта, соответственно.

Шкала интервалов и единица измерения в этой шкале всегда привязаны к определенной процедуре или к конкретной психодиагностической методике. В шкале интервалов минимально возможное значение измеряемого признака всегда равно нулю, последующие значения отличаются от предыдущего на единицу измерения δx .

В качестве примера на рис. 2.4. представлена шкала интервалов для измерения объема памяти по методике "Запоминание 12 слов". Единицей измерения в этой шкале является *слово*, которое указывается на шкале в скобках рядом с наименованием измеряемого признака. Минимально возможное значение признака равно 0 слов, максимально возможное - 12 слов. Количество градаций или возможных значений признака равно 13 (от 0 до 12).

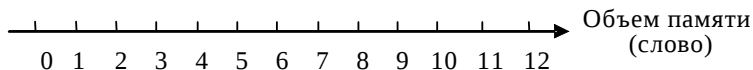


Рис. 2.4. Шкала интервалов для измерения объема памяти по методике "Запоминание 12 слов".

В этой шкале, как и в предыдущих, допускается преобразование исходных значений измеряемого признака в другие, так называемые, производные значения. Допустимым в шкалах интервалов является линейное преобразование, определяемое соотношением:

$$a_i = a * x_i + b \quad (2.3)$$

где a и b – произвольные коэффициенты.

При таком преобразовании соотношение $a_2 - a_1 = k * \delta a$ сохраняется эквивалентным исходному соотношению $x_2 - x_1 = k * \delta x$, но уже в других единицах измерения, а именно в единицах δa . И в исходных единицах измерения δx , и в новых δa первый объект отличается от второго на k единиц.

Значения показателей, полученные при измерении по шкале интервалов, можно преобразовать в соответствующие значения шкал более низкого уровня: шкалы порядка или шкалы наименований.

Для преобразования значений признака x , измеренных по шкале интервалов, в значения, соответствующие шкале порядка, достаточно про ранжировать объекты, обладающие данным признаком, по значениям x_i . Если два или несколько объектов будут иметь одинаковые значения признака x , им присваиваются одинаковый ранг, значение которого равно среднему от суммы рангов, на которые они претендуют. Например, если у четырех объектов (пятого, седьмого, десятого и двенадцатого) одинаковые значения признака, т. е. $x_5 = x_7 = x_{10} = x_{12}$, и они претендуют на 2, 3, 4 и 5 ранги, каждому присваивается ранг $(2+3+4+5)/4$, т. е. 3,5. Следующему объекту с большим значением признака, если он один с таким значением признака, присваивается ранг 6, поскольку предыдущие ранги уже были учтены или задействованы. Однако, если такое же значение признака есть еще и у другого объекта, то им обоим присваивается средний ранг $(6+7)/2=6,5$.

Шкала отношений

В отличие от шкалы интервалов шкала отношений (или пропорциональная шкала) позволяет определять не только, насколько отличаются объекты по измеряемому признаку, но и во сколько раз один объект отличается по этому признаку от другого объекта. Это наиболее мощная измерительная шкала из всех рассмотренных.

Шкала отношений характеризуется наличием:

- о направления в изменении измеряемого признака (свойство шкалы порядка),
- о единицы измерения или минимально измеряемого интервала в значениях признака (свойство шкалы интервалов),
- о абсолютной "нулевой" точки отсчета, указывающей на абсолютное отсутствие измеряемого признака (особое свойство шкалы отношений).

В шкале отношений допустимо следующее преобразование исходных значений измеряемого признака:

$$a_i = k * x_i \quad (2.4)$$

где k - произвольный коэффициент.

В результате соответствующих преобразований отношение a_2/a_1 сохраняется равным отношению x_2/x_1 .

В психологических исследованиях при измерении параметров различных психических процессов и явлений обычно используются лишь первые три типа измерительных шкал.

Шкала отношений практически не используется, поскольку трудно найти такое психическое явление, для характеристики проявлений которого можно выделить признаки, имеющие нулевые значения в абсолютном, а не относительном смысле. Тем не менее, попытки создать особую шкалу типа шкалы отношений для использования ее в психологических исследованиях предпринимались неоднократно. Г. Раш (G. Rasch, 1960), в частности, предложил такого рода шкалу для измерения интеллектуальных способностей человека.

Раш исходил из того, что человек решает или не решает ту или иную проблему (задачу) с определенной вероятностью. Если в качестве характеристики проблемы (задачи) рассматривать показатель ее трудности, то в качестве некоторой интеллектуальной характеристики личности можно рассматривать степень вероятности, с которой данная личность способна решать эти проблемы (задачи).

Процедура шкалирования, согласно Рашу, состоит в измерении интеллектуальной способности личности, используя соотношение между вероятностью решения задач и их трудностью (см. рис. 2.5).

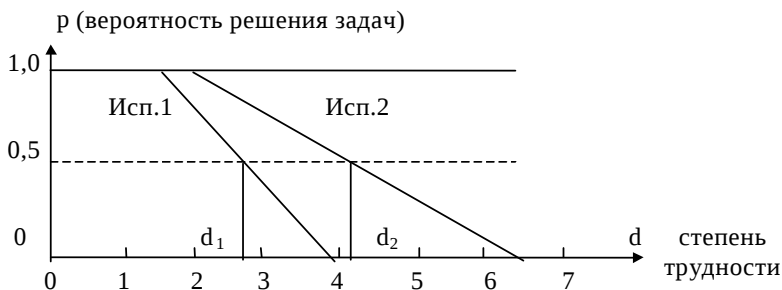


Рис. 2.5. Зависимость вероятности (p) решения задач от степени их трудности (d).

Для оценки трудности задач выбирается некоторая выборка испытуемых, на которой эмпирически определяется частота решения тех или иных задач. Задачи с одинаковой частотой решения, группируются в серию задач соответствующей степени трудности. Таким образом, строится линейная шкала трудности серий задач, в которой степень трудности задач (d) определяется как отношение общего количества испытуемых (N) к числу испытуемых (n), решивших задачи данной трудности, т. е. $d = N/n$.

Далее задачи различных по трудности серий предъявляются контрольным испытуемым, и определяется трудность задач, решение которых осуществляется ими с вероятностью 50% (или $p = 0,5$). Если один испытуемый (Исп.1) на уровне $p = 0,5$ решает задачи, уровень трудности которых равен d_1 , а другой испытуемый (Исп.2) при том же уровне вероятности решает задачи более высокого уровня трудности $d_2 > d_1$, это означает, что его интеллектуальные способности выше интеллектуальных способностей первого испытуемого в d_2/d_1 раз.

Согласно данным, приведенным на рис. 2.5, уровень трудности задач, решаемых испытуемыми с вероятностью 0,5, который характеризует их интеллектуальные способности, составляет для испытуемого Исп.1 $d_1 = 2,8$, для испытуемого Исп.2 $d_2 = 4,2$. Следовательно, уровень интеллектуальных способностей второго испытуемого выше соответствующих интеллектуальных способностей первого испытуемого в $4,2 / 2,8 = 1,5$ раза.

Процедура шкалирования, предложенная Рашем, не нашла, однако, особой поддержки и дальнейшего развития, главным образом, по причине сложности в создании и соответствующей калибровке заданий по степени трудности.

2.3. Виды измерений

Измерения в психологических исследованиях различаются не только типом используемой шкалы, но и критерием, так называемой, нормы, т. е. тем, что рассматривается в качестве основы для сравнения измеряемого показателя. В зависимости от того, с чем сравнивается измеряемый показатель, различают следующие виды измерения⁹:

- ☐ *нормативное,*
- ☐ *критериальное,*
- ☐ *ипсативное.*

Нормативное измерение предполагает сравнение индивидуальных значений измеряемого показателя с некоторой нормой, получаемой, как правило, путем усреднения соответствующих значений показателя по достаточно большой группе испытуемых. При таком измерении оценка индивидуального значения показателя оказываются зависимой от значений этого показателя у других испытуемых. Индивидуальное значение может рассматриваться как высокое в сравнении с одной выборкой испытуемых и как низкое - в сравнении с другой.

Критериальное измерение основано на прямой оценке индивидуального значения показателя без учета соответствующих значений показателя других испытуемых. Сравнение осуществляется с независимым стандартом (критерием), определяемым из рассмотрения содержания тестовых заданий и возможных уровней их выполнения. Наиболее часто данный вид измерения используется в педагогической практике для оценки знаний, умений и навыков учащихся, а также в профессиональном отборе.

Примером критериального измерения может быть, например, измерение уровня умственного развития или уровня усвоения учебного материала с помощью известных тестов ШТУР, ШТОМ. Если нормативное измерение направлено, главным образом, на выявление индивидуальных различий, и в этом смысле должно стремиться по возможности к максимальной дисперсии индивидуальных значений показателя, то критериальное измерение не предполагает достижение значительной дисперсии. Бо-

⁹ см. Кулагин Б.В. Основы профессиональной психодиагностики. –Л.: Медицина, 1984.

лее того, результаты тестов и других измерительных процедур, используемых при критериальном измерении, могут давать очень малую дисперсию, и это будет свидетельствовать лишь о том, что каждый из испытуемых достиг примерно одинакового уровня компетентности в области знаний, умений и навыков, определяемой заданным критерием.

Ипсативное измерение предполагает оценку индивидуального значения показателя с, так называемой, индивидуальной нормой, т. е. с нормой, характерной для самого индивида. При ипсативном измерении оцениваются, следовательно, не межиндивидуальные соотношения показателей, а внутрииндивидуальные. Например, ответы отстающего ученика оцениваются не по отношению к ответам отличников, а по отношению к его же ответам, зарегистрированным ранее. Соответственно и ответы отличника оцениваются по отношению к его же ранее зарегистрированным ответам, а не в сравнении с ответами других, в том числе, и отстающих учеников. Поэтому плохой результат, показанный успевающим учеником, несмотря на то, что по своим показателям он не хуже результата неуспевающего ученика, при ипсативном измерении может иметь оценку более низкую по сравнению с соответствующей оценкой результата неуспевающего ученика. При нормативном и критериальном измерениях более высокие абсолютные значения измеряемого показателя в сравнении с более низкими всегда оцениваются выше.

2.4. Основные понятия и процедуры статистического анализа

Понятие выборки

В исследованиях, имеющих статистический характер, одним из центральных понятий является понятие "выборки". В психологических исследованиях, чтобы получить эмпирические данные, касающиеся особенностей, механизмов и закономерностей психических явлений, проводят соответствующие измерения на определенной группе испытуемых. Вот эта группа испытуемых и представляет собой выборку. То есть выборка – это некоторая группа испытуемых, обладающая определенными общими признаками, выделенными в соответствии с целями исследования. Принято выделять *генеральную выборку* (генеральную совокупность испытуемых) и *конечную или ограниченную выборку* (см. рис 2.6).

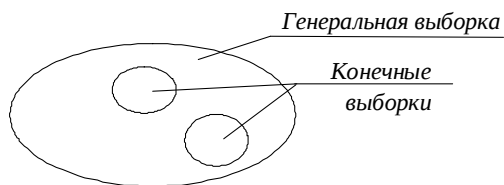


Рис. 2.6. Соотношение понятий генеральной и конечной выборок.

Генеральная совокупность охватывает наиболее широкий круг лиц, имеющих необходимые для целей исследования характеристики, и которые в силу наличия этих характеристик, в принципе, могли бы стать испытуемыми (объектами непосредственных измерений). Очевидно, что в психологических исследованиях измерению подлежат определенные признаки, параметры, показатели тех особенностей, механизмов и закономерностей психических явлений, которые проявляются в поведении и деятельности различных субъектов, являющихся носителями психики. Чем больше таких субъектов привлечено в качестве испытуемых, т. е. чем больше объем выборки, тем больше можно получить эмпирических данных, относящихся к исследуемым психическим явлениям, тем более достоверные выводы можно делать на основании результатов исследования. Однако провести измерительные процедуры в полном объеме на всей генеральной выборке испытуемых не реально. Поэтому в практических исследованиях ограничиваются конечной (или ограниченной) выборкой. Иногда такую выборку называют реальной или просто выборочной совокупностью.

В психологических исследованиях выделение испытуемых, которых можно было бы включить в генеральную или ограниченную выборки, является довольно сложной и ответственной задачей. Основная проблема в формировании генеральной выборки – это проблема адекватности выборки целям исследования. Необходимо, прежде всего, четко сформулировать объект и предмет исследования. Затем необходимо проанализировать и выделить совокупность зависимых и независимых переменных (факторов, признаков, параметров), характеризующих предмет исследования и подлежащих измерению. И лишь затем на основании выделенных признаков определиться с тем, какая категория испытуемых должна войти в генеральную совокупность. Далее необходимо решить вопрос о той группе испытуемых, которые могут и должны войти в состав непосредственно обследуемой выборки. Если состав и параметры генеральной выборки оп-

ределяются, по существу, характерными признаками (параметрами) предмета исследования, адекватного целям исследования, то состав и параметры отдельной выборочной совокупности определяются, главным образом, двумя факторами: репрезентативностью выборки и необходимым уровнем достоверности результатов измерений.

Репрезентативность (в пер. с фр. *representatif* - показательный) есть свойство выборочной совокупности представлять характеристики генеральной совокупности. Выборка испытуемых может считаться репрезентативной, если ее основные параметры, являющиеся в исследовании независимыми переменными (или факторами) (например, распределение по полу, возрасту, социальному статусу, уровню интеллекта и пр.) не отличаются от соответствующих параметров всей генеральной совокупности испытуемых данной категории.

Если выборочные совокупности испытуемых нерепрезентативны, полученные на них результаты исследований не могут быть распространены на генеральную совокупность, и считаются действительными лишь для данных отдельных выборок.

Необходимость в формировании репрезентативных выборок особо значима при проверках и обосновании валидности используемых в исследовании тестов, опросников и других психодиагностических процедур, а также при их стандартизации и построении стандартизованных диагностических шкал. Выборки, используемые в последнем случае, часто называются выборками стандартизации или нормативными выборками¹⁰.

Другой, не менее важной проблемой (наряду с проблемой репрезентативности), возникающей при формировании выборочной совокупности испытуемых, является проблема объема выборки. Сколько испытуемых должно принять участие в психологическом исследовании? Каким должен быть объем выборки? Очевидно, что чем больше объем выборки, тем больше измерений необходимо проделать. А это требует, соответственно, больших материальных, финансовых, и временных затрат. При малом объеме выборки начинают возникать сомнения в достоверности результатов проведенного исследования. Следовательно, при проведении исследования необходимо определить некоторое оптимальное значение объема выборки.

¹⁰ Анастаси А. Психологическое тестирование: Книга 1. –М.: Педагогика, 1982.

Как будет показано ниже, это значение непосредственно связано с уровнями достоверности и точности результатов измерений, которых мы хотим достичь.

Проводя психологические исследования на ограниченной (конечной) выборке испытуемых, мы, тем не менее, всегда предполагаем, что с некоторой точностью и вероятностью полученные нами результаты можно распространить и на генеральную выборку, т. е. на всю совокупность испытуемых. Тот факт, что выборка ограничена и порой неслучайна, обуславливает появление в результатах измерений, так называемых, *ошибок репрезентативности* выборки.

Различают *ошибки репрезентативности* двух видов:

- о систематические,
- о случайные.

Систематические ошибки возникают в том случае, если не выполнены условия случайности отбора объектов исследования, включаемых в выборку. Не случайность отбора (формирования выборки) приводит к смещению средних значений отдельных выборок относительно среднего или, так называемого, истинного значения, характеризующего генеральную выборку.

Смещение среднего значения $\bar{X}_1$ относительно истинного значения $X_{ист}$ называется систематической ошибкой и равно:

$$d_1 = X_{ист} - \bar{X}_1 \quad (2.5)$$

Случайные ошибки возникают за счет того, что для анализа всей совокупности используется только часть ее.

Допустим, что мы исключили систематические ошибки и $\delta = 0$. Тогда наличие случайных ошибок можно отразить следующим образом:

$$X_{ист} = X_{cp_i} + e_i \quad (2.6)$$

где X_{cp_i} - среднее значение показателя x i -ой выборки,

e_i - ошибка измерения истинного значения показателя $X_{ист}$ на i -ой выборке.

Множество отдельных выборок даст множество возможных значений x_{cp} . Если просуммировать средние значения отдельных k выборок, то получим:

$$\sum_{i=1}^k x_{cp_i} = k * x_{уст} - \sum_{i=1}^k e_i \quad (2.7)$$

Поскольку ошибки выборки случайны, их сумма $\sum e_i = 0$. В связи с этим можно записать:

$$\sum x_{cp_i} = k * x_{уст} \quad (2.8)$$

откуда находится:

$$x_{уст} = \frac{\sum x_{cp_i}}{k} \quad (2.9)$$

Получается, что среднее значение всех значений x_{cp_i} находится вблизи истинного значения $x_{уст}$ всей генеральной совокупности.

Статистическая совокупность значений показателя

В результате измерений какого-либо показателя исследуемого психического явления, проведенных на группе (выборке) из N испытуемых, мы получаем, так называемые, первичные или исходные значения показателя. Эти первичные значения показателя, измеренные у различных испытуемых, чаще всего представлены неупорядоченной последовательностью чисел, в которой сразу бывает трудно разобраться. Поэтому их подвергают специальной обработке, в результате которой вся совокупность первичных (исходных) данных приводится к виду, удобному для анализа.

Набор чисел, полученный в результате измерений, принято называть *статистической совокупностью* значений измеряемого показателя, а числа, характеризующие изменение (вариацию) в отдельных значениях показателя, - *вариантами*.

Если упорядочить совокупность вариантов, т. е. отличающихся значений показателя (признака), в убывающем или в возрастающем порядке, то

получим, так называемый, *ранжированный* ряд. Если этот ряд дополнить рядом чисел, показывающих, как часто встречаются отдельные значения вариант, то получим ряд, который называется *вариационным*.

Рассмотрим в качестве примера распределение по возрасту некоторой группы (выборки) испытуемых, где возраст рассматривается как измеряемый показатель (признак).

Статистическую совокупность значений измеряемого признака (возраста) для выборки из 50 испытуемых можно представить следующим образом:

Распределение испытуемых по возрасту

Таблица 2.1

№ исп.	1	2	3	4	5	6	7	8	9	10
Возраст	17	18	18	18	19	18	20	20	19	18
№ исп.	11	12	13	14	15	16	17	18	19	20
Возраст	18	21	19	22	23	18	19	19	19	21
№ исп.	21	22	23	24	25	26	27	28	29	30
Возраст	21	18	18	18	18	22	19	18	20	18
№ исп.	31	32	33	34	35	36	37	38	39	40
Возраст	19	18	20	19	21	20	22	18	19	21
№ исп.	41	42	43	44	45	46	47	48	49	50
Возраст	19	19	22	23	19	20	21	22	17	19

Ранжированный ряд отличающихся значений показателя (т. е. ряд вариант) в возрастающем варианте для статистической совокупности, представленной в таблице 2.1, будет выглядеть следующим образом:

Значение вариант в распределении испытуемых по возрасту

Таблица 2.2

Показатель	В а р и а н т а						
Возраст	17	18	19	20	21	22	23

Вариационный ряд той же статистической совокупности исходных данных может быть представлен в различной форме:

а) в форме таблицы:

Вариационный ряд распределения испытуемых по возрасту

Таблица 2.3

Возраст	Кол-во испытуемых данного возраста	Доля испытуемых данного возраста (%)
17	2	4
18	15	30
19	14	28
20	6	12
21	6	12
22	5	10
23	2	4
Итого:	50	100

б) в графической форме:

- в форме *гистограммы*:

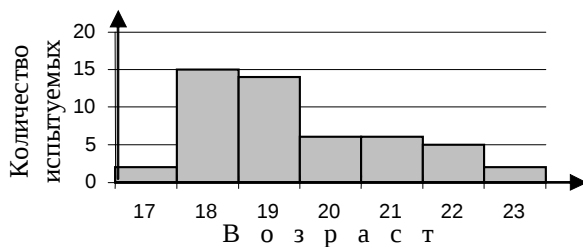


Рис. 2.7. Гистограмма распределения испытуемых по возрасту.

- в форме *полигона частот* (кривой распределения соответствующих долей испытуемых, имеющих заданное значение показателя:

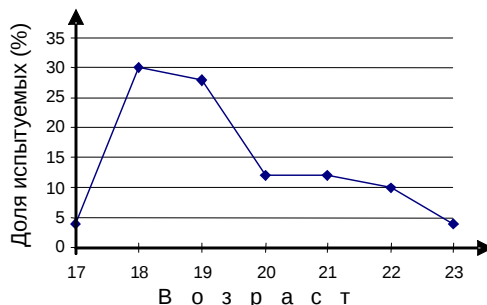


Рис. 2.8. Полигон частот (кривая распределения) значений возраста.

- в форме *кумуляты* (кривой распределения накопленных частот):

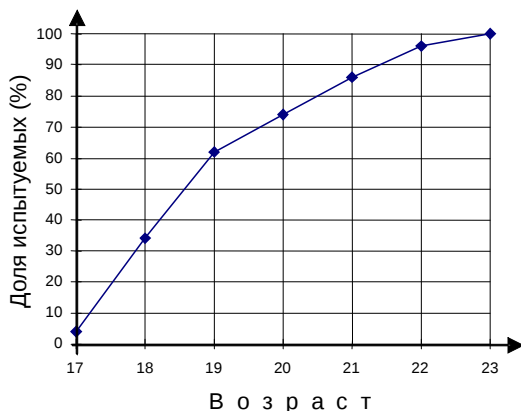


Рис. 2.9. Кумулята распределения испытуемых по возрасту.

Ордината кривой кумулятивного распределения показывает долю испытуемых, имеющих значения показателя (признака) не превышающие соответствующее значение, заданное по оси абсцисс, т. е. значение соответствующей варианты.

В зависимости от вида вариации различают дискретные и интервальные вариационные ряды.

В *дискретных вариационных рядах* каждое отдельное (дискретное) значение варианты представлено как самостоятельное значение. Примером дискретного вариационного ряда является рассмотренный нами вариационный ряд, характеризующий количество или долю испытуемых определенного возраста.

В *интервальных вариационных рядах* исходные варианты объединяются в отдельные группы. Значения вариантов, входящих в группы, образуют определенный интервал. При этом частотные показатели измеряемого признака, соответствующие отдельным вариантам, входящим в образованный интервал, суммируются. Такая процедура "сжатия" вариационного ряда используется для еще более компактного представления статистической совокупности исходных данных.

Первым шагом при построении интервального вариационного ряда является выбор принципа, по которому будут объединяться исходные варианты измеряемого признака. Выбор этого принципа зависит от степени

однородности рассматриваемой совокупности вариант. Если совокупность вариант однородна, то при построении интервального ряда используют принцип равных интервалов, если совокупность вариант неоднородна - используют принцип неравных интервалов. При этом стремятся добиться качественной однородности объектов внутри интервалов. Иногда при построении интервального ряда используют типологический принцип, когда границы интервала выделяют новое качество признака.

Например, если среди испытуемых каждая возрастная категория представлена в равной мере, то ряд можно строить равноинтервальный. Если отдельные возрастные категории испытуемых не представлены, или представлены не равномерно, то ряд целесообразно строить с разными интервалами, каждый из которых охватывает достаточно однородную совокупность испытуемых. Если каждая такая совокупность испытуемых отличается наличием некоторого особого качества (например, молодые, среднего возраста и пожилые), то можно говорить о типологическом принципе построения интервального ряда.

В рассматриваемом нами примере дискретный вариационный ряд можно преобразовать в равноинтервальный, если, например, сгруппировать варианты попарно: 17-18, 19-20, 21-22, 23-24 (см. таблица 2.4).

Равноинтервальный вариационный ряд

Таблица 2.4

Возрастной интервал	17-18	19-20	21-22	23-24
Кол-во испытуемых	17	20	11	2
Доля испытуемых	34	40	22	4

Однако, как видно из таблицы 2.4, в соответствии с такой группировкой количество испытуемых, охватываемых каждым интервалом, оказывается неравномерным. Следовательно, можно построить интервальный ряд, исходя из другого принципа, например, типологического. В частности, можно выделить четыре типа испытуемых, относящихся к разным возрастным категориям: а) до 18 лет, б) от 18 до 20 лет, в) от 20 до 23 лет, г) старше 23 лет. В этом случае новый вариационный ряд будет выглядеть следующим образом (см. Таблица 2.5):

Таблица 2.5

Тип испытуемого	I	II	III	IV
Возрастной интервал	до 18	18-19	20-22	старше 22
Кол-во испытуемых	2	29	17	2
Доля испытуемых (%)	4	58	34	4

Вторым шагом при построении интервального ряда является выбор величины интервала. Для равноинтервального ряда величина интервала (h) определяется соотношением:

$$h = \frac{R}{k} \quad (2.10)$$

где: R - размах вариации, определяемый как разница между максимальным ($x_{\max}$) и минимальным ($x_{\min}$) значением вариант;

k - количество интервалов, численное значение которого зависит от количества вариант n :

$$k = 1 + 3,2 * \lg n \quad (2.11)$$

В результате оптимальная величина интервала для совокупности из n вариант определяется соотношением:

$$h = \frac{x_{\max} - x_{\min}}{1 + 3,2 * \lg n} \quad (2.12)$$

Для нашего примера $x_{\max} = 23$, $x_{\min} = 17$, количество вариант $n = 7$. Следовательно, оптимальное количество интервалов $k = 1 + 3,2 * \lg 7 = 3,7$.

Округляем значение k до $k = 4$ и находим величину интервала $h = (23-17)/4=1,5$.

В результате можно образовать следующий равноинтервальный ряд:
1) от 17 до 18,5; 2) от 18,5 до 20; 3) от 20 до 21,5; 4) от 21,5 до 23.

Для интервального ряда с разными интервалами величины интервалов определяются требованиями однородности объектов внутри интервала.

Построение вариационных рядов является наиболее простой процедурой и первым шагом в изучении совокупности статистических данных, полученных в результате измерения. Для более глубокого изучения результатов измерений необходимы обобщенные количественные показате-

ли, характеризующие общие свойства статистической совокупности значений измеряемого признака. Эти показатели, нивелируя случайные индивидуальные отклонения, показывают общие тенденции в развитии изучаемого признака (явления) и позволяют сравнивать результаты измерений, полученные на различных выборках испытуемых.

В психодиагностической практике используются две группы характеристик статистической совокупности значений измеряемого признака:

- 1) *меры уровня или средние*,
- 2) *меры рассеяния*.

В связи с тем, что статистические совокупности значений измеряемого признака, полученные при измерениях по различным шкалам, имеют

Меры уровня некоторую специфику, они описываются различными **(средние)** статистическими характеристиками, в том числе и различными "средними" или, по-другому, "мерами центральной тенденции"¹¹.

Наиболее употребительными являются три вида "средних":

- о *мода*,
- о *медиана*,
- о *арифметическая средняя*.

Общим для каждого вида "средних" является требование качественной однородности статистической совокупности и наличие ясного смысла в значении "средней".

В совокупности значений признака, полученной при измерении по *шкале наименований*, "средняя" характеризует наиболее часто встречаемый класс объектов. Другими словами, в качестве "среднего" здесь выступает наименование класса, который является наиболее наполненным, к которому отнесено наибольшее количество объектов. Поскольку наименование класса в шкале наименований, по существу, является вариантом, "средняя" - это наиболее часто встречаемая в статистической совокупности варианта. Данный вид "среднего" называется "*модой*".

¹¹ см. Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии. – М.: Прогресс, 1976. С.58.

Если дискретные значения возраста испытуемых в рассматриваемом нами примере (см. Таблица 2.3) считать наименованиями отдельных классов (типов) испытуемых, то модой или "средней" величиной распределения испытуемых по возрасту будет 18 лет.

В практике психологических исследований иногда возникают ситуации, когда в статистической совокупности исходных данных имеется не одна, а две наиболее часто встречаемые варианты. В таком случае считается, что распределение исходных данных имеет две моды и называется бимодальным. Если варианты с наибольшими частотами встречаемости расположены рядом, т. е. являются смежными, что может иметь место в шкалах интервалов, мода определяется как среднее значение этих двух вариантов.

Другой "средней" или "мерой центральной тенденции" является "медиана", которая по определению есть значение признака, по обе стороны от которого располагается равное количество измеряемых объектов (испытуемых), т. е. половины объектов имеют значение признака меньше "среднего", другая половина - больше "среднего". Поскольку определение медианы связано с понятием "больше - меньше", данная форма среднего может иметь место в шкалах более высокого уровня, чем шкала наименований, т. е. в шкалах порядка, интервалов и отношений.

Для нахождения медианы удобно воспользоваться кумулятой распределения совокупности исходных данных (см. рис. 2.9). Достаточно провести линию от точки 50 на оси ординат (которая соответствует 50 % или половине испытуемых) до пересечения с кумулятой и затем опустить ее вниз до пересечения с осью абсцисс, на которой отложены варианты значений измеряемого признака x . Значение варианты в точке пересечения линии с осью абсцисс и есть искомое значение медианы.

В рассматриваемом нами примере значение медианы для совокупности значений возраста испытуемых составляет 18,6 лет. Именно это значение возраста делит всю выборку испытуемых на две равные части – на тех, кто старше и кто младше 18,6 лет. Изложенный порядок нахождения медианы показан на рис. 2.10.

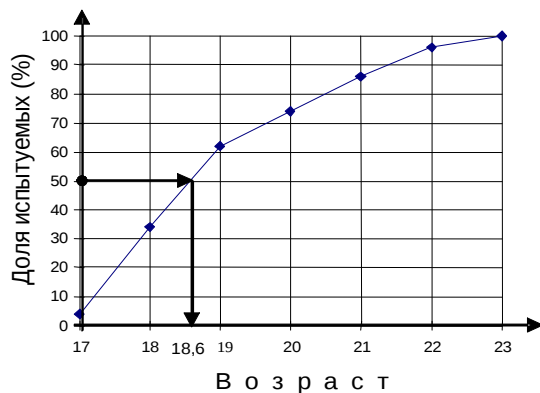


Рис. 2.10. Нахождение медианы по кумуляте распределения исходных данных.

В *шкале интервалов* кроме нахождения "средних" в форме моды и медианы можно вычислить "среднее" как "средне арифметическое", которое определяется по формуле:

$$x_{cp} = \frac{\sum_{i=1}^N x_i}{N} \quad (2.13)$$

где: x_i - индивидуальное значение показателя i -го испытуемого;

N - общее количество испытуемых;

или

$$x_{cp} = \frac{\sum_{j=1}^k (n_j * x_j)}{\sum_{j=1}^k n_j} \quad (2.14)$$

где: n_j - количество испытуемых, имеющих значение показателя, равное значению j -ой варианты;

x_j - значение j -ой варианты;

k - число отдельных вариантов.

В нашем примере среднее арифметическое значение возраста испытуемых x_{cp} равно 19,4 года.

Меры рассеяния

Для характеристики статистической совокупности знание "средней" не всегда является достаточным. При одном и том же значении "средней" совокупности могут иметь различные распределения индивидуальных значений.

Рассмотрим, например, два следующих распределения испытуемых по возрасту:

1-ое распределение: 31 32 36 40 41

2-ое распределение: 14 15 15 66 70 .

Значения "средних" обоих распределений одинаковы ($x_{cp} = 36$), однако разброс индивидуальных значений в первой группе испытуемый значительно меньше, чем во второй. Поэтому, для характеристики совокупности данных наряду со "средними" используются различные меры рассеяния (или меры разброса).

Основными *мерами рассеяния* являются:

□ *размах вариации* - R , определяемый как разница между наибольшим и наименьшим значением показателя:

$$R_x = x_{\max} - x_{\min} \quad (2.15)$$

Несмотря на простоту вычислений, существенным недостатком этой меры рассеяния является то, что характеристика всей совокупности данных дается из рассмотрения всего лишь двух ее крайних значений. Кроме того, значение R существенно зависит от случайных колебаний выборки.

□ *среднее арифметическое отклонение* - σ_x , вычисляемое по формуле:

$$\sigma_x = \frac{\sum_{i=1}^N |x_i - x_{cp}|}{N} \quad (2.16)$$

При вычислении среднего арифметического отклонения используется каждое значение показателя из общей совокупности, что, очевидно, повышает значимость σ_x как интегрального показателя, характеризующего не просто степень разброса отдельных значений показателя, но и указывающего степень однородности исследуемой совокупности.

□ *дисперсия* - S^2 , численное значение которой может быть вычислено по следующим эквивалентным формулам:

$$S_x^2 = \frac{\sum_{i=1}^N (x_i - x_{cp})^2}{N - 1} \quad (2.17)$$

или

$$S_x^2 = \frac{\sum_{i=1}^N x_i^2 - N \cdot x_{cp}^2}{N - 1} \quad (2.18)$$

Как и в среднем арифметическом отклонении $\bar{x}$, в дисперсии S_x^2 учитывается каждое значение показателя. Деление на $N - 1$, а не на N в указанных формулах обусловлено тем, что именно в этом случае значение S_x^2 , полученное на ограниченной выборке, оказывается в большей степени несмещенной оценкой истинного значения $S_{ист}^2$, которое характеризует генеральную выборку.

Смысл введения и использования дисперсии в качестве меры рассеяния связан с возможностью рассмотрения статистической совокупности значений показателя, как некоторого "множества". Значение "средней арифметической" этого множества есть некоторая его центральная точка. Отдельное i -ое значение показателя, входящее в состав множества, удалено от центральной точки на некоторое расстояние d_i , определяемое следующим соотношением:

- для одномерного множества $d_i = x_i - x_{cp}$
- для двухмерного множества $d_i^2 = (x_i - x_{cp})^2 + (y_i - y_{cp})^2$
- для n -мерного множества $d_i^2 = \sum_{j=1}^n (x_i^j - x_{cp}^j)^2$,

где j - индекс шкалы множества.

Очевидно, что для общего случая в качестве меры удаления от центральной тенденции (среднего значения) берется не значение $d_i = x_i - x_{cp}$

или $d_i = |x_i - x_{cp}|$, а значение $d_i^2 = (x_i - x_{cp})^2$.

Соответственно среднее значение удаленности точек множества от центральной тенденции будет характеризоваться значением дисперсии

$$S_x^2 = \frac{\sum_{i=1}^N (x_i - x_{cp})^2}{N - 1}.$$

□ *стандартное или среднее квадратичное отклонение* - S (сигма), определяемое как значение квадратного корня из дисперсии:

$$S_x = \sqrt{S_x^2} = \sqrt{\frac{\sum_{i=1}^N (x_i - x_{cp})^2}{N - 1}} \quad (2.19)$$

Рассмотренные меры рассеяния (размах R , среднее арифметическое отклонение dx , дисперсия S^2 , среднее квадратичное отклонение S) имеют определенную размерность, т. е. выражены в каких-то единицах измерения. Однако очень часто для сравнения разных статистических совокупностей исходных данных возникает необходимость в, так называемых, безразмерных или относительных показателях, характеризующих совокупности независимо от используемых в них единиц измерения. Такими показателями являются:

□ *относительное среднее арифметическое отклонение* Δ :

$$\Delta = \frac{dx}{x_{cp}} \quad (2.20)$$

□ *коэффициент вариации* V :

$$V = \frac{S}{x_{cp}} \quad (2.21)$$

Коэффициент вариации V и относительное отклонение Δ позволяют, в частности:

- сравнивать вариацию одного и того же признака у разных групп испытуемых (например, при дифференциации и индивидуализации обучения, когда возникает необходимость формирования гомогенных групп учащихся);
- выявлять смещение значений одного и того же признака у одной и той же группы в разное время (при проведении психокоррекционной работы);

– сопоставлять вариацию разных признаков у одних и тех же групп испытуемых (например, при выделении наиболее значимых признаков, необходимых для учета в процессе индивидуализации и дифференциации обучения).

**Стандартизация
значений
показателя**

Любое множество значений показателя со средним значением x_{cp} и средне квадратичным отклонением S можно преобразовать в другое множество, среднее значение которого будет равно 0, а средне квадратичное отклонение – равно 1. Необходимость в таком преобразовании возникает, например, в тех случаях, когда требуется сопоставить значения показателей, имеющих разную размерность, т. е. измеренных по шкалам с различными единицами измерения (например, интеллект в баллах, память в количестве слов, глазомер в миллиметрах, распределяемость внимания в секундах и т. д.). Такое преобразование носит название стандартизации или нормирования и позволяет получить, так называемые, *стандартизованные* или *нормированные* значения исходных данных. Стандартизация (нормирование) осуществляется по формуле:

$$z_i = \frac{x_i - x_{cp}}{S} \quad (2.22)$$

Смысл этой процедуры состоит в том, что на шкале интервалов вводится новая единица измерения, равная значению средне квадратичного отклонения S . В связи с чем, исходное значение показателя x_i и его отклонение от среднего ($x_i - x_{cp}$) также начинают измеряться в единицах этого средне квадратичного отклонения.

Значения z , полученные в результате стандартизации (нормирования), могут быть отрицательными и дробными, что не всегда удобно. Но и эти значения можно преобразовать и привести к более удобной форме. Для этого следует воспользоваться допустимым в шкале интервалов линейным преобразованием значений показателя типа:

$$z = a + b * \frac{x_i - x_{cp}}{S_x} \quad (2.23)$$

где a и b - произвольные коэффициенты.

Из наиболее известных шкал, образованных путем указанной процедуры стандартизации исходных значений показателя, можно отметить:

а) шкалу Гилфорда, построенную им для оценки интеллекта:

$$IQ = 100 + 15 * \frac{x_i - x_{cp}}{S_x} \quad (2.24)$$

б) шкалу Векслера, которая была построена для тех же целей:

$$IQ = 10 + 3 * \frac{x_i - x_{cp}}{S_x} \quad (2.25)$$

в) шкалу общего назначения Мак-Колла, известную под именем шкалы Т-баллов:

$$T = 50 + 10 * \frac{x_i - x_{cp}}{S_x} \quad (2.26)$$

Возможности использование такого рода шкал в психодиагностической практике можно продемонстрировать на примере построения, так называемого, профиля личности – графика, на котором отражены разнообразные индивидуально-психологические характеристики испытуемого.

Допустим, что на некоторой выборке испытуемых были проведены измерения особенностей познавательных процессов, например, воображения по методике "круги Гилфорда", вербального мышления по тесту Векслера, а также ряда личностных свойств, например, общительности и самостоятельности методом экспертных оценок по десяти балльной шкале. Допустим далее, что нас интересует испытуемый А, который имеет следующие числовые значения измеренных показателей:

- ☐ воображение 13
- ☐ мышление 92
- ☐ общительность 6
- ☐ самостоятельность ... 6

Очевидно, что на основании одних лишь индивидуальных значений показателей трудно сказать, что именно и насколько у испытуемого развито. Нужны эталоны или какие-то значения показателей, которые могли бы служить в качестве критериев для сравнения. Обычно в качестве таких критериев выступают средние арифметические значение x_{cp} и средние квадратичные отклонения S_x , получаемые на группе испытуемых. Соответствующие данные для нашего примера приведены в таблице 2.6.

Показатель	Исходное индивид. значение x_i	Средне арифметич. по группе x_{cp}	Средне квадратич. отклонение S_x	Стандарт. значение T – баллы
1. Воображение	13	10,2	2,3	62
2. Мышление	92	103	12,4	41
3. Общительность	6	5,4	1,3	55
4. Самостоят-ть	6	7,7	1,6	39

Сравнивая в таблице 2.6 исходные индивидуальные значения показателей x_i со средними по группе, можно сделать вывод, что у нашего испытуемого показатели воображения достаточно высокие ($13 > 10,2$), уровень мышления несколько снижен ($92 < 103$). Общительность и самостоятельность при одинаковых исходных оценках (по 6 баллов) находятся на разных уровнях по отношению к соответствующим средним групповым оценкам: общительность более высокая ($6 > 5,4$), а самостоятельность более низкая ($6 < 7,7$). Однако остаются еще, по крайней мере, два вопроса:

- во-первых, насколько сильно отличаются индивидуальные значения испытуемого от средних групповых значений;
- во-вторых, какое из измеренных качеств по сравнению с другими качествами развито у испытуемого больше, а какое меньше.

На все эти вопросы можно найти ответы, используя описанные выше процедуры стандартизации исходных данных. В качестве примера в таблице 2.6 приведены результаты преобразований исходных значений показателей в соответствующие значения стандартной шкалы Мак-Колла или в значения Т-баллов.

Поскольку значения Т-баллов безразмерны, их можно представить в единой системе координат в форме графика, который представляет собой профиль индивидуально-психологических особенностей рассматриваемого испытуемого (см. рис. 2.11).

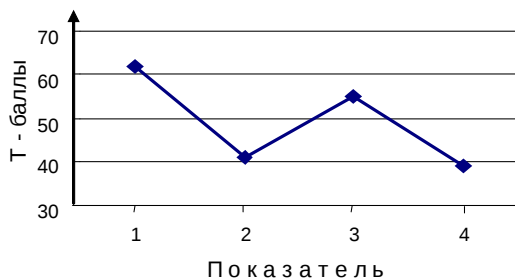


Рис. 2.11. Профиль индивидуально-психологических особенностей испытуемого.

Из графика видно, что у нашего испытуемого уровень воображения (показатель 1) имеет очень высокое значение по сравнению с другими испытуемыми ($T_1 = 62 \gg 50$), и в то же время данное свойство более развито по сравнению с другими измеренными свойствами (значение $T_1=62$ является наибольшим среди других значений T). В наименьшей степени развитым по сравнению с другими свойствами (и с другими испытуемыми) оказывается самостоятельность (значение $T_4 = 39 \ll 50$ и является наименьшим из всех). При одинаковых экспертных оценках общительности и самостоятельности (показатели 3, 4), оказывается, что уровень развития общительности у испытуемого намного выше, чем уровень развития самостоятельности, и занимает второе место (после воображения) в структуре индивидуально-психологических характеристик.

Понятие нормального распределения

Если в гистограмме распределения или полигоне частот (см. рис. 2.7, 2.8) уменьшать интервалы и одновременно увеличивать число испытуемых, то в пределе вместо ломаной линии получится плавная кривая линия. Эта линия называется *кривой распределения*. Формы кривой распределения могут быть самые разные.

Рассмотрим в качестве примера кривую распределения испытуемых по возрасту, представленную на рис. 2.8. Нетрудно убедиться, что количество испытуемых, возраст которых выше и ниже значения "моды", равного 18 лет, неодинаково.

Испытуемых до 18 лет в анализируемой выборке всего лишь 34 %, остальные испытуемые старше. Очевидно, такую асимметрию нельзя признать "нормой" для генеральной совокупности. Соответственно и само распределение нельзя назвать "нормальным".

Наличие асимметричности распределения проявляется и при рассмотрении медианы, которая по определению делит выборку пополам. Как было показано выше, медиана данного распределения равна 18,6 года. Однако из рис. 2.8. видно, что распределение испытуемых по вариантам не одинаково и явно асимметрично. Опять же это не нормально. Еще более асимметрично распределение испытуемых относительно средне арифметического значения возраста, которое составляет 19,4 года. И это тоже нельзя признать нормальным для генеральной выборки. "Нормальным", очевидно, следует признать такое распределение, которое симметрично относительно своих средних, и сами средние (мода, медиана и среднее арифметическое значение) равны между собой.

Понятие нормального распределения и условия, по которым эмпирическое распределение можно считать "нормальным", были сформулированы известным математиком де Муавром (1667-1754) в результате решения задачи, связанной с подбрасыванием монет. Суть задачи состоит в следующем: "Какова вероятность, что из 10 подбрасываний монеты "орел" выпадет 0 раз, 1 раз, 2 раза и т. д. до 10 раз?".

График эмпирического распределения вероятностей, полученного путем многократных подбрасываний монет, показан на рис. 2.12.

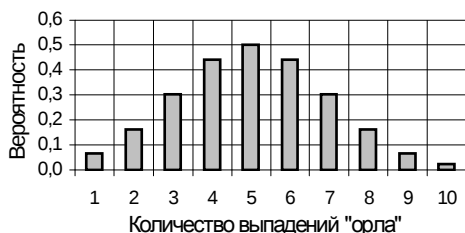


Рис. 2.12. Эмпирическое распределение вероятности выпадения "орла" при 10-ти бросаниях монеты.

Де Муавру удалось показать, что уравнение кривой, наиболее точно описывающей полученное эмпирическое распределение вероятностей, имеет следующий вид:

$$f = \frac{1}{s \cdot \sqrt{2\pi}} e^{-\frac{(x-m)^2}{2s^2}} \quad (2.27)$$

где: f - значение вероятности, x - значение варианты, m - значение "средней" статистической совокупности, S - стандартное отклонение.

Кривая, описываемая уравнением 2.27, это и есть кривая нормального распределения.

В общем случае имеется множество кривых нормального распределения, отличающихся значениями μ и σ . Для нормированных (стандартных) значений z_i измеряемого признака, когда статистическая совокупность характеризуется параметрами: $\mu = 0$, $\sigma = 1$, кривая нормального распределения имеет вид симметричной колоколообразной кривой (см. рис. 2.13).

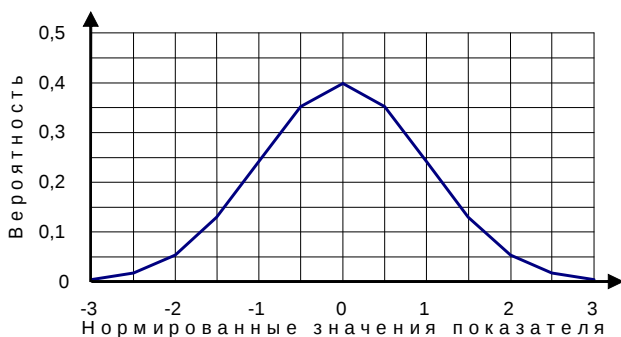


Рис. 2.13. Кривая нормального распределения.

Кривая, представленная на рис. 2.13, носит название *единичной нормальной кривой*, поскольку площадь под ней, т. е. площадь между кривой и осью значений z_i измеряемого показателя, точно равна 1. Данное свойство единичной нормальной кривой позволяет рассматривать площади отдельных областей нормального распределения как доли или проценты целого.

Именно для нормального распределения характерны, в частности, следующие соотношения:

- а) площадь под кривой, в пределах значений показателя $z = \mu \pm 1\sigma$, равна 68,26 % всей площади;
- б) площадь под кривой, в пределах значений показателя $z = \mu \pm 2\sigma$, равна 95,44 % всей площади;
- в) площадь под кривой, в пределах значений показателя $z = \mu \pm 3\sigma$, равна 99,72 % всей площади.

Смысл значения площади под кривой состоит в том, что оно указывает, какой процент испытуемых из статистической совокупности имеет значение показателя в заданных пределах. Если обратиться к кумулятив-

ной форме представления данных статистической совокупности, то трудно обнаружить полное соответствие значения кумуляты и площади под кривой нормального распределения. Поэтому для нормального распределения можно утверждать, что доля испытуемых, имеющих значение показателя z в пределах $\mu \pm 1\sigma$, составляет $\approx 68\%$, и далее по аналогии.

Статистические распределения, отличающиеся от нормального, характеризуются, главным образом, двумя параметрами:

- о степени *асимметрии*,
- о степени *выпуклости*.

Степень асимметрии статистического распределения эмпирических данных можно определить путем вычисления соответствующего *показателя асимметрии* A :

$$A = \frac{\sum (x_i - x_{cp})^3}{N \cdot s^3} \quad (2.28)$$

Для нормальных распределений $A = 0$. При левосторонней асимметрии $A > 0$. При правосторонней асимметрии $A < 0$.

В случае ненормальных распределений, когда возникает та или иная асимметрия, средние значения, определяемые модой (mo), медианой (md) и средне арифметической (x_{cp}), не совпадают. При левосторонней асимметрии: $x_{cp} > md > mo$. При правосторонней асимметрии: $x_{cp} < md < mo$. В связи с этим для оценки степени асимметрии иногда используют значение разницы ($x_{cp} - md$). В этом случае величина A определяется по формуле:

$$A = \frac{3 \cdot (x_{cp} - md)}{s} \quad (2.29)$$

Степень выпуклости распределения определяется *показателем эксцесса* E и зависит от соотношения срединной (ограниченной значениями показателя $x_{cp} \pm \sigma$) и крайних частей статистической совокупности. Если доля срединной части больше 68%, эксцесс положительный и распределение более выпуклое, если доля срединной части меньше 68%, эксцесс отрицательный и распределение более пологое.

Численное значение эксцесса определяется по формуле:

$$E = \frac{\sum (x_i - x_{cp})^4}{N \cdot s^4} - 3 \quad (2.30)$$

Для нормальных распределений $E = 0$.

**Нормализация
значений
показателя**

Распределение совокупности данных, получаемых в результате психодиагностики, может иметь различную форму и к тому же отличную от нормальной. В этом случае применение некоторых математических методов и статистических критериев при сравнении и оценке таких распределений становится некорректным. Даже простое сравнение арифметических средних для распределений, имеющих различную асимметрию, вообще говоря, не является корректным, поскольку соответствующие статистические совокупности неоднородны и неоднородны по-разному. Некорректным также является использование коэффициентов корреляции Пирсона, а также *t*-критерия Стьюдента.

Чтобы добиться сопоставимости статистических совокупностей, имеющих распределения различной формы, необходимо воспользоваться нелинейным преобразованием исходных данных и привести форму распределения к какому-то одному виду. Обычно в качестве эталона принимается форма нормального распределения. Процедура приведения распределения к нормальному виду носит название *нормализации*, а преобразованные исходные данные называются *нормализованными*.

Нормализованные значения могут быть найдены с помощью таблиц, в которых приводится процент случаев (процентили) различных отклонений в единицах σ от среднего значения для случая нормального распределения измеряемого показателя (см. Приложение 1). Сначала определяется процент испытуемых в исследуемой выборке с тем же или более высоким исходным значением показателя, т. е. вычисляются соответствующие значения кумуляты распределения. Затем этот процент отыскивается в таблице нормального распределения частот и по нему находится соответствующее значение нормализованного стандартного показателя Z .

Далее нормализованные значения Z путем соответствующего линейного преобразования могут быть преобразованы в другие более удобные для последующего использования и анализа числа, например в *T*-баллы шкалы Мак-Колла.

В качестве примера процедура преобразования исходных значений измеряемого показателя к нормальному виду на основе значений кумуляты представлена в таблице 2.7.

Распределение испытуемых по возрасту

Таблица 2.7

Исходные значения возраста (лет)	17	18	19	20	21	22	23
Кол-во испытуемых	2	15	14	6	6	5	2
Доля испытуемых (%)	4	30	28	12	12	10	4
Кумулята распределения (%)	4	34	62	74	86	96	100
Нормализованные значения возраста в единицах z	-1,75	-0,41	0,31	0,64	1,08	1,75	3,09
Значения Т-баллов	32	46	53	56	61	68	81

Нормализованные значения возраста в единицах z находятся с помощью специальных статистических таблиц (см. Приложение 1) в следующей последовательности:

1. Для исходного значения показателя находим соответствующее значение кумуляты. (Например, для значения показателя $x_1 = 17$ находим значение кумуляты $K_1 = 4$).

2. В статистической таблице нормального распределения (см. Приложение 1) находим процентиль P_1 , значение которого равно значению кумуляты K_1 . (В нашем случае $P_1 = K_1 = 4\%$).

3. Находим по статистической таблице нормализованное значение z_1 , соответствующее значению процентилля P_1 . (При $P_1 = 4\%$ искомое значение $z_1 = -1,75$).

Далее аналогичным образом находятся нормализованные значения z_i для других исходных значений возраста x_i .

Нормализованные значения z_i можно далее преобразовать в значения Т-баллов, используя соотношение:

$$T_i = 50 + 10 * z_i \quad (2.31)$$

В итоге, в результате двойного преобразования, исходные значения возраста испытуемых, распределение которых имеет левостороннюю асимметрию, могут быть преобразованы в безразмерные значения Т-баллов, распределение которых является нормальным со статистическими параметрами $T_{cp} = 50$, $S = 10$.

Характер кривой нелинейного преобразования, позволившего осуществить процедуру нормализации исходных данных и перевести исходное ненормальное распределение значений возраста в нормальное распределение значений Т-баллов, можно видеть из графика, представленного на рис. 2.14.

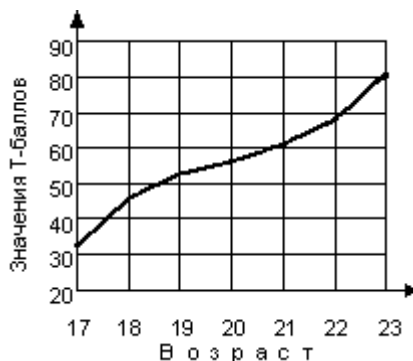


Рис. 2.14. Кривая нелинейного преобразования, лежащая в основе нормализации исходных данных.

Другим достаточно известным нелинейным преобразованием, позволяющим осуществлять нормализацию данных, является преобразование с помощью 9-ти балльной *шкалы станайнов*. Название шкалы образовано путем сокращения выражения "standard nine", т. е. стандартная девятка.

Шкала станайнов является интервальной шкалой с неравномерным распределением доли испытуемых, имеющих соответствующие значения станайнов. Численные значения этих долей в процентах при нормальном распределении значений измеряемого показателя представлены в таблице 2.8. Здесь же приведены значения накопленных частот для каждого из станайнов для построения кумуляты нормального распределения.

Шкала станайнов

Таблица 2.8

Станайн (<i>Ст</i>)	1	2	3	4	5	6	7	8	9
Доля испытуемых (%)	4	7	12	17	20	17	12	7	4
Накопленные частоты (%)	4	11	23	40	60	77	89	96	100

Нормализация исходных значений показателя x_i , распределение которых не является нормальным, осуществляется с помощью шкалы станайнов следующим образом.

Сначала строится вариационный ряд эмпирического распределения значений показателя x_i в форме кумуляты, которая представляется либо в виде графика (см. рис. 2.9), либо в виде таблицы численных значений соответствующих накопленных частот (см. Таблица 2.9).

Значения накопленных частот эмпирического распределения

Таблица 2.9

<i>Возраст испытуемого (лет)</i>	17	18	19	20	21	22	23
<i>Накопленные частоты (%)</i>	4	34	62	74	86	96	100

Далее путем соотнесения значений накопленных частот из таблицы 2.9 со значениями частот, приведенными в таблице 2.8, для каждой варианты x_i исходного распределения находятся наиболее соответствующие значения станайнов.

Значения станайнов для исходных значений вариантов нашего эмпирического распределения представлены в таблице 2.10.

*Возраст испытуемых в значениях станайнов
нормального распределения*

Таблица 2.10

<i>Возраст испытуемого (лет)</i>	17	18	19	20	21	22	23
<i>Значение станайнов</i>	1	3	5	5	6	8	9

**Понятие
доверительного
интервала**

В результате психодиагностики, когда мы, используя соответствующие психодиагностические средства, измеряем какое-либо свойство объекта (в частности, индивидуально-психологические особенности испытуемого), мы получаем значение показателя X_i , которое, в общем случае, отличается от истинного значения показателя действительно присущего объекту.

Другими словами, при измерении всегда имеется определенная ошибка измерения, которая может обуславливаться как погрешностью измерительной процедуры, так и возможной нестабильностью состояния диагностируемого объекта (испытуемого). В связи с этим исходные значения показателя X_i , полученные при неоднократном измерении, как правило, образуют определенный интервал, внутри которого с некоторой вероятностью находится истинное значение показателя $X_{ист}$.

Интервал, внутри которого может находиться истинное значение показателя, называется *доверительным интервалом*.

Степень вероятности, с которой истинное значение показателя $X_{ист}$ будет находиться в доверительном интервале, носит название *доверительной вероятности*.

В психодиагностической практике обычно используют следующие значения доверительной вероятности: 90%, 95% и 99%. Это означает, что при измерениях стремятся достичь такой точности и надежности измерений, при которой вероятность ошибки (α) не превышала бы, соответственно, 10%, 5% и 1%.

Если статистическая совокупность исходных значений показателя X_i имеет нормальное распределение, то 68% всех значений показателя лежит в диапазоне $X_{cp} \pm S$. Для нормального распределения можно утверждать, что с вероятностью 68% истинное значение среднего $X_{cp}^{ист}$ лежит в диапазоне $X_{cp} \pm S$.

С помощью таблицы нормального распределения можно определить границы интервалов, при которых вероятность нахождения $X_{cp}^{ист}$ внутри интервалов будет составлять, соответственно, 90%, 95% и 99%. Границы интервалов выражаются в долях S_x и составляют:

для вероятности 90% - $\pm 1,64 S_x$

для вероятности 95% - $\pm 1,96 S_x$

для вероятности 99% - $\pm 2,58 S_x$.

Указанные соотношения справедливы для случая нормального распределения значений показателя X_i по всей генеральной совокупности. Однако на практике мы всегда имеем дело с конечными или ограниченными выборками. Если в рамках генеральной выборки выделить k конечных выборок, то измерения на каждой i -ой отдельной конечной выборке дадут свои значения S_x^i . При наличии множества конечных выборок значение S_x для генеральной выборки находится из соотношения:

$$S_x^2 = \sum_{i=1}^k (S_x^i)^2 \quad (2.32)$$

Если конечные выборки являются репрезентативными, то для них выполняется условие равенства всех значений S_x^i , и в этом случае соотношение 2.27 будет выглядеть следующим образом:

$$S_x^2 = k * (S_x^i)^2 \quad \text{или} \quad S_x^i = \frac{S_x}{\sqrt{k}} \quad (2.33)$$

Используя данное соотношение, мы можем записать соотношение, определяющее границы доверительного интервала для истинного значения показателя X_i^{ucm} i -ой конечной выборки, в следующем виде:

$$X_i^{ucm} = X_i \pm \frac{t \cdot S_x}{\sqrt{k}} \quad (2.34)$$

где значение t выбирается в соответствии с требуемым уровнем доверительной вероятности α : при $\alpha = 0,10$; $0,05$ и $0,01$ значение t равны, соответственно, $1,64$; $1,96$ и $2,58$.

Как видно из приведенных соотношений, достоверность измерений, определяемая величиной доверительного интервала, существенным образом зависит от числа k : чем больше k , тем уже доверительный интервал и, соответственно, более достоверные измерения.

Если перейти от рассмотрения отдельных конечных выборок к рассмотрению отдельных испытуемых и учесть, что в этом случае число k соответствует количеству испытуемых N , то соотношение 2.34 можно использовать для определения границ доверительного интервала, в котором находится истинное значение диагностируемого признака x у i -ого испытуемого.

**Определение
объема
выборки**

Соотношение 2.34 позволяет не только находить границы доверительного интервала истинных значений показателей, но и рассчитывать объем выборки испытуемых, необходимый для получения заданной достоверности результатов измерений.

Для минимально допустимой погрешности измерения (e_i) относительное значение погрешности измерения (Δx) определяется соотношением: $\Delta x = \frac{e_i}{x_{cp}}$. Поскольку e_i можно рассматривать как определенный доверительный интервал, в рамках которого находится истинное значение измеряемого показателя, можно записать:

$$\frac{t \cdot S_x}{\sqrt{N}} \leq e_i \quad (2.35)$$

Учитывая, что $e_i = \Delta x \cdot x_{cp}$, на основе уравнения 2.35 нетрудно получить соотношение для определения объема выборки:

$$N \geq \left(\frac{t \cdot S_x}{\Delta x \cdot x_{cp}} \right)^2 \quad (2.36)$$

Учитывая далее, что $\frac{S_x}{x_{cp}}$ есть коэффициент вариации V_x , можно записать:

$$N \geq \left(\frac{t \cdot V_x}{\Delta x} \right)^2 \quad (2.37)$$

При значениях погрешности измерения $\Delta x = 0,1$ (10%), величины коэффициента вариации $V_x = 0,3$ (что достаточно типично для измерений в психологии) и доверительной вероятности 95%, которой соответствует значение $t = 1,96$, минимальный объем выборки в соответствии с соотношением 2.32 составляет $N \geq 34,6$. Увеличение доверительной вероятности до 99% ($t = 2,58$) при заданной погрешности измерения можно достичь за счет увеличения объема выборки до $N \geq 59,9$.

Снижение погрешности измерения, например, с 10% до 5% (что означает увеличение достоверности измерения в два раза) достигается, как следует из соотношения 2.37, за счет четырехкратного увеличения объема

выборки. То есть, чтобы достичь 95% доверительной вероятности ($t = 1,96$) при погрешности ($\Delta x = 0,05$), объем выборки должен составлять не менее $34,6 * 4 = 138,4$. Чтобы при тех же условиях ($\Delta x = 0,05$; $V_x = 0,3$) достичь доверительной вероятности 99% ($t = 2,58$), потребуется выборка объемом не менее 235,9.

Как видно из соотношения 2.37, объем необходимой выборки существенным образом зависит от вариативности значений показателя, характеризуемой коэффициентом вариативности V_x . Если показатель, который мы собираемся диагностировать на предполагаемой выборке испытуемых, имеет значительную степень вариативности, объем выборки, при котором обеспечивается необходимая степень достоверности результатов измерения, должен быть значительно увеличен: при увеличении вариативности в два раза, объем выборки увеличивается в 4 раза.

**Выявление
различий
в значениях
показателя**

До сих пор мы рассматривали свойства отдельных статистических совокупностей, взятых сами по себе. Однако очень часто возникает необходимость в сравнении двух или нескольких совокупностей, полученных при измерении одного и того же показателя на разных выборках или на одной и той же выборке, но в разных условиях или обстоятельствах. Вопрос, например, может ставиться так: "Есть ли разница между значениями показателя x , полученными на выборке мальчиков и на выборке девочек?". Другой задачей на сравнение значений показателя может быть решение вопроса о том, произошли ли изменения (или сдвиги) в значениях показателя на выборке испытуемых после проведения с ними какого-либо психологического тренинга или оказания на них определенного психологического воздействия. В обоих случаях решение о наличии или отсутствии различий в значениях показателя принимается на основании анализа двух или нескольких статистических совокупностей исходных данных и в соответствие с определенным, так называемым, статистическим критерием.

Статистический критерий – это правило, обеспечивающее принятие выдвигаемой гипотезы как истинной и отклонение как ложной с высоким уровнем достоверности.

Гипотезы, проверяемые в соответствии со статистическим критерием, носят название *статистических гипотез* и подразделяются на:

- нулевые,
- альтернативные,
- ненаправленные,
- направленные.

Нулевая гипотеза (H_0) – это гипотеза об отсутствии значимых различий между двумя и более совокупностями значений показателя.

Альтернативная гипотеза (H_1) – это гипотеза о наличии значимых различий между совокупностями значений показателя.

Гипотеза называется *ненаправленной*, если ее формулировка предполагает определение лишь отличий или не отличий без указания направления этих отличий (в большую или меньшую сторону). Если гипотеза содержит указание на направление отличий, например, χ_{cp1} не превышает χ_{cp2} (для гипотезы H_0) или χ_{cp1} превышает χ_{cp2} (для гипотезы H_1), то гипотеза рассматривается как *направленная*.

Поскольку при проверке статистической гипотезы мы располагаем лишь выборочными совокупностями исходных данных, полученных на отдельных ограниченных выборках, всегда существует риск совершить ошибку: либо отклонить верную гипотезу, либо принять неверную. Поэтому в психодиагностике при формулировке выводов и принятии каких-либо решений всегда указывается степень их достоверности (или уровень значимости).

Статистические критерии делятся на:

- *параметрические*,
- *непараметрические*.

К *параметрическим критериям* относятся критерии, в вычислении которых используются параметры распределения (средние значения и средние квадратичные отклонения, например, *t*-критерий Стьюдента, *F*-критерий Фишера).

Преимуществом параметрических критериев является то, что они позволяют прямо оценить различия в средних или в дисперсиях сравниваемых выборок. Однако их применение является корректным лишь по отношению к статистическим совокупностям, имеющим нормальные или близкие к нормальным распределения. Кроме того, процедура их вычисления в сравнении с непараметрическими критериями несколько более сложная.

К *непараметрическим критериям* относятся критерии, не требующие для расчета знание параметров распределения. Для их вычисления достаточно иметь значения частот или рангов. Перечень непараметрических критериев с указанием задач психодиагностики, которые решаются с их помощью, приведены в таблице 2.11¹².

Применение непараметрических критериев в психодиагностике

Таблица 2.11

Задача психодиагностики	Условия психодиагностики	Статистический критерий
1. Выявление различий в уровне исследуемого признака	2 выборки испытуемых	Q – критерий Розенбаума; U – критерий Манна-Уитни.
	3 и более выборок испытуемых	S – критерий тенденций Джонкира; H – критерий Крускала-Уоллиса.
2. Оценка сдвига значений исследуемого признака	2 замера на одной и той же выборке испытуемых	T – критерий Вилкоксона; G – критерий знаков.
	3 и более замеров на одной выборке испытуемых	S_r^2 – критерий Фридмана; L – критерий тенденций Пейджа.
3. Анализ изменений признака под влиянием контролируемых условий	под влиянием одного фактора	S – критерий тенденций Джонкира; L – критерий тенденций Пейджа; Однофакторный дисперсионный анализ Фишера.
	под влиянием двух факторов одновременно	Двухфакторный дисперсионный анализ Фишера.

¹² см. Сидоренко Е.В. Методы математической обработки в психологии. –СПб.: ООО "Речь", 2002.

Непараметрические критерии вычисляются проще, но они позволяют оценить лишь средние тенденции, например, ответить на вопрос, ~~лишь в выборке А среднее значение показателя выше, чем в выборке Б – более низкие.~~ Тем не менее, если непараметрические критерии показывают наличие различий - значит, эти различия действительно есть. Однако если различия с помощью непараметрических методов не выявляются, то это еще не говорит о том, что их нет. Возможно, они есть, но для их выявления необходимо воспользоваться более мощными параметрическими критериями.

В связи с тем, что особенности применения непараметрических критериев очень подробно и замечательно изложены в указанной выше работе Е.В. Сидоренко, мы рассмотрим особенности применения параметрических критериев.

В качестве примера, рассмотрим психодиагностическую задачу определения различий в объеме памяти детей разных возрастных групп, или имеющих разный уровень интеллекта. Возможна и другая, можно сказать, "обратная" задача: выделить в рамках одной выборки испытуемых различные группы, отличающиеся по уровню измеряемого показателя. Простейшим случаем такой задачи является деление выборки на две группы: а) "более хорошие" и б) "более плохие" или на три группы: а) "более успешные", б) "средние" и в) "менее успешные" и т. д.

Определение достоверности различий Допустим, имеется две выборки испытуемых объемом n_1 и n_2 со средними значениями показателя x , равными x_{cp1} и x_{cp2} , соответственно. Если средние значения x_{cp1} и x_{cp2} не равны между собой, то очевидно, что между ними есть определенная разница, которая указывает на наличие различий между выборками по показателю x . Однако возникает вопрос, насколько эти различия можно считать существенными, и насколько достоверны эти различия? Если статистические совокупности значений показателя в обеих выборках имеют распределения близкие к нормальным и для них известны среднеквадратичные отклонения S_1 и S_2 , достоверность различий между средними можно определить с помощью параметрического критерия, называемого t -критерием Стьюдента, численное значение которого вычисляется по формуле:

$$t = \frac{x_{cp1} - x_{cp2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (2.38)$$

Различия в средних значениях x_{cp1} и x_{cp2} считаются достоверными, если вычисленное или эмпирическое значение t -критерия Стьюдента превышает критическое значение $t_{кр}$, соответствующее заданному уровню достоверности. Значение $t_{кр}$ находится из специальных таблиц распределения критических значений t -критерия Стьюдента (см. Приложение 2) в зависимости от числа степеней свободы n , вычисляемого по формуле $n = n_1 + n_2 - 2$, и желаемого уровня достоверности α .

Определение достоверности различий в средних значениях показателя различных выборок с использованием t -критерия Стьюдента является весьма полезным и целесообразным, в частности, при формировании общей выборки испытуемых. Если, например, при сравнении средних значений какого-либо показателя в группе мальчиков и группе девочек наблюдаются отличия, но при проверке по t -критерию Стьюдента достоверность этих отличий не подтверждается, то, в принципе, эти две группы испытуемых по измеряемому признаку могут быть объединены и отнесены к одной общей выборке.

Обратная задача, т. е. задача деления общей выборки на подгруппы с достоверно отличающимися значениями измеряемого показателя, осуществляется следующим образом.

При делении выборки на две подгруппы в первую подгруппы включаются все испытуемые, имеющие значения показателя $x_i > x_{cp}$, во вторую – те, у которых $x_i < x_{cp}$.

При делении выборки на три подгруппы в качестве центральной подгруппы выделяются испытуемые, для которых значения показателя x_i находится в пределах $x_{cp} \pm S_x$. Во вторую подгруппу выделяются испытуемые со значениями показателя, превышающими $x_{cp} + S_x$. Третью группу образуют испытуемые, у которых значение показателя ниже $x_{cp} - S_x$.

Значения показателей центральной подгруппы испытуемых обычно рассматриваются в качестве нормы; второй и третьей подгрупп – соответственно, выше и ниже нормы.

Графически это может быть представлено следующим образом (см. рис. 2.15):

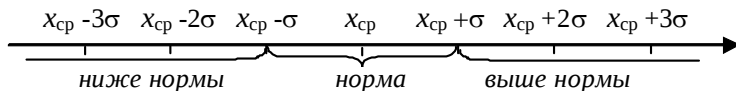


Рис. 2.15. Схема деления выборки испытуемых на три подгруппы.

Однако при таком делении численность испытуемых в подгруппах оказывается неравномерной, в центральной подгруппе около 68 %, в других подгруппах по 16 %. Чтобы подгруппы были более равномерны по численности испытуемых, необходимо в качестве нормы принять диапазон значений показателя от $x_{ср} - 0,5S$ до $x_{ср} + 0,5S$. В таком случае в центральной подгруппе будет около 38 % испытуемых, в двух других – по 31 %.

Использование средне квадратичного отклонения S_x позволяет не только выделять отдельные группы испытуемых, но и, фактически, преобразовывать измерительные шкалы интервального типа либо в номинальные, либо в шкалы порядка. Наименования подгрупп можно рассматривать как некоторые названия (или обозначения) отличающихся классов испытуемых в полном соответствии с правилами, принятыми в номинальной шкале. Количество испытуемых, имеющих значения показателя в пределах соответствующих интервалов, можно рассматривать как показатель наполненности этих классов, и можно говорить о значении моды как о мере центральной тенденции для данного распределения классов испытуемых. Поскольку выделенные группы (или классы) испытуемых отличаются по уровню признака и их можно расположить в порядке возрастания или убывания признака, каждой группе испытуемых в соответствии с ее положением относительно других групп можно присвоить определенный ранг. Например, при выделении пяти групп испытуемых и наличии, соответственно, пяти градаций значений измеряемого показателя группам можно присвоить ранги от 1 до 5.

Как было показано выше, использование статистического критерия типа t -критерия Стьюдента позволяет определять наличие или отсутствие различий между двумя статистическими совокупностями значений показателя, полученными при измерении одного и того же показателя на двух различных выборках (или при повторных измерениях на одной и той же выборке). А как быть, если на одной и той же выборке испытуемых получены результаты измерений по двум разным показателям x и y ? Очевидно, что вопрос о соотношении средних значений показателей не имеет смысла, поскольку показатели разные. Однако уместен вопрос о соотношении индивидуальных значений показателей, который решается с использованием процедуры стандартизации исходных данных или соответствующих процедур корреляционного анализа. Если исходные данные получены при измерении разных показателей на разных выборках, то вопрос о связи и соотношениях между ними не имеет смысла.

2.5. Методы корреляционного анализа

Корреляционный анализ представляет собой мощное средство математической обработки и анализа результатов психодиагностики, полученных при измерении различных показателей на одной и той же выборке испытуемых или различных оценок разными выборками испытуемых, но одной и той же совокупности диагностируемых объектов.

Корреляция как мера статистической связи показателей

Само слово "корреляция" в переводе с латинского означает "связь". Следовательно, корреляционный анализ – это анализ различных связей между измеряемыми показателями. Для обозначения степени или силы связи вводится понятие коэффициента корреляции. Очевидно, чем больше коэффициент корреляции, тем сильнее связь между анализируемыми показателями.

Существует два вида связей:

- функциональные,
- статистические.

При функциональной связи каждому значению одного показателя (x) всегда соответствует определенное значение другого (y): $x_i \Leftrightarrow y_i$.

При статистической связи одному и тому же значению одного показателя (x) могут соответствовать различные значения другого (y).

Наличие статистических связей между двумя и более показателями обуславливается, главным образом:

- многомерностью психологических явлений, подлежащих диагностике;
- наличием случайных и не учитываемых факторов, воздействие которых приводит к изменению в проявлениях диагностируемых показателей.

В результате множественного воздействия взаимно пересекающихся факторов связь между показателями проявляется не в каждом отдельном случае, как при функциональной связи, а только в тенденции и "в среднем". Выявление наличия и степени статистической связи между различными показателями и составляет основную задачу корреляционного анализа. В задачу корреляционного анализа входит также оценка значимости или достоверности связи между показателями. Кроме того, анализ статистической связи предполагает выявление формы связи. Однако решение данной задачи осуществляется с помощью методов регрессионного анализа.

Один из важных вопросов, возникающих в изучении взаимосвязей показателей, это вопрос о том, можно ли наличие связи между ними рассматривать как проявление зависимости, как наличие влияния одного показателя на другого. Например, при наличии связи между показателями успеваемости, памяти и мышления можно ли говорить о том, что память и мышление влияют на успеваемость, или, наоборот, что развития памяти и мышления зависит от успеваемости?

Понятие "зависимость" обычно используется в научных исследованиях для обозначения причинно-следственных отношений между событиями. Применительно к психодиагностике, где в качестве событий выступают отдельные значения диагностируемых показателей, наличие зависимости означало бы, что, получив при измерении показателя x значение $x=a$, мы с неизбежностью при измерении показателя y должны получить $y=b$. При этом первое событие ($x=a$) выступает как причина второго ($y=b$), а второе – как следствие первого. Зависимость y от x описывается некоторой функцией f , которая записывается следующим образом: $y = f(x)$.

В случае статистических связей одно и то же значение показателя x может сочетаться с разными значениями показателя y . Следовательно, обнаружение корреляционной связи между показателями не может трактоваться как проявление зависимости между ними. Наличие корреляций ни в коей мере не является свидетельством наличия причинно-следственных отношений или зависимостей между коррелируемыми показателями.

При использовании методов корреляционного анализа и определении соответствующих коэффициентов корреляции могут возникать ряд проблем, среди которых следует отметить следующие:

1) Статистические совокупности исходных данных получены при измерении показателей с использованием измерительных шкал разного типа. Например, пол испытуемых измеряется по шкале наименований, особенности памяти – по шкале порядка, а уровень интеллекта – по интервальной шкале.

2) Различные совокупности исходных данных имеют различные статистические параметры (различный объем выборки, различную вариативность значений показателя, различную форму распределения, в том числе далекую от нормального).

3) Связь между измеряемыми показателями может быть не только линейной, но и нелинейной.

Решение первой и частично третьей проблемы дает нам разнообразие в процедурах вычисления коэффициентов корреляции и, соответственно, разнообразие типов коэффициентов корреляции.

Решение второй и третьей проблемы приводит к формулированию условий, при которых применение тех или иных процедур корреляционного анализа является корректным, и к понятию достоверности коэффициентов корреляции.

В зависимости от используемых при проведении психодиагностики измерительных шкал, применяются разные методы корреляционного анализа и разные коэффициенты корреляции.

**Корреляция показателей
в шкалах интервалов**

Корреляция показателей, при измерении которых использовались интерваль-

ные шкалы, осуществляется путем вычисления коэффициента корреляции Пирсона.

*Коэффициент
корреляции
Пирсона*

Данный коэффициент относится к статистическим критериям параметрического типа, который позволяет определить не только наличие или отсутствие достоверной связи между показателями, но и характер этой связи, в частности, силу и направление связи. Следует отметить также, что коэффициент корреляции Пирсона – это коэффициент парной и линейной корреляции, т. е. указывает на наличие и степень лишь линейной связи между двумя показателя x и y .

При определении коэффициента корреляции Пирсона сопоставляются две статистические совокупности значений показателя x и y , полученные на одной и той же выборке испытуемых и при использовании одного и того же типа шкалы – шкалы интервалов.

Рассмотрим в качестве примера статистические совокупности исходных данных, полученных при измерении уровня интеллекта и успеваемости на выборке учащихся объемом $N = 12$. Исходные значения указанных показателей приведены в таблице 2.12.

*Значения показателей интеллекта и успеваемости
в группе учащихся*

Таблица 2.12

Показатель	Номер учащегося											
	1	2	3	4	5	6	7	8	9	10	11	12
Интеллект	120	112	110	120	103	126	113	114	106	108	128	109
Успеваемость	31	25	19	24	17	28	18	20	16	15	27	19

Анализируемые статистические совокупности имеют следующие параметры:

Интеллект: $x_{\text{ср}} = 114,1$ $S_x = 7,86$

Успеваемость: $y_{\text{ср}} = 21,6$ $S_y = 5,23$.

Вопрос, который касается корреляции показателей интеллекта и успеваемости, состоит в следующем: "Если испытуемый имеет более высо-

кий уровень интеллекта, то имеет ли он так же более высокий уровень успеваемости или такого соответствия не наблюдается?".

Характер связи между показателями интеллекта и успеваемости можно наглядно представить в виде диаграммы рассеяния (см. рис 2.16).

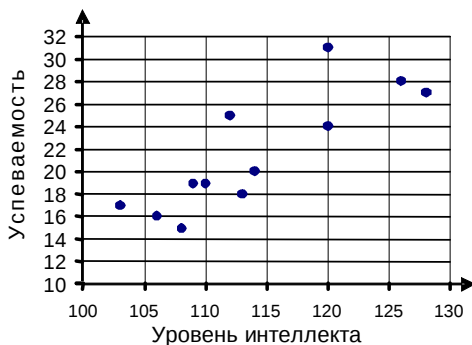


Рис. 2.16. Диаграмма рассеяния значений уровня интеллекта и успеваемости.

Положение учащегося относительно остальных в выборке по показателям x и y определяется величиной и знаком отклонений:

$$\Delta x_i = x_i - x_{cp} \quad \text{и} \quad \Delta y_i = y_i - y_{cp}.$$

Определим учащихся, имеющих уровень успеваемости и уровень интеллекта выше среднего. Если это будут одни и те же учащиеся – значит связь между успеваемостью и уровнем интеллекта есть. В противном случае следует признать, что связи нет.

Анализ таблицы 2.12 показывает, что более высокая успеваемость ($y_i > y_{cp}$) у учащихся под номерами: 1, 2, 4, 6, 11, более высокий интеллект ($x_i > x_{cp}$) у учащихся под номерами: 1, 4, 6, 11. Очевидно, что более высокий уровень успеваемости и более высокий уровень интеллекта имеют практически одни и те же учащиеся. Это позволяет высказать предположение о связи успеваемости с уровнем интеллекта. На наличие связи, причем линейной, указывает и характер распределения точек на диаграмме рассеяния (см. рис. 2.16) – точки расположены в среднем вдоль диагонали.

Чтобы определить степень связи, рассмотрим произведение $\Delta x_i * \Delta y_i$. Если между x и y существует прямая связь, т. е. большим значениям x_i соответствуют большие значения y_i и наоборот, то большин-

ство произведений $\Delta x_i * \Delta y_i$ будет положительным. Соответственно общая сумма произведений $\sum_{i=1}^N \Delta x_i * \Delta y_i$ будет большой и положительной. В случае отсутствия прямой связи некоторые произведения $\Delta x_i * \Delta y_i$ могут быть отрицательными, и тогда общая сумма произведений будет меньше и может быть равной нулю. Если все произведения $\Delta x_i * \Delta y_i$ будут отрицательными, это означает, что большим значениям x_i соответствуют меньшие значения y_i и наоборот. В этом случае имеет место, так называемая, обратная связь. И в том, и в другом случаях величина общей суммы произведений $\sum_{i=1}^N \Delta x_i * \Delta y_i$ будет указывать на степень связи. Что касается знака суммы, то он указывает на направление связи: если сумма положительная – связь прямая, если отрицательная – связь обратная.

Однако величина суммы $\sum_{i=1}^N \Delta x_i * \Delta y_i$ зависит от объема выборки, и это не позволяет сравнивать результаты, полученные на разных выборках испытуемых. Чтобы можно было сравнивать степень связи между показателями x и y в двух выборках, необходимо произвести усреднение суммы произведений по объему выборки, поделив ее на $N - 1$. В результате получаем:

$$S_{xy} = \frac{\sum_{i=1}^N (x_i - x_{cp})(y_i - y_{cp})}{N - 1} \quad (2.39)$$

где S_{xy} носит название *ковариации* показателей x и y .

Значение ковариации S_{xy} не зависит от средних значений x_{cp} и y_{cp} и от объема выборки N . Чтобы сделать его независимым также и от влияния стандартных отклонений S_x и S_y , необходимо разделить S_{xy} на соответствующие значения S_x и S_y . В результате мы получаем меру связи показателей x и y , которая называется *коэффициентом корреляции Пирсона* и обозначается r_{xy} :

$$r_{xy} = \frac{S_{xy}}{S_x \cdot S_y} \quad (2.40)$$

Вычисление численного значения коэффициента корреляции Пирсона по исходным значениям показателей X_i и Y_i осуществляется по формуле:

$$r_{xy} = \frac{N \cdot \sum x_i y_i - (\sum x_i) \cdot (\sum y_i)}{\sqrt{[N \cdot \sum x_i^2 - (\sum x_i)^2] \cdot [N \cdot \sum y_i^2 - (\sum y_i)^2]}} \quad (2.41)$$

Численные значения коэффициентов корреляции Пирсона находятся в пределах от +1,0 до -1,0. Общий смысл этих значений иллюстрируется на рис. 2.17.

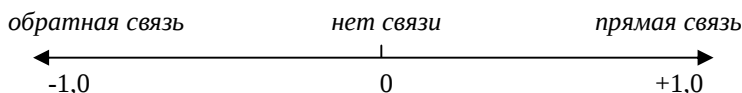


Рис. 2.17. Смысл численных значений коэффициента корреляции.

НАПОМИНАНИЕ! При проведении корреляционного анализа и вычислении коэффициентов корреляции необходимо помнить, что наличие взаимосвязи двух показателей не означает наличие между ними причинно-следственных отношений. Если показатели x и y взаимосвязаны, это не означает, что x влияет на y или наоборот.

Интерпретация коэффициента корреляции существенно зависит от параметров выборки и характера распределения исходных значений анализируемых показателей.

Допустим, что в результате психодиагностики были получены данные, свидетельствующие о том, что уровень тревожности у девочек более высокий по сравнению с мальчиками. Кроме того, известно, что и успеваемость у девочек более высокая, чем у мальчиков. Означает ли это, что между тревожностью и успеваемостью учащихся есть какая-то связь?

Возьмем выборку из 30 учащихся, среди которых 15 девочек и 15 мальчиков, и определим коэффициент корреляции двух показателей – тревожности (x) и успеваемости (y). Распределение исходных данных и

диаграмма рассеяния приведены, соответственно, в таблице 2.13 и на рис.2.18.

Исходные значения показателей тревожности x_i и успеваемости y_i мальчиков и девочек

Таблица 2.13

		М а л ь ч и к и														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i		4	5	3	4	3	5	3	4	4	5	4	5	3	4	5
y_i		18	24	20	16	18	20	23	16	16	18	24	16	18	22	22
		Д е в о ч к и														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i		7	5	7	6	8	7	6	6	5	7	6	5	8	7	6
y_i		25	28	22	18	30	27	30	28	27	30	30	28	30	22	24

Статистические совокупности значений показателей, представленных в таблице 2.13, имеют следующие параметры:

по группе мальчиков: тревожность $x_{cp}^M = 4,1$ $s_x^M = 0,64$

успеваемость $y_{cp}^M = 19,4$ $s_y^M = 8,83$;

по группе девочек: тревожность $x_{cp}^D = 6,4$ $s_x^D = 0,97$

успеваемость $y_{cp}^D = 26,6$ $s_y^D = 13,54$.

Вычисление коэффициента корреляции Пирсона для значений показателей x_i и y_i по общей выборке мальчиков и девочек дает значение $r_{xy}=0,621$. Казалось бы, исходное предположение о наличии связи между успеваемостью и уровнем тревожности подтверждается. Из диаграммы рассеяния (см. рис. 2.18) также видно, что связь есть, и она положительная

(прямая), так как распределение точек достаточно сильно вытянуто вдоль основной диагонали.

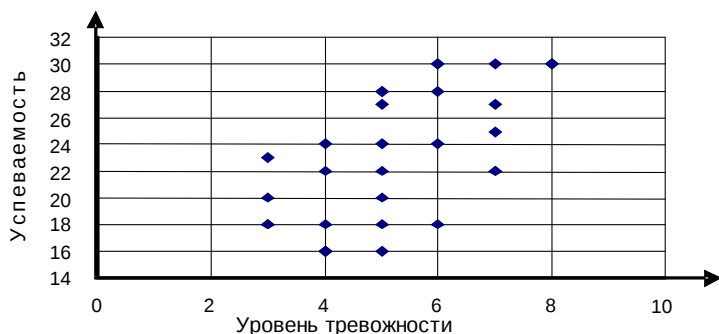


Рис. 2.18. Диаграмма рассеяния значений уровня тревожности и успеваемости.

Однако это не так, поскольку при проведении корреляционного анализа не было соблюдено условие однородности выборок.

Выборка, состоящая из девочек и мальчиков, неоднородна – средние значения уровня тревожности и успеваемости группы девочек и группы мальчиков различны, и это различие статистически значимо. Так, эмпирические значения t -критерий Стьюдента для различий средних между мальчиками и девочками (при уровне достоверности $\alpha = 0,01$, объемах выборки $n_m = n_d = 15$ и критическом значении $t_{кр} = 2,76$) составляют:

по уровню тревожности 7,66 ,

по уровню интеллекта 3,52 .

Эти данные свидетельствуют о достоверности различий между мальчиками и девочками и о необходимости выделения мальчиков и девочек в отдельные независимые выборки.

Корреляционный анализ взаимосвязи показателей тревожности и успеваемости, проведенный отдельно по группе мальчиков и группе девочек, дает следующие значения коэффициентов корреляции Пирсона:

для мальчиков $r_{xy}=0,048$,

для девочек $r_{xy}=0,067$.

Столь низкие значения коэффициентов корреляции свидетельствуют об отсутствии связей между анализируемыми показателями.

Однако, если значение коэффициента корреляции Пирсона равно нулю, это вовсе не означает, что между показателями x и y вообще нет никакой связи. При $r_{xy}=0$ можно говорить лишь об отсутствии линейной связи между показателями. Но между ними может быть и нелинейная связь. Однако коэффициент корреляции Пирсона наличие нелинейной связи не определяет. В связи с этим прежде, чем вычислять коэффициент корреляции Пирсона, целесообразно построить диаграмму рассеяния значений коррелируемых показателей и визуально проанализировать характер взаимосвязи между ними.

Проблема с интерпретацией коэффициента корреляции Пирсона возникает и в том случае, когда распределения значений показателей x и y имеют *разнонаправленную асимметрию*. В силу разнонаправленной асимметрии количество значений x_i , превышающих среднее значение x_{cp} , не уравнивается соответствующим количеством значений y_i , превышающих среднее значение y_{cp} . В результате даже при абсолютной связи показателей сумма произведений $\sum_{i=1}^N \Delta x_i \cdot \Delta y_i$ не может достигать максимально возможного значения, и коэффициент корреляции оказывается значительно меньше 1,0.

Проверка значимости коэффициента корреляции Поскольку значение коэффициента корреляции r_{xy} определяется на конечной выборке, в его определении имеется некоторая погрешность. Средне квадратичная ошибка для значений r_{xy} определяется выражением:

$$S_r = \frac{1 - r_{xy}^2}{\sqrt{N}} \quad (2.42)$$

где r_{xy} - коэффициент корреляции показателей x и y для генеральной выборки.

Очевидно, что чем больше значение r_{xy} отличается от r_{xy} , тем менее вероятно, что оно случайно. Величина нормированного отклонения z для r_{xy} определяется значением $z = \frac{r_{xy} - r_{xy}}{S_r}$.

Для проверки значимости коэффициента корреляции, т. е. достоверности его отличия от нуля, необходимо принять $r_{xy} = 0$ и вычислить эмпирическое значение $t_{эмп}$:

$$t_{эмп} = \sqrt{N} \cdot |r_{xy}| \quad (2.43)$$

Если $t_{эмп}$ больше критического значения $t_{кр}$, определяемого по таблице граничных значений t (см. Приложение 2) при уровне достоверности α и числе степеней свободы $n = N - 1$, то коэффициент корреляции r_{xy} считается значимым или достоверным на соответствующем уровне достоверности.

Для малой выборки $N < 30$ значение $t_{эмп}$ определяется из выражения:

$$t_{эмп} = \sqrt{N - 2} \cdot \frac{|r_{xy}|}{\sqrt{1 - r_{xy}^2}} \quad (2.44)$$

Например, для значения $r_{xy} = 0,621$, полученного на общей выборке $N = 30$, имеем $t_{эмп} = \frac{\sqrt{28} \cdot 0,621}{\sqrt{1 - 0,386}} = 4,19$. Критическое значение $t_{кр}$ для данного объема выборки и уровня достоверности $\alpha = 0,01$ составляет $t_{кр} = 2,74$. Поскольку $t_{эмп} > t_{кр}$, полученное значение коэффициента корреляции следует считать достоверным или значимым.

ВНИМАНИЕ! *Значимость коэффициента корреляции ни в коей мере не говорит о степени связи коррелируемых показателей. Она лишь указывает на то, что полученному значению коэффициента корреляции можно доверять с вероятностью ошибки не более α . Сама же степень связи характеризуется численным значением коэффициента корреляции. При больших объемах выборки значимыми могут быть и очень малые значения коэффициента корреляции.*

Соотношения (2.43) и (2.44) позволяют не только оценить степень значимости или достоверности коэффициента корреляции, но и опреде-

лить минимальный объем выборки, который мог бы обеспечить получение результатов корреляционного анализа с требуемой степенью достоверности.

Допустим, нам нужно обеспечить достоверность коэффициентов корреляции, имеющих значения $r_{xy} \geq 0,3$, на уровне $\alpha = 0,05$. Из соотношения (2.43) находим

$$N \geq \frac{t_{кр}^2}{r_{xy}^2} \quad (2.45)$$

Далее воспользуемся методом *итераций* (последовательных вычислений). На первом шаге примем $t_{кр}=1,96$, соответствующее $\alpha = 0,05$ и $N = \infty$. Подставляя в уравнение (2.45) значения $r_{xy} = 0,3$ и $t_{кр}=1,96$, получаем $N = 40,8$. На втором шаге находим по таблице (см. Приложение 2) при $N = 41$ значение $t_{кр}=2,02$. Вновь подставляя в уравнение (2.45) значение $r_{xy} = 0,3$ и новое значение $t_{кр}=2,02$, получаем $N = 43,3$. В принципе, на этом шаге вычисления можно закончить и окончательно принять значение минимального объема выборки, равное $N = 43$.

Аналогичным образом можно найти, что для того, чтобы обеспечить достоверность тех же значений коэффициента корреляции ($r_{xy} = 0,3$), но на уровне достоверности $\alpha = 0,01$, потребуется выборка объемом $N = 73$. Для достижения такой же степени достоверности, но значений коэффициента корреляции, начиная с $r_{xy} = 0,2$, потребуется выборка испытуемых объемом не менее $N = 167$.

Очевидно, что при больших выборках ($N > 200$) значимыми оказываются достаточно малые коэффициенты корреляции, но это говорит не о наличии связи, а о достоверности ее отсутствия. Если, например, коэффициент корреляции $r_{xy} = 0,3$ оказался значимым на высоком уровне значимости ($\alpha = 0,01$), это вовсе не означает, что между показателями x и y имеется сильная корреляционная связь. Это говорит лишь о том, что между показателями с очень высокой степенью достоверности ($\alpha = 0,01$) следует ожидать наличие слабой связи ($r_{xy} = 0,3$).

Множественный
коэффициент
корреляции

Иногда в психологических исследованиях возникает необходимость определить, насколько связан измеряемый показатель не с одним, а сразу с двумя другими показателями, которые выступают по отношению к исходному (результативному) показателю в качестве некоторых факторов. Для этих целей используют процедуру вычисления множественного (или совокупного) коэффициента корреляции

$R_{1.23}$, где индекс "1" означает номер результирующего признака, индексы "2", "3" – номера факторных признаков.

Величина $R_{1.23}$ определяется по формуле:

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}} \quad (2.46)$$

Численное значение множественного коэффициента корреляции $R_{1.23}$ лежит в диапазоне от 0 до 1,0. Если коэффициент $R_{1.23}$ равен нулю, это говорит о том, что связь между показателем "1" и двумя другими факторными показателями "2", "3", отсутствует. Равенство коэффициента $R_{1.23}$ единице свидетельствует о наличии функциональной связи.

Проверка значимости множественного (совокупного) коэффициента корреляции осуществляется с помощью F-критерия – критерия Фишера, эмпирическое значение которого вычисляется по формуле:

$$F_{\text{эмп}} = \frac{(N - 3) \cdot R_{1.23}^2}{2 \cdot (1 - R_{1.23}^2)} \quad (2.47)$$

Критическое значение F-критерия определяется по специальным таблицам в зависимости от объема выборки N и уровня достоверности или значимости α (см. Приложение 3). О наличии связи показателя "1" с другими показателями "2", "3" или о наличии совместного влияния факторов "2" и "3" на показатель "1" можно говорить в том случае, если $F_{\text{эмп}} > F_{\text{кр}}$.

Корреляция показателей в шкалах порядка

Рассмотренные нами коэффициенты корреляции Пирсона, характеризующие степень связи между двумя и более показателями, применимы в психодиагностических исследованиях при условии, что значения показателей измерены в шкалах интервалов и их распределение близко к нормальному. Однако достаточно часто, особенно при небольших выборках, распределение значений измеряемого показателя не является нормальным. Кроме того, в психодиагностике при измерении отдельных показателей очень часто используются процедуры ранжирования, в результате которых значения диагностируемых показателей представляются не в шкалах интервалов, в шкалах порядка. В связи с этим в психодиагностической практике широкое распространение получил метод корреляционного анализа, основанный на вычислении коэффициентов ранговой корреляции.

Ранговая корреляция представляет собой процедуру, которая позволяет определять меру взаимосвязи двух показателей, значения которых представлены в форме соответствующих рангов.

Допустим, мы имеем выборку из N испытуемых, отличающихся по показателям x и y . Измерение значений этих показателей с помощью шкал порядка (методом ранжирования испытуемых по степени выраженности признака) дает нам два ряда рангов r_i^x и r_i^y . Значения рангов изменяются в диапазоне от 1 до N .

Если каждый из испытуемых будет иметь одинаковые значения рангов по обоим показателям, то, очевидно, это будет свидетельствовать о наличии связи между показателями, поскольку каждому большему значению ранга r_i^x будет соответствовать большее значение r_i^y , и наоборот – меньшему значению r_i^x будет соответствовать меньшее значение r_i^y . В данном случае говорят о наличии прямой или положительной корреляционной связи. О наличии обратной или отрицательной связи можно говорить в том случае, когда большему значению r_i^x соответствует меньшее значение r_i^y , а меньшему значению r_i^x соответствует большее значение r_i^y . Мерой связи в ранговых корреляциях выступают коэффициенты ранговой корреляции.

Для определения связи между двумя показателями обычно используется коэффициент ранговой корреляции Спирмена¹³. Для определения согласованности в изменениях сразу нескольких показателей используется коэффициент множественной ранговой корреляции, который по-другому называется коэффициентом конкордации.

Коэффициент ранговой корреляции Спирмена	Использование коэффициента ранговой корреляции Спирмена получило наибольшее распространение в психодиагностике в силу простоты его вычисления и отсутствия необходимости в проверке распределения значений диагностируемого показателя на нормальность.
---	--

¹³ Кендэлл М. Ранговые корреляции. –М.: Статистика, 1975.

Численное значение коэффициента ранговой корреляции Спирмена находится по формуле:

$$r = 1 - \frac{6 \cdot \sum_{i=1}^N (r_i^x - r_i^y)^2}{N \cdot (N^2 - 1)} \quad (2.48)$$

Как и в случае парного коэффициента корреляции Пирсона r_{xy} , значение коэффициента ранговой корреляции Спирмена ρ меняется в диапазоне от +1 до -1.

Процедура вычисления коэффициента ранговой корреляции Спирмена для распределений рангов 10 испытуемых по показателям x и y в качестве примера представлена в таблице 2.14.

Пример вычисления коэффициента ранговой корреляции Спирмена

Таблица 2.14

Этап вычисления	Испытуемый									
	1	2	3	4	5	6	7	8	9	10
r_i^x	5	3	7	2	1	8	10	4	6	9
r_i^y	6	2	3	5	4	10	9	7	10	8
$r_i^x - r_i^y$	-1	1	4	-3	-3	-2	1	-3	-4	1
$(r_i^x - r_i^y)^2$	1	1	16	9	9	4	1	9	16	1
$\sum_{i=1}^{10} (r_i^x - r_i^y)^2$	1+1+16+9+9+4+1+9+16+1=67									
$\frac{6 \cdot \sum_{i=1}^{10} (r_i^x - r_i^y)^2}{10 \cdot (10^2 - 1)}$	$\frac{6 \cdot 67}{10 \cdot 99} = 0,41$									
$1 - \frac{6 \cdot \sum_{i=1}^n (r_i^x - r_i^y)^2}{n \cdot (n^2 - 1)}$	$1 - 0,41 = 0,59$									

В рассмотренном примере значение r для выборке из 10 испытуемых оказалось равным 0,59. Это означает, что между показателями x и y , измеренными на малой выборке с использованием метода ранжирования,

обнаружена определенная связь: она прямая по характеру (т. к. Γ имеет положительное значение) и сравнительно небольшая по степени (т. к. Γ значительно меньше 1,0).

Значимость полученного значения коэффициента ранговой корреляции, как и значимость коэффициента корреляции Пирсона, определяется с помощью t -критерия Стьюдента с использованием соотношений (2.43) или (2.44).

Используя для нашего примера соотношение (2.44), получаем $t_{эмп}=2,07$. Учитывая, что критическое значение t -критерия Стьюдента для уровня достоверности $\alpha = 0,05$ и $N = 10$ (см. Приложение 2) составляет $t_{кр}=2,23$ и превышает эмпирическое значение, можно сделать вывод, что между показателями x и y отсутствует достоверная корреляционная связь на данном уровне достоверности. Однако, поскольку для уровня достоверности $\alpha = 0,10$ $t_{кр}=1,81$ и в этом случае наше эмпирическое значение $t_{эмп}$ превышает $t_{кр}$, можно сделать сказать, что связь между показателями x и y существует, но достоверность связи невысокая. В психологических исследованиях этот уровень достоверности считается нижней границей приемлемости.

Определение коэффициента ранговой корреляции Спирмена по формуле (2.48) является корректным, если коррелируемые ряды рангов не содержат, так называемые, связанные ранги. Связанные ранги возникают в том случае, когда происходит преобразование значений показателей, измеренных в шкалах интервалов, в значения шкалы порядка. Другими словами, когда множество чисел (индивидуальных значений показателя) x_i подвергается процедуре ранжирования¹⁴. Испытуемые, имеющие одинаковые индивидуальные значения показателя, объединяются в группы, и каждому испытуемому из группы, в которую он входит, присваивается среднее значение тех рангов, на которые они все вместе претендуют.

В качестве примера рассмотрим распределение учащихся по уровню тревожности и успеваемости, взяв за основу исходные данные для мальчиков, приводимые ранее в таблице 2.13. Как следует из этой таблицы, у многих учащихся имеются одинаковые значения, как показателей тревожности, так и показателей успеваемости. Так, по показателю тревожности

¹⁴ Вообще говоря, ранжируются не числа, а испытуемые по их индивидуальным значениям показателя.

среди учащихся выделяются три группы с одинаковыми значениями показателя x_i : первую группу составляют 5 учащихся с уровнем тревожности $x_i=5$, вторую – 6 учащихся, для которых $x_i=4$, и третью – 4 учащихся с $x_i=3$.

Обозначим количество групп учащихся с одинаковыми значениями показателя буквой k , а число учащихся в группе – буквой t . В этом случае можно записать: $k=3$, $t_1=5$, $t_2=6$, $t_3=4$.

В соответствии с правилами для связанных рангов каждому учащемуся из первой группы присваивается ранг 3, т. е. среднее значение рангов 1, 2, 3, 4, 5. Учащимся из второй группы присваивается ранг 8,5 (среднее значение рангов 6, 7, 8, 9, 10, 11). Учащимся из третьей группы присваивается ранг 13,5 (среднее значение рангов 12, 13, 14, 15). Аналогичным образом находятся ранги учащихся по показателю успеваемости. По этому показателю количество групп учащихся, имеющих одинаковые значения показателя, равно 5.

По аналогии с обозначениями k и t , для распределения учащихся по показателю y введем обозначения l и u . В этом случае получаем: $l=5$, $l_1=2$, $l_2=2$, $l_3=2$, $l_4=4$, $l_5=4$.

Итоговые распределения рангов учащихся по обоим показателям приведены в таблице 2.15.

Распределения рангов учащихся r_i^x и r_i^y по значениям тревожности x_i и успеваемости y_i

Таблица 2.15

	Но м е р у ч а щ е г о с я														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i	4	5	3	4	3	5	3	4	4	5	4	5	3	4	5
r_i^x	8,5	3	13,5	8,5	13,5	3	13,5	8,5	8,5	3	8,5	3	13,5	8,5	3
y_i	18	24	20	16	18	20	23	16	16	18	24	16	18	22	22
r_i^y	9,5	1,5	6,5	13,5	9,5	6,5	3	13,5	13,5	9,5	1,5	13,5	9,5	4,5	4,5

Определение коэффициента ранговой корреляции Спирмена между показателями x и y для случая связанных рангов осуществляется по формуле:

$$r = 1 - \frac{6 \cdot \left[\sum_{i=1}^n (r_i^x - r_i^y)^2 \right] + T + U}{n \cdot (n^2 - 1)} \quad (2.49)$$

$$\text{где: } T = \frac{1}{12} \cdot \sum_{j=1}^k (t_j^3 - t_j) \text{ и } U = \frac{1}{12} \cdot \sum_{j=1}^l (u_j^3 - u_j) \quad (2.50)$$

Для нашего примера:

$$T = \frac{1}{12} \cdot (5^3 - 5 + 6^3 - 6 + 4^3 - 4) = 32,5$$

$$U = \frac{1}{12} \cdot [3 \cdot (2^3 - 2) + 2 \cdot (4^3 - 4)] = 11,5.$$

Подставляя численные значения T и U в формулу (2.49), получаем:

$$r = 1 - \frac{6 \cdot (476,5 + 32,5 + 11,5)}{15 \cdot 224} = 1 - 0,93 = 0,07.$$

Без учета связности рангов коэффициент корреляции Спирмена равен:

$$r = 1 - \frac{6 \cdot 476,5}{15 \cdot 224} = 1 - 0,85 = 0,15.$$

Ранее полученный для этой же выборки и для тех же значений показателей коэффициент корреляции Пирсона составил 0,048. Как видим, процедура вычисления коэффициента ранговой корреляции Спирмена с учетом связности рангов дает результаты, более согласованные с коэффициентом корреляции Пирсона, хотя и несколько завышенные.

Коэффициент множественной ранговой корреляции Множественный коэффициент ранговой корреляции, который называется также коэффициентом конкордации, предназначен для определения взаимосвязи сразу нескольких показателей.

Рассмотрим в качестве примера результаты определения ценностных ориентаций на выборке из 5 испытуемых, полученные методом ранжирования 10 личностных качеств по значимости (см. таблица 2.16).

*Распределение 10 качеств личности по значимости
на выборке из 5 испытуемых*

Таблица 2.16

		К а ч е с т в а л и ч н о с т и									
		1	2	3	4	5	6	7	8	9	10
		Честность	Доброта	Трудолюбие	Выносливость	Вежливость	Ум	Смелость	Самостоятельность	Настойчивость	Справедливость
Испытуемый	1	3	1	10	2	4	5	9	6	8	7
	2	3	6	2	4	10	1	7	5	8	9
	3	1	2	7	10	6	8	9	3	5	4
	4	2	8	4	6	10	7	9	3	5	1
	5	6	7	10	8	5	1	3	2	4	9
Сумма рангов		15	24	33	30	35	22	37	19	30	30

Испытуемые независимо друг от друга проранжировали 10 качеств личности по их значимости. Каждое индивидуальное распределение качеств по значимости характеризует систему ценностных ориентаций отдельного испытуемого. Возникает вопрос: "Насколько в данной группе испытуемых совпадают их индивидуальные системы ценностных ориентаций?".

Очевидно, что при совпадении ценностных ориентаций значимость каждого качества будет оцениваться всеми испытуемыми одинаково. Соответственно, каждое качество будет иметь одинаковые ранги: какое-то – ранги "1", какое-то – ранги "2" и т.д. В этом случае суммы рангов у отдельных качеств будут значительно отличаться: у наиболее значимого качества (где ранги "1") сумма будет равна 5, а наименее значимого (где ранги "10") – 50. В случае расхождения в оценках значимости отдельных качеств суммы рангов для каждого качества будут примерно одинаковы.

Среднее значение суммы рангов по полному рассогласованию систем ценностей испытуемых определяется соотношением:

$$s = \frac{N \cdot (m + 1)}{2}, \quad (2.51)$$

где: N – количество испытуемых, m – число качеств.

Чем больше фактическая сумма рангов отличается от среднего значения, тем более согласованы распределения рангов (в нашем примере это распределения качеств по значимости) и тем больше степень множественной корреляции рангов.

Для вычисления множественного коэффициента ранговой корреляции сначала необходимо для каждого j -го качества вычислить разность между фактической суммой рангов и средним значением, определяемым соотношением (2.51), т. е. $(\sum_{i=1}^N r_{ij} - s)$. Далее необходимо возвести полученные значения в квадрат и все квадраты просуммировать по всем качествам, т. е. найти сумму:

$$S = \sum_{j=1}^m (\sum_{i=1}^N r_{ij} - s)^2 \quad (2.52)$$

Далее значение суммы S необходимо нормировать относительно максимально возможного значения, которое находится по формуле:

$$S_{\max} = \frac{N^2 \cdot (m^3 - m)}{12} \quad (2.53)$$

В результате деления выражения (2.52) на (2.53) получается то, что называется коэффициентом конкордации W :

$$W = \frac{12 \cdot S}{N^2 \cdot (m^3 - m)} \quad (2.54)$$

Численное значение коэффициента конкордации изменяется в диапазоне от 0 до 1,0. В случае $W = 0$ имеет место полная несогласованность рассматриваемых рядов рангов; если же $W = 1$, то рассматриваемые ряды рангов полностью совпадают.

Для нашего примера среднее значение суммы рангов S равно $\frac{5 \cdot 11}{2} = 27,5$. Фактические суммы рангов для каждого качества приведены в таблице 2.16. Значение суммы S , определяемой соотношением (2.53), равно 484,5.

Подставляя соответствующие значения в формулу (2.54), получаем $W = \frac{12 \cdot 484,5}{5^2 \cdot (10^3 - 10)} = 0,23$. Поскольку значение коэффициента конкордации значительно меньше единицы, можно сделать вывод о том, что индивидуальные системы ценностей в исследуемой группе испытуемых связаны очень слабо.

Достоверность (значимость) коэффициента множественной ранговой корреляции определяется с помощью критерия χ^2 (хи-квадрат), эмпирическое значение которого вычисляется по формуле:

$$c_{\text{эмп}}^2 = \frac{12 \cdot S}{N \cdot m \cdot (m + 1)} \quad (2.55)$$

Критическое значение критерия χ^2 находится по специальной таблице (см. Приложение 3) в зависимости от числа степеней свободы $k = m - 1$ и уровня достоверности α . Для рассматриваемого примера $c_{\text{эмп}}^2 = 10,57$. По таблице критических значений χ^2 определяем, что для $k=9$ и уровня значимости $\alpha = 0,01$ $\chi_{\text{кр}}^2 = 14,7$. Поскольку $c_{\text{эмп}}^2 < c_{\text{кр}}^2$, делаем вывод, что коэффициент конкордации $W = 0,23$ не может считаться достоверным. Следовательно, мы не можем с достоверностью утверждать, что индивидуальные системы ценностных ориентаций в группе из 5 человек имеют хоть какую-то степень согласованности.

Корреляция номинальных признаков Для определения взаимосвязи показателей, измеренных в номинальных шкалах, строятся таблицы (или матрицы) сопряженности и определяются соответствующие коэффициенты сопряженности анализируемых показателей. Матрица сопряженности для двух показателей x и y , имеющих альтернативные признаки ("да" – "нет" или "1" – "0"), представлена в таблице 2.17.

Числа a, b, c, d в матрице сопряженности соответствуют количеству испытуемых, обладающих сочетанием соответствующих значений показателей x и y . Так, например, число b равно количеству испытуемых, у которых показатель y имеет значение "1", а показатель x – значение "0". Очевидно, что сумма всех чисел $a + b + c + d$ равна

Матрица сопряженности
альтернативных признаков

Таблица 2.17

		Показатель x		Сумма
		x_1	x_0	
Показатель y	y_1	a	b	$a + b$
	y_2	c	d	$c + d$
Сумма		$a + c$	$b + d$	N

общему количеству испытуемых N , т. е. объему выборки.

Коэффициенты
ассоциации и
сопряженности

Условие статистической независимости значений показателей x и y эквивалентно требованию, чтобы доля испытуемых, имеющих значение "1" по показателю y в группе испытуемых, имеющих значение "1" по показателю x , была равна доле испытуемых, имеющих значение "1" по показателю y во всей выборке испытуемых. Другими словами, должно выполняться соотношение $\frac{a}{a+c} = \frac{a+b}{a+b+c+d}$ или $a \cdot d = b \cdot c$. Если же

$a \cdot d - b \cdot c \neq 0$, то рассматриваемые признаки являются взаимосвязанными.

Для оценки степени связи альтернативных признаков обычно используются два коэффициента корреляции:

коэффициент ассоциации:
$$Q = \frac{ad - bc}{ad + bc} \quad (2.56)$$

коэффициент сопряженности:
$$j = \frac{ad - bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}} \quad (2.57)$$

Оба коэффициента принимают значения в диапазоне от +1,0 до -1,0. В случае независимости признаков значения обоих коэффициентов равно нулю.

Коэффициент ассоциации Q равен +1,0, когда имеется полная связь показателей: все испытуемые, имеющие значения "1" по показателю x , одновременно имеют такие же значения и по показателю y ($c = 0$), либо

все испытуемые, имеющие значения "1" по показателю y , одновременно имеют такие же значения по показателю x ($b = 0$). В случае обратной связи ($a = 0$ или $d = 0$), коэффициент ассоциации Q равен $-1,0$.

Коэффициент сопряженности j , в отличие от коэффициента ассоциации Q , предназначен для измерения степени двухсторонней связи (корреляции) между значениями коррелируемых показателей, когда одновременно выполняются оба условия $c=0$ и $b=0$ (в этом случае $j = +1$) или $a=0$ и $d=0$ (в этом случае $j = -1$). Поскольку коэффициент сопряженности является более жестким по сравнению с коэффициентом ассоциации, его значение всегда несколько ниже, т. е. $j < Q$.

Рассмотрим процедуры вычисления коэффициентов ассоциации и сопряженности на примере анализа распределений испытуемых по показателям тревожности и успеваемости, приводимых ранее (см. таблица 2.13).

Чтобы воспользоваться указанными коэффициентами, необходимо преобразовать значения показателей тревожности и успеваемости, приведенные в таблице 2.13, в альтернативные признаки. Присвоим всем значениям тревожности, превышающим среднее значение, число "1", а всем значениям, меньшим среднего – число "0". Точно также поступим со значениями показателя успеваемости. Преобразованные таким образом исходные данные представлены в таблице 2.18.

*Распределения показателей тревожности X_i и успеваемости Y_i
в значениях альтернативных признаков*

Таблица 2.18

Показатель	Номер учащегося														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X_i	1	1	0	1	0	1	0	1	0	1	0	1	0	0	1
Y_i	0	1	1	0	0	1	1	0	0	0	1	0	0	1	1

Соответствующая матрица сопряженности для альтернативных значений показателей тревожности и успеваемости представлена в таблице 2.19.

Коэффициент ассоциации Q , характеризующий взаимосвязь показателей тревожности и успеваемости, вычисленный по формуле (2.56), составляет

$$Q = \frac{3 \cdot 3 - 4 \cdot 5}{3 \cdot 3 + 4 \cdot 5} = -0,38.$$

Аналогичный коэффициент сопряженности j , вычисленный по формуле (2.57), составляет

$$j = \frac{3 \cdot 3 - 4 \cdot 5}{\sqrt{7 \cdot 8 \cdot 7 \cdot 8}} = -0,20.$$

Таблица сопряженности показателей тревожности и успеваемости

Таблица 2.19

		Показатель x		Сумма
		x_1	x_0	
Показатель y	y_1	3	4	7
	y_0	5	3	8
Сумма		8	7	15

В сравнении с коэффициентом корреляции Пирсона $r_{xy}=0.048$, полученным на той же выборке учащихся, значения коэффициентов Q и j отличаются: они отрицательны и имеют более высокое значение. Это объясняется тем, что при преобразовании значений показателя из шкал более высокого порядка (пропорциональные) в шкалы более низкого порядка (номинальные) происходит огрубление данных и частичная потеря информации. В связи с этим в корреляционном анализе не рекомендуется делать какие-либо выводы относительно степени связи показателей на основе сравнения коэффициентов корреляции, относящихся к различным типам.

*Множественный
коэффициент
квадратичной
сопряженности*

Показатели в номинальных шкалах могут характеризоваться не только двумя альтернативными признаками, но и целой группой признаков. В этом случае можно говорить о сложном показателе, имеющем определенную структуру характеризующих его признаков. К таким показателям относятся, например, характер, интеллект, мотивационная сфера. Для определения корреляционных связей между такими сложными психологическими образованиями используется множественный коэффициент квадратичной сопряженности χ^2 .

Как и в случае с альтернативными признаками, для нахождения численного значения коэффициента χ^2 строится матрица сопряженности, размерность которой определяется размерностью коррелируемых показателей.

Допустим, имеются два сложных номинальных показателя x и y , в состав которых входят, соответственно, l и m номинальных признаков. Матрица сопряженности, соответствующая указанному множеству признаков, представлена в таблице 2.20.

В таблице 2.20 k_{ij} обозначает количество испытуемых, имеющих i -ый признак по показателю y и j -ый признак по показателю x .

Таблица сопряженности двух множеств номинальных признаков

Таблица 2.20

		Признаки показателя x						Σ
		1	2	...	j	...	l	
Признаки показателя y	1	k_{11}	k_{12}	...	k_{1j}	...	k_{1l}	a_1
	2	k_{21}	k_{22}	...	k_{2j}	...	k_{2l}	a_2
	...	...	...	...	...	...	...	...
	i	k_{i1}	k_{i2}	...	k_{ij}	...	k_{il}	a_i
	...	...	...	...	...	...	...	...
	m	k_{m1}	k_{m2}	...	k_{mj}	...	k_{ml}	a_m
Σ		b_1	b_2	...	b_j	...	b_l	N

Обозначим через $\bar{k}_{ij}$ - количество испытуемых, которое получилось бы в соответствующих ячейках матрицы сопряженности в случае независимости коррелируемых показателей.

При заданных значениях a_i и b_j доля испытуемых, имеющих одновременно i -ый признак показателя y и j -ый признак показателя x в случае независимости показателей, составляет:

$$f_{ij} = \frac{\bar{k}_{ij}}{N} = \frac{a_i}{N} \cdot \frac{b_j}{N} \quad (2.58)$$

Отсюда находим

$$\bar{k}_{ij} = \frac{a_i \cdot b_j}{N} \quad (2.59)$$

Очевидно, что чем больше фактические значения k_{ij} отличаются от значений $\bar{k}_{ij}$, тем в меньшей степени их можно считать независимыми. Соответственно, тем в большей степени проявляется их взаимосвязь. Для оценки различий в значениях k_{ij} и $\bar{k}_{ij}$ находится коэффициент квадратичной сопряженности

$$C^2 = \sum_{i=1}^m \sum_{j=1}^l \frac{(k_{ij} - \bar{k}_{ij})^2}{\bar{k}_{ij}} \quad (2.60)$$

Для удобства вычислений формула для коэффициента квадратичной сопряженности записывается в виде

$$C^2 = N \left(\sum_{i=1}^m \sum_{j=1}^l \frac{k_{ij}^2}{a_i \cdot b_j} - 1 \right) \quad (2.61)$$

В случае независимости коррелируемых показателей значение χ^2 равно нулю. Чем больше величина χ^2 , тем более вероятно, что между показателями x и y существует взаимосвязь. Наличие корреляции показателей считается достоверной, если эмпирическое значение $C_{\text{эм}}^2$ превышает критическое $C_{\text{кр}}^2$, которое определяется из таблиц граничных значений $C_{\text{кр}}^2$ (см. Приложение 3) в зависимости от числа степеней свободы $k = (l - 1) \cdot (m - 1)$ и уровня значимости α .

В качестве примера вычисления χ^2 рассмотрим, имеется ли связь между типом темперамента и предпочитаемым цветом в группе детей старшего дошкольного возраста по данным, полученным в исследовании С.З. Сыдыковой¹⁵.

В данном примере в качестве сложных коррелируемых показателей будут выступать темперамент, включающий 4 признака, и предпочитаемый цвет, включающий 6 признаков. Матрица сопряженности соответствующих множеств номинальных признаков представлена в таблице 2.21.

¹⁵ Сыдыкова С.З. Взаимосвязь доминирующего цвета и типа темперамента в избирательной деятельности детей: дисс. канд. психол. наук. –Казань, 2002.

*Матрица сопряженности двух множеств
номинальных признаков*

Таблица 2.21

		Тип темперамента				Σ
		М	Ф	С	Х	
Предпочитаемый цвет	красный	-	1	2	6	9
	желтый	-	1	15	4	20
	синий	4	5	1	1	11
	зеленый	-	8	2	2	12
	розовый	2	7	1	3	13
	коричневый	-	6	-	-	6
Σ		6	28	21	16	71

Используя выражение (2.61), находим промежуточные значения C_i^2 для каждой отдельной строчки таблицы сопряженности.

$$C_1^2 = \left[\frac{0}{9 \cdot 6} + \frac{1}{9 \cdot 28} + \frac{4}{9 \cdot 21} + \frac{36}{9 \cdot 16} \right] = 0,28 \quad C_2^2 = \left[\frac{0}{20 \cdot 6} + \frac{1}{20 \cdot 28} + \frac{125}{20 \cdot 21} + \frac{16}{20 \cdot 16} \right] = 0,59$$

$$C_3^2 = \left[\frac{16}{11 \cdot 6} + \frac{25}{11 \cdot 28} + \frac{1}{11 \cdot 21} + \frac{1}{11 \cdot 16} \right] = 0,33 \quad C_4^2 = \left[\frac{0}{12 \cdot 6} + \frac{64}{12 \cdot 28} + \frac{4}{12 \cdot 21} + \frac{4}{12 \cdot 16} \right] = 0,23$$

$$C_5^2 = \left[\frac{4}{13 \cdot 6} + \frac{49}{13 \cdot 28} + \frac{1}{13 \cdot 21} + \frac{9}{13 \cdot 16} \right] = 0,23 \quad C_6^2 = \left[\frac{0}{6 \cdot 6} + \frac{36}{6 \cdot 28} + \frac{0}{6 \cdot 21} + \frac{0}{6 \cdot 40} \right] = 0,21$$

Суммируя промежуточные значения, окончательно находим

$$C^2 = 71 \cdot (0,28 + 0,59 + 0,33 + 0,23 + 0,23 + 0,21 - 1) = 71 \cdot 0,87 = 61,77$$

По таблице граничных значений $C_{кр}^2$ (см. Приложение 3) для числа степеней свободы $k = (4 - 1) \cdot (6 - 1) = 15$ и уровня значимости $\alpha = 0,01$ находим $C_{кр}^2 = 30,6$. Поскольку $C_{эмп}^2 > C_{кр}^2$, можно сделать вывод, что тип темперамента и предпочитаемый цвет у детей старшего дошкольного возраста достоверно связаны между собой.

Однако, установив факт наличия достоверной связи, желательно определить и ее силу или степень связи. Для этого используются следующие коэффициенты корреляции:

$$\text{коэффициент Чупрова } T^2 = \frac{c^2}{N \cdot \sqrt{(l-1) \cdot (m-1)}} \quad (2.62)$$

$$\text{коэффициент Крамера } K^2 = \frac{c^2}{N \cdot \min(l-1, m-1)} \quad (2.63)$$

где $\min(l-1, m-1)$ - минимальное из чисел $l-1, m-1$.

Численные значения обоих коэффициентов лежат в диапазоне от 0 до 1. Чем ближе значения коэффициентов к единице, тем сильнее связь между коррелируемыми показателями. Коэффициент Чупрова может достигать значения, равного 1, только в том случае, если размерность коррелируемых показателей одинакова, т. е. $l = m$. Значения коэффициента Крамера всегда несколько выше соответствующего значения коэффициента Чупрова.

Для нашего примера коэффициент Чупрова равен $\frac{61,77}{71 \cdot \sqrt{3 \cdot 5}} = 0,22$,

коэффициент Крамера равен $\frac{61,77}{71 \cdot 3} = 0,29$.

Полученные данные свидетельствуют о том, что связь между исследуемыми показателями хотя и есть, но не сильная, а точнее, достаточно слабая.

Для корректного использования множественного коэффициента квадратичной сопряженности необходимо, чтобы среднее значение чисел в ячейках матрицы сопряженности было не менее $3 \div 5$, что имеет место при объеме выборки $N \geq (3 \div 5) \cdot l \cdot m$.

Корреляция показателей в смешанных шкалах

В психологических исследованиях довольно часто приходится иметь дело с показателями, значения которых представлены в измерительных шкалах разного типа. Например, при изучении влияния фактора, связанного с наличием или отсут-

ствием высшего образования у родителей, на развитие интеллектуальной сферы и качеств личности ребенка в зависимости от его пола мы можем иметь:

а) номинальные показатели альтернативного типа – образование родителей (наличие высшего образования – "1", отсутствие – "0"), пол ребенка (мужской – "1", женский – "0");

б) показатели по тестам интеллекта, представленные в шкале интервалов;

в) экспертные оценки личностных качеств, выраженные в шкале порядка.

Взаимосвязь показателей x и y , один из которых (x) представлен альтернативными признаками "1" и "0", а другой (y) численными значениями y_i шкалы интервалов, определяется с помощью *точечного бисериального коэффициента корреляции Пирсона*, который вычисляется по формуле:

$$r_{pb} = \frac{\bar{y}_1 - \bar{y}_0}{S_y} \sqrt{\frac{n_1 \cdot n_0}{N \cdot (N - 1)}} \quad (2.64)$$

где: $\bar{y}_1$ и $\bar{y}_0$ - средние значения показателя y (y_i) у испытуемых, имеющих по показателю x значения, соответственно, "1" и "0";

S_y - среднее квадратичное отклонение по всем значениям y_i ;

n_1 и n_0 - количество испытуемых, имеющих по показателю x значения "1" и "0", соответственно;

N – общее количество испытуемых.

В случае, когда один показатель (x) представлен альтернативными признаками, а другой (y) выражен в значениях рангов, взаимосвязь между показателями определяется с помощью *коэффициента бисериальной ранговой корреляции* r_{br} :

$$r_{br} = \frac{2}{N} (\bar{y}_1 - \bar{y}_0) \quad (2.65)$$

где $\bar{y}_1$ и $\bar{y}_0$ - средние значения рангов испытуемых, имеющих по показателю x значения, соответственно, "1" и "0".

Уровень значимости коэффициентов корреляции смешанных признаков определяется с использованием t-критерия Стьюдента, вычисляемого по формулам (2.43) или (2.44) в зависимости от объема выборки.

В случаях, когда одни показатели измеряются в шкалах интервалов, а другие – в шкалах порядка, для нахождения связи между показателями предварительно осуществляется преобразование значений интервальных шкал в значения рангов соответствующих порядковых шкал, и затем используются обычные процедуры ранговой корреляции.

Классификация коэффициентов корреляции, используемых для нахождения связи между показателями, измеренными в разных шкалах, приведена в таблице 2. 22.

Классификация коэффициентов корреляции

Таблица 2.22

		Тип шкалы значений показателя x		
		Интервалов	Порядка	Наименований
Тип шкалы значений показателя y	Интервалов	коэффициент корреляции Пирсона r_{xy}		точечный бисериальный коэффициент корреляции r_{pb}
	Порядка		а) коэффициент корреляции Спирмена ρ ; б) коэффициент конкордации W	коэффициент бисериальной ранговой корреляции r_{br}
	Наименований	точечный бисериальный коэффициент корреляции r_{pb}	коэффициент бисериальной ранговой корреляции r_{br}	а) коэффициент ассоциации Q и сопряжен-сти ϕ ; б) множествен. коэффициент квадратичной сопряжен-сти χ^2

**Построение
корреляционных
плеяд**

Результаты корреляционного анализа, проводимого на множестве коррелируемых показателей, обычно представляются в форме, так называемой, интеркорреляционной матрицы или матрицы взаимных корреляций. Далее данные матрицы интеркорреляций могут быть представлены в графической форме – в форме корреляционных плеяд.

В состав интеркорреляционной матрицы включаются, как правило, все коэффициенты корреляции, представляющие определенный интерес для исследователя и имеющие достаточно высокий уровень значимости. Уровень значимости коэффициента корреляции, включенного в матрицу интеркорреляций, принято отмечать звездочками. Если коэффициент корреляции значим на уровне $\alpha = 0,05$, он отмечается одной звездочкой (например $0,712^*$), если значим на уровне $\alpha = 0,01$ – двумя звездочками (например $0,564^{**}$), если значим на уровне $\alpha = 0,001$ – тремя звездочками (например $(0,823^{***})$).

В качестве примера построения матрицы интеркорреляций и корреляционных плеяд рассмотрим результаты психодиагностики особенностей интеллектуальных процессов у разных учащихся с высоким и низким уровнем успеваемости. Тестирование учащихся проводилось по восьми субтестам теста Векслера. В тестировании определялись 4 показателя, характеризующие особенности вербального интеллекта, и 4 показателя, характеризующие особенности невербального интеллекта.

В качестве показателей вербального интеллекта были выделены:

1. Осведомленность.
2. Понятливость.
3. Способность к логическому обобщению.
4. Объем памяти.

К показателям невербального интеллекта были отнесены:

5. Перцептивные способности (зрительное восприятие).
6. Пространственное воображение.
7. Структурирование целого по отдельным его частям.
8. Концентрация и переключаемость внимания.

На основе указанных 8-ми показателей были образованы еще три комплексных показателя:

9. Интегральный показатель вербального интеллекта, включающий первые четыре показателя 1, 2, 3, 4.

10. Интегральный показатель невербального интеллекта, включающий последующие четыре показателя 5, 6, 7, 8.

11. Показатель общего интеллекта, объединяющий все восемь базовых показателей.

Для каждой пары показателей отдельно по группе "успевающих" и "неуспевающих" учащихся были определены коэффициенты корреляции Пирсона. Далее для каждой группы учащихся были построены матрицы интеркорреляций (см. Таблицы 2.23 и 2.24), в которые были включены коэффициенты корреляции, значения которые превышали 0,65.

*Матрица интеркорреляций показателей интеллекта
группы "успевающих" учащихся*

Таблица 2.23

		Показатели интеллекта										
		1	2	3	4	5	6	7	8	9	10	11
Показатели интеллекта	1		.70							.67		
	2	.70								.90		
	3											
	4											
	5											
	6											
	7											
	8										.70	
	9	.67	.90									.80
	10								.70			.94
	11									.80	.94	

Матрица интеркорреляций показателей интеллекта
группы "неуспевающих" учащихся

Таблица 2.24

		Показатели интеллекта										
		1	2	3	4	5	6	7	8	9	10	11
Показатели интеллекта	1				.67		.77			.69		
	2									.74		
	3							.88				
	4	.67				.66				.86	.88	.94
	5				.66						.73	.71
	6	.77								.67		
	7			.88								
	8											
	9	.69	.74		.86		.67				.70	.80
	10				.88	.73				.70		.98
	11				.94	.71				.80	.98	

Корреляционные плеяды, отражающие взаимосвязи между выделенными показателями, в группах "успевающих" и "неуспевающих" учащихся, представлены на рис. 2.19.

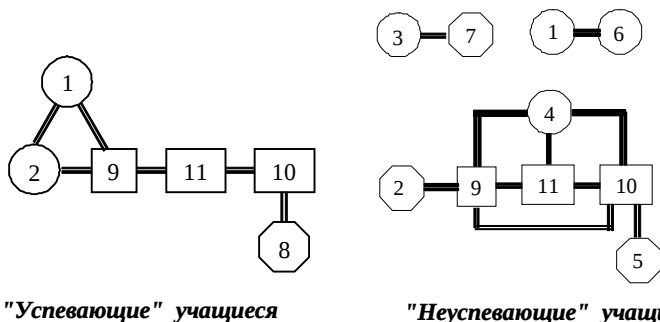


Рис. 2.19. Корреляционные плеяды взаимосвязей показателей интеллектуальных процессов в группе "успевающих" и "неуспевающих" учащихся.

Анализ корреляционных плеяд показывает, что для группы "неуспевающих" учащихся характерным является изолированность показателя осведомленности (наличия знаний) от других показателей вербального интеллекта, в том числе и от понимания. Кроме того, очевидно повышение значимости в структуре интеллектуальных процессов показателей памяти, а также невключенность в эту структуру показателей внимания. На основании этих данных можно предположить, что в усвоении учебного материала "слабые" учащиеся "вылезают" за счет образного мышления и механической памяти. Проблемы в учебе возникают у них в связи с бессистемным характером знаний при отсутствии понимания и низком уровне внимания. В усвоении учебного материала "успевающие" учащиеся в отличие от "неуспевающих" ориентируются на понимание, основанное на осведомленности (наличии знаний) при наличии высокого уровня внимания.

2.6. Понятие о факторном анализе

Факторный анализ представляет собой вариант многомерного статистического анализа, в основе которого лежат процедуры корреляционного анализа. По существу факторный анализ можно рассматривать как логическое продолжение корреляционного анализа для случаев, когда испытуемый или другой объект подвергается психодиагностике по множеству показателей.

При наличии множества диагностируемых показателей, характеризующих объект психодиагностики, возникает необходимость в адекватной интерпретации большого объема эмпирических данных и в их представлении в более компактной форме. Одной из форм компактного представления результатов психодиагностики является рассмотренная нами выше процедура построения корреляционных плеяд, которая позволяет скомпоновать отдельные взаимосвязанные показатели в определенные блоки или плеяды. При факторном анализе также выделяются определенные блоки или группы взаимосвязанных показателей. Однако в отличие от метода корреляционных плеяд при факторном анализе количество блоков, в которые необходимо скомпоновать диагностируемые показатели, устанавливается предварительно до проведения анализа. Фактически это процедура

вынужденного разбиения (или обобщения) показателей на определенное количество групп.

Теоретическим обоснованием такого разбиения является предположение о существовании некоторых "скрытых" факторов (отсюда и название факторного анализа), благодаря действию которых, собственно, и образуются связи между показателями. Количество "скрытых" или общих для ряда взаимосвязанных показателей факторов и, соответственно, количество выделяемых групп показателей устанавливается из расчета $2 \div 3$ показателя на один фактор. Например, при общем количестве диагностируемых показателей, равным 10, количество факторов в факторном анализе устанавливается от 3 до 5. Максимальное количество факторов равно общему количеству диагностируемых показателей.

Основная стратегия при выделении факторов заключается в оценке гипотезы о минимальном количестве общих факторов, действие которых обуславливает имеющиеся корреляции. При этом сами общие факторы должны быть ортогональны, т. е. независимы друг от друга. Если нет каких-либо предположений о возможном количестве факторов, то начинают с однофакторной модели. Далее переходят к двухфакторной и т. д. до тех пор, пока не будут достигнуты удовлетворительные результаты. В конечном счете, в результате факторного анализа происходит сокращение размерности общих характеристик диагностируемого объекта от большого множества исходных показателей до небольшого множества обобщенных факторов.

Выделение в ходе анализа данных общего для ряда переменных фактора позволяет решать одну непростую задачу — оценивать некоторую скрытую от непосредственного наблюдения переменную (фактор) опосредованно или косвенно через ее проявление в других, прямо измеряемых переменных. Подобным образом в психодиагностике личности были обнаружены, выделены и измерены многие личностные свойства. Классическим примером является работа, сделанная Кеттеллом, который с помощью факторного анализа выделил из очень большого множества личностных качеств 16 базовых и создал широко известный в настоящее время 16-ти факторный опросник личности. Другим примером является создание не менее известного личностного теста интерперсональных отношений

Т. Лири, в котором на основе диагностики 128 поведенческих признаков (отдельных личностных свойств) выделяются 6 обобщенных свойств личности.

С математической точки зрения сущность факторного анализа заключается в предположении, что выделяемый фактор представляет собой линейную комбинацию диагностируемых показателей и выражается уравнением:

$$F_i = \sum_{j=1}^k (w_{ij} \cdot x_j) \quad (2.66)$$

где: w_{ij} – коэффициент, указывающий, в какой степени i -ый фактор оказывает влияние на j -ый диагностируемый показатель x_j ;

k – количество диагностируемых показателей.

Коэффициент w_{ij} в факторном анализе носит название *факторной нагрузки*.

Графически смысл процедуры факторного анализа может быть представлен следующим образом (см. рис. 2.20):

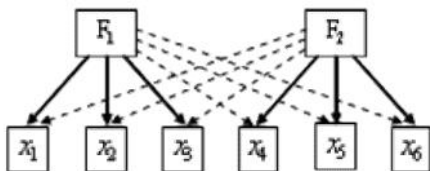


Рис. 2.20. Пример выделения двух общих факторов (F_1, F_2) для объяснения корреляционных связей между шестью показателями ($x_1 \div x_6$).

На рис. 2.20 факторные нагрузки обозначены стрелками. Сплошными стрелками обозначены более высокие факторные нагрузки, пунктирными – более низкие. Как видно из рисунка, первый фактор F_1 оказывает наибольшее влияние на показатели x_1, x_2, x_3 и лишь незначительное на показатели x_4, x_5, x_6 . Фактор F_2 , наоборот, на первые три показателя оказывает слабое влияние, а сильное – на последние три.

В настоящее время процедуры факторного анализа осуществляются с помощью специальных компьютерных программ, например, Statistica или SPSS. Благодаря этим программам осуществляется факторизация корреляционной матрицы (матрицы интеркорреляций) и вычисление опти-

мальных значений факторных нагрузок из множества других возможных вариантов.

Одним из наиболее распространенных методов факторного анализа является метод *варимакса*, который обеспечивает минимизацию количества показателей, связанных с каждым из выделяемых факторов. Это достигается за счет группировки вокруг каждого фактора только тех показателей, которые связаны с ним в большей степени, чем остальные.

Существует и другой метод факторного анализа – метод *квартимакса*, который в определенном смысле противоположен варимаксу, поскольку минимизирует количество факторов, необходимых для объяснения того или иного показателя.

При проведении факторного анализа необходимо помнить, что выделяемые факторы – это факторы гипотетические, поэтому обращение к ним и их интерпретация должны осуществляться при наличии солидных теоретических оснований.

2.7. Кластерный анализ

В психодиагностике достаточно часто возникают задачи, когда необходимо выявить взаимосвязи не только между измеряемыми показателями, но и между обследуемыми объектами (испытуемыми), имеющими сходные значения показателей. По существу, это задачи, связанные с необходимостью в классификации объектов психодиагностики, в разбиении множества диагностируемых объектов на определенные классы или типы. Решение задач данного типа обеспечивается применением методов автоматической классификации, частным случаем которых является кластерный анализ. Суть кластерного анализа состоит в следующем¹⁶.

Допустим, мы имеем выборку из N испытуемых, на которой проведено измерение m показателей с использованием интервальных шкал. По результатам измерений каждый испытуемый может быть охарактеризован соответствующей совокупностью значений показателей $\{x_{ij}\}$, где x_{ij} – значение j -го показателя у i -го испытуемого. Представим далее, что совокупность показателей составляет некоторое многомерное пространство, в котором m координат, каждая из которых соответствует одному из показате-

¹⁶ Бородкин Л.И., Промахина И.М. Методы многомерного статистического анализа // Количественные методы в исторических исследованиях. – М.: Высш. шк., 1984. – С. 267-275.

лей. В этом m -мерном пространстве каждый из испытуемых будет представлен некоторой точкой с координатами $(x_{i1}, x_{i2}, \dots, x_{im})$. Характер распределения точек, обозначающих испытуемых, будет отражать структуру сходства и различия испытуемых в заданной системе показателей.

О сходстве испытуемых можно судить по расстоянию между соответствующими точками m -мерного пространства показателей. Смысл такого понимания сходства заключается в том, что чем более схожи испытуемые, тем меньше у них различия в значениях одноименных показателей, и тем ближе между собой точки в m -мерном пространстве.

Определение расстояния d_{ik} между i -ой и k -ой точками в m -мерном пространстве осуществляется по формуле:

$$d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - x_{kj})^2} \quad (2.67)$$

В результате вычисления множества значений d_{ik} получается некоторая матрица расстояний между испытуемыми, которая и служит основанием для проведения кластерного анализа и построения соответствующей классификации испытуемых.

Процедура классификации состоит из последовательности шагов, на каждом из которых производится объединение наиболее сходных испытуемых или сходных групп испытуемых. Группы сходных испытуемых в кластерном анализе называются *кластерами*. На первом шаге выделяются несколько пар испытуемых, имеющих между собой минимальные значения d_{ik} ¹⁷. Эти пары объединяются в кластеры, которые, фактически, образуют основания (или критерии) последующей классификации и выделения соответствующих классов испытуемых. При образовании кластеров, естественно, размерность матрицы расстояний уменьшается на количество испытуемых, включаемых в образуемые кластеры. Поскольку образуемые кластеры занимают определенное место в пространстве, между ними и другими испытуемыми, еще не объединенными в кластеры, вновь вычисляются значения соответствующих расстояний.

При вычислении расстояний между двумя кластерами находится среднее значение квадрата расстояний между всеми парами испытуемых, в

¹⁷ Обычно выделяются две ÷ четыре пары.

которых один испытуемый принадлежит к одному кластеру, а другой – к другому (см. уравнение 2.68).

$$D_{pq}^2 = \sum_{i \in X_p} \sum_{k \in X_q} \frac{d_{ik}^2}{n_p \cdot n_q} \quad (2.68)$$

где: D_{pq}^2 - мера близости между кластерами с номерами p и q ;

X_p и X_q - кластеры с номерами p и q , соответственно;

n_p и n_q - количество испытуемых в соответствующих кластерах;

$i \in X_p$ и $k \in X_q$ означает, что i -ый и k -ый испытуемые принадлежат, соответственно, к кластерам X_p и X_q .

Затем относительно образованных кластеров находятся наименее удаленные отдельные испытуемые или другие кластеры. Эти испытуемые (или кластеры) включаются в исходные, в результате чего образуются новые кластеры. Далее вся процедура вычисления новых расстояний и образования новых кластеров повторяется. На каждом шаге кластеризации отдельные, невязаносвязанные кластеры образуют соответствующие классы испытуемых. С каждым шагом количество таких кластеров и, соответственно, количество выделяемых классов испытуемых уменьшается. В конце концов, все испытуемые будут объединены в один большой кластер, который представляет собой всю выборку испытуемых.

В связи с необходимостью многократного пересчета расстояний между кластерами процедура кластерного анализа обычно проводится с использованием специальных компьютерных программ.

Несколько проще в вычислительном отношении, но все же весьма утомительно, осуществляется кластеризация испытуемых, если показатели, составляющие координаты многомерного пространства, измеряются в дихотомических шкалах, т. е. могут иметь всего лишь два альтернативных значения: "1" или "0". В этом случае в качестве меры близости обычно используются два показателя:

$$\square \text{ расстояние по Хеммингу } d_{ik} = \sum_{j=1}^m |x_{ij} - x_{kj}| \quad (2.69)$$

$$\square \text{ коэффициент композиционного сходства } S_{ik} = \frac{P_{ik}}{m + q_{ik}} \quad (2.70)$$

где P_{ik} и Q_{ik} - числа, равные количеству показателей, имеющих одинаковые и, соответственно, отличающиеся значения для i -го и k -го испытуемых.

Значения расстояний по Хеммингу лежат в диапазоне от 0 (в случае полного совпадения признаков) до m (в случае максимального расхождения признаков). Коэффициент композиционного сходства изменяется от 0 до 1 и имеет противоположную направленность, т. е. равен 0 при максимальном отличии признаков и 1 – при полном совпадении.

Рассмотрим в качестве примера применение процедуры кластерного анализа для классификации учителей по типам и видам педагогических установок, определяемых с помощью теста СПУЛ (см. Приложение 4). Исходные данные тестирования группы их 14 учителей приведены в таблице 2.25.

Особенности структуры педагогических установок личности

Таблица 2.25

Установка		Номер учителя													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Тип установки	1. Авторит.	54	36	42	49	54	36	57	78	58	34	57	39	46	62
	2. Демократ.	57	67	64	70	52	62	68	54	53	67	56	67	65	50
	3. Либерал.	60	68	65	52	65	73	46	39	60	70	58	65	60	59
Вид установки	4. На ученика	62	57	66	55	42	62	59	55	57	55	42	42	55	38
	5. На УВП	57	65	62	78	72	56	73	59	58	57	51	61	65	55
	6. На учителя	52	49	43	38	57	53	39	57	56	59	82	68	51	78

Мера близости учителей по показателям типа и вида установок определялась в соответствии с формулой (2.67). В результате вычислений получилась матрица расстояний, которая представлена в таблице 2.26.

Из таблицы видно, что наибольшее сходство по значениям диагностируемых показателей имеют учителя под номером 1 и 9 – расстояние между ними в многомерном пространстве диагностируемых показателей составляет 9,1 балла.

*Матрица расстояний между учителями
по особенностям педагогических установок*

Таблица 2.26

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1		24,2	18,5	30,8	26,6	22,9	27,7	33,3	9,1	26,6	36,9	32,2	16,0	37,2
2	24,2		13,5	27,1	30,0	13,3	33,2	53,7	29,2	13,6	46,7	25,1	13,7	48,6
3	18,5	13,5		25,7	34,1	16,2	28,2	49,1	26,0	22,6	50,7	35,1	15,8	52,0
4	30,8	27,1	25,7		33,0	37,9	12,4	44,7	34,2	38,1	56,0	40,6	21,2	55,2
5	26,6	30,0	34,1	33,0		34,1	35,3	40,0	21,7	32,5	33,9	26,5	22,9	29,4
6	22,9	13,3	16,2	37,9	34,1		41,3	55,4	27,9	11,6	44,5	27,6	20,6	47,2
7	27,7	33,2	28,2	12,4	35,3	41,3		35,1	30,8	42,3	54,0	44,4	23,7	53,1
8	33,3	53,7	49,1	44,7	40,0	55,4	35,1		29,3	55,5	40,9	51,7	40,9	37,8
9	9,1	29,2	26,0	34,2	21,7	27,9	30,8	29,3		29,9	31,2	31,2	19,5	29,9
10	26,6	13,6	22,6	38,1	32,5	11,6	42,3	55,5	29,9		39,2	18,1	19,7	43,2
11	36,9	46,7	50,7	56,0	33,9	44,5	54,0	40,9	31,2	39,2		28,3	39,3	11,1
12	32,2	25,1	35,1	40,6	26,5	27,6	44,4	51,7	31,2	18,1	28,3		23,8	31,9
13	16,0	13,7	15,8	21,2	22,9	20,6	23,7	40,9	19,5	19,7	39,3	23,8		40,2
14	37,2	48,6	52,0	55,2	29,4	47,2	53,1	37,8	29,9	43,2	11,1	31,9	40,2	

На основании таблицы расстояний наиболее близкие по значениям установок учителя объединялись в общие кластеры. На первом шаге кластерного анализа было образовано 4 общих кластера с учителями: (1, 9); (4, 7); (6, 10); (11, 14).

Далее при расчетах расстояний между образованными кластерами и другими учителями использовалась формула (2.68). Наиболее близкими на втором шаге кластеризации оказались кластер (6, 10) и учитель 2, а также учителя 3 и 13. В результате было образовано два новых кластера: (6, 10, 2) и (3, 13). На следующем третьем шаге кластеризации наиболее близкими оказались именно эти два кластера. После объединения они образовали новый общий кластер (6, 10, 2, 3, 13). На четвертом шаге к этому кластеру присоединился кластер (1, 9). Далее образовался новый кластер (5, 12), который затем на седьмом шаге объединился с кластером (11, 14).

В итоге на 7-ом шаге кластеризации выделились четыре группы учителей:

I группа – из семи учителей: 6, 10, 2, 3, 13, 1, 9;

II группа – из четырех учителей: 11, 14, 5, 12;

III группа – из двух учителей: 4, 7;

IV группа – из одного учителя: 8.

Графически результаты кластерного анализа представляются в виде дендрограммы (дерева иерархической классификации). Дендрограмма для результатов кластерного анализа, проведенного нами в качестве примера, представлена на рис. 2.21.

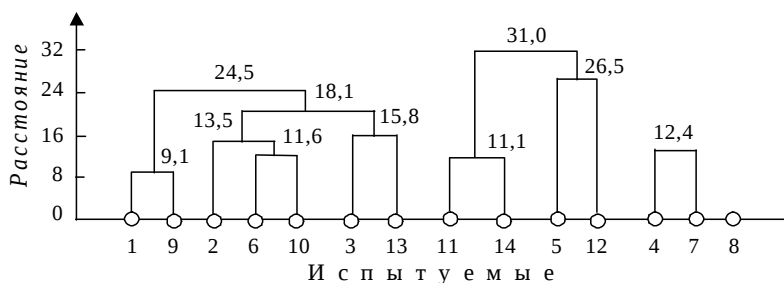


Рис. 2.21. Дендрограмма распределения учителей по группам с совпадающими педагогическими установками.

Присвоение определенных наименований или характеристик выделенным в результате кластерного анализа группам диагностируемых объектов осуществляется на основании содержательного анализа особенностей объектов, вошедших в соответствующую группу.

Например, в нашем примере можно дать следующие характеристики группам учителей:

I группа – "авторитарная с ориентацией на личность учителя";

II группа – "демократическая с ориентацией на личность ученика";

III группа – "с ориентацией на личность учителя при отсутствии выраженной авторитарной установки";

IV группа – "либеральная с ориентацией на ученика".

ГЛАВА 3

ПОКАЗАТЕЛИ ЭФФЕКТИВНОСТИ ПСИХОДИАГНОСТИЧЕСКИХ ПРОЦЕДУР

Выводы, которые делаются на основании данных психодиагностики, во многом зависят от качества используемых психодиагностических процедур. Показателями качества психодиагностических процедур или показателями, обеспечивающими эффективность их использования, являются:

- 1) объективность,
- 2) валидность,
- 3) надежность,
- 4) дискриминативность.

3.1. Объективность

Согласно определению Дж. Кэттелла, психодиагностическая процедура может считаться объективной, если ее цель (цель психодиагностики) скрыта от испытуемого (что не позволяет ему фальсифицировать результаты диагностики) и данные, полученные с ее помощью, могут быть оценены независимо от лица, проводящего диагностику. Другими словами, под объективностью психодиагностической процедуры понимается независимость получаемых с ее помощью данных от двух факторов:

- а) от воли и установок испытуемого,
- б) от воли и установок психодиагноста.

Чтобы оценить, насколько сложна проблема объективности психодиагностической процедуры, имеет смысл рассмотреть динамику тех внутренних психических (а стало быть, и субъективных) процессов, которые протекают в ходе диагностики в психике, как испытуемого, так и психодиагноста.

Очевидно, что при проведении психодиагностики предполагается наличие определенных заданий, предъявляемых испытуемому, наличие определенных воздействий, оказываемых на испытуемого, наличие определенных ситуаций, в которые ставится испытуемый. Вся эта совокупность заданий, воздействий, ситуаций, так или иначе, призвана вызвать проявление у испытуемого того свойства или качества, которое интересует

психодиагноста. Очевидно также, что свойства и качества, интересующие психодиагноста, могут проявляться у испытуемого лишь в определенных его поведенческих актах: движениях, действиях, высказываниях.

Обозначим совокупность заданий, воздействий, ситуаций символом S , а совокупность соответствующих поведенческих актов (поведенческих признаков) символом R . Получается классическая бихевиористская формула $S \Rightarrow R$. Однако известно, что особенности поведения R определяются не только совокупностью внешних воздействий и обстоятельств S , но и внутренними субъективными факторами, связанными, в частности, с субъективностью процессов восприятия, анализа и понимания воспринимаемого, а также оценки значимости происходящего. Испытуемый, таким образом, может вносить существенные искажения в соотношение $S \Rightarrow R$, причем как неосознаваемые (за счет неосознаваемых установок), так и осознаваемые, т.е. прямые фальсификации. Если бы действительно можно было подобрать такую совокупность S (такую диагностическую процедуру), которая независимо от возможных субъективных влияний испытуемого приводила бы к ожидаемым результатам R (т. е. к таким результатам, которые характеризуют действительную меру диагностируемого качества или свойства), то тогда эта совокупность и эта диагностическая процедура могла бы считаться объективной.

Но процедура психодиагностики предполагает не только наличие испытуемого и соответствующих совокупностей S и R , но и психодиагноста, который непосредственно проводит диагностику. В его психике и в его сознании также протекают внутренние субъективные процессы, которые могут оказывать не меньшее влияние, как на результаты измерения, так и на выводы, которые из них следуют, поскольку эти процессы самым непосредственным образом увязываются с указанными совокупностями S и R . Процессы восприятия психодиагноста направлены, с одной стороны, на восприятие совокупности S и, с другой стороны, на восприятие совокупности R . В силу субъективности процессов восприятия восприятие совокупностей S и R у испытуемого и у психодиагноста, в принципе, различно. Более того, психодиагност, зная о характере и целях психодиагностической процедуры, видит за внешними поведенческими актами испытуемого поведенческие признаки (проявления) диагностируемых качеств, чего не знает и не видит испытуемый. В результате в сознании психодиагноста создается модель личности испытуемого и на ее основе вырабаты-

ваются представления о том, какой должна быть совокупность заданий, воздействий, ситуаций (или как она должна быть откорректирована), чтобы эффективность диагностики была более высокой. Таким образом, формирование системы воздействий, характер восприятия и интерпретации результатов диагностики оказываются в зависимости от особенностей личности психодиагноста, от его осознаваемых и неосознаваемых установок.

Соотношение субъективных психических процессов, происходящих в психике и сознании испытуемого и психодиагноста в ходе реализации психодиагностической процедуры, представлены на рис. 3.1.

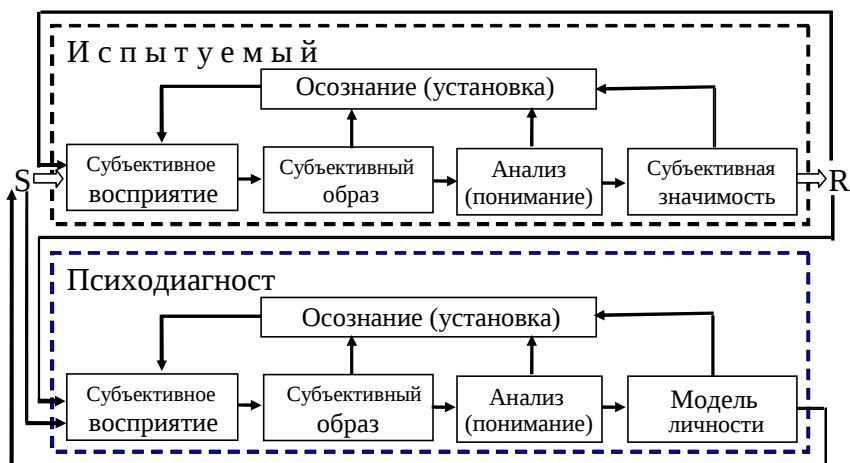


Рис. 3.1. Соотношение субъективных психических процессов испытуемого и психодиагноста, оказывающих влияние на результаты психодиагностики.

Как следует из выше изложенного и из анализа схемы, представленной на рис. 3.1, добиться независимости измерения психологических свойств и качеств от влияния субъективных факторов, связанных с особенностями личности испытуемого и исследователя, практически невозможно. Следовательно, в психодиагностике не может быть абсолютно объективных психодиагностических процедур, если под объективностью понимать именно независимость их от воли и установок испытуемого и исследователя.

Более мягкой формулировкой понятия объективности психодиагностической процедуры является следующая: психодиагностическая процедура считается объективной, если ее использование разными психодиагностами при диагностике одного и того же показателя (признака) на эквивалентных выборках дает одинаковый результат. В данном определении смещены акценты с требования независимости на требование сохранения одинакового результата. Как было показано выше, добиться независимости результатов измерения от особенностей личности испытуемого и исследователя практически невозможно. Но вот добиться сохранения одинакового результата, в принципе, можно. Для этого предпринимаются меры по стандартизации основных составляющих диагностической процедуры, начиная от выбора заданий, предъявляемых испытуемым, инструкций по их выполнению, времени, отводимому на выполнение заданий, и кончая порядком регистрации, обработки и интерпретации результатов диагностики.

Объективность психодиагностической процедуры является предпосылкой ее надежности и валидности.

3.2. Валидность

Понятие валидности	"Валидность" представляет собой научный термин, который используется для характеристики способности психодиагностической процедуры диагностировать действительно то, для диагностики чего она предназначена. Если, например, мы берем тест, чтобы провести диагностику общительности, и тест действительно диагностирует общительность, значит он валидный для диагностики общительности. Если для диагностики общительности мы возьмем тест, валидность которого неизвестна, возможно, что мы, сами того не зная, будем измерять вовсе не общительность, а какое-то другое качество личности.
----------------------------------	---

Степень валидности психодиагностической процедуры показывает, в какой мере данная процедура диагностирует то, что необходимо диагностировать. Тест может иметь высокую степень валидности для измерения одного показателя и низкую степень валидности для измерения другого.

Смысл понятия "валидности" в графической форме можно представить следующим образом.

Допустим, имеется объект психодиагностики, который характеризуется некоторыми двумя качествами K_1 и K_2 (см. рис. 3.2).

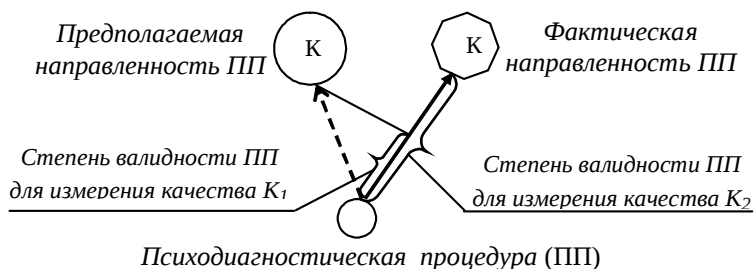


Рис. 3.2. Понятие валидности психодиагностической процедуры.

Допустим далее, что мы используем некоторую психодиагностическую процедуру для измерения качества K_1 и считаем, что результаты измерения дают нам характеристику именно этого качества. Однако, если в действительности данная процедура валидна по отношению не к качеству K_1 , а к качеству K_2 , то возникает ошибка диагностики. Отсутствие такого рода ошибки и характеризуется понятием "валидности". Чем больше валидность диагностической процедуры по отношению к тому, что мы измеряем, тем меньше соответствующая ошибка диагностики. Величина ошибки определяется тем, насколько качество K_2 отличается от качества K_1 . По существу, применение невалидной диагностической процедуры приводит к ошибке диагностики, величину которой можно представить как разность между длиной вектора $\overrightarrow{ПП-K_2}$ и проекцией вектора $\overrightarrow{ПП-K_1}$ на вектор $\overrightarrow{ПП-K_2}$ (см. рис. 3.2). Если качества K_1 и K_2 ортогональны (т. е. невязимосвязаны, независимы друг от друга), соответствующие вектора $\overrightarrow{ПП-K_1}$ и $\overrightarrow{ПП-K_2}$ образуют прямой угол, и проекция одного вектора на другой равна нулю. В этом случае ошибка в диагностике качества K_1 невалидной диагностической процедурой будет максимальна. В других случаях ошибка будет меньше, поскольку в силу взаимосвязи рассматриваемых качеств, результаты фактической диагностики K_2 частично будут содержать данные о качестве K_1 .

Отсюда можно сделать вывод, что чем больше ошибка измерения, тем меньше валидность измерительной процедуры.

**Виды валидности
и методы
их определения**

В зависимости от того, каким образом определяется степень валидности психодиагностической процедуры, выделяют следующие четыре вида валид-

ности:

- ☐ очевидная (внешняя),
- ☐ содержательная,
- ☐ концептуальная,
- ☐ эмпирическая.

Очевидная (или внешняя) валидность – валидность с точки зрения "здорового смысла", когда о психодиагностической процедуре складывается впечатление (в частности, у испытуемых), что она диагностирует именно то, для диагностики чего она используется. Например, если испытуемому предъявляется ряд слов для запоминания, то очевидно, что диагностируются особенности вербальной памяти, а именно объема памяти. Собственно для этого данная процедура и используется. Следовательно, процедура предъявления ряда слов для запоминания обладает очевидной валидностью в отношении диагностики объема памяти. Степень очевидной валидности психодиагностической процедуры определяется долей испытуемых, которые правильно называют назначение диагностической процедуры.

Содержательная валидность – валидность, определяемая тем, насколько содержание заданий адекватно диагностируемому показателю. Например, для диагностики уровня вербального интеллекта тест должен содержать задания на установление соотношений между словами и понятиями, а не на запоминание списка слов или понятий.

Разработка психодиагностических методик, обладающих содержательной валидностью, включает в себя следующие процедурные моменты¹⁸:

- определение предмета психодиагностики, т. е. того качества (или свойства), которое подлежит диагностике;
- определение поведенческих признаков или показателей, характеризующих проявления исследуемого качества;

¹⁸ Клайн П. Справочное руководство по конструированию тестов. –Киев, 1994.

– передача полученного перечня показателей специалистам-экспертам для проверки на предмет их соответствия исследуемому качеству и отсутствия каких-либо упущений;

– преобразование перечня поведенческих признаков диагностируемого качества в систему заданий и измерительных шкал, необходимых для предъявления испытуемым;

– представление заданий и соответствующих измерительных шкал специалистам-экспертам для проверки и оценки адекватности произведенного преобразования.

Степень содержательной валидности психодиагностической процедуры определяется отношением реальной оценки, которую дал эксперт всем составляющим диагностической процедуры, к максимально возможной.

Концептуальная валидность – валидность, определяемая тем, насколько результаты диагностики соответствуют теоретической концепции, положенной в основу диагностической процедуры.

Для практического определения уровня концептуальной валидности психодиагностической процедуры может быть использован предложенный Д.Т. Кэмпбеллом и Д.В. Фиске метод оценки *конвергентной и дискриминантной валидности*. Согласно этому методу, исследователь заранее предсказывает, с результатами каких известных измерительных процедур будут коррелировать результаты валидизируемой процедуры, а с результатами каких – не будут или будут иметь низкий уровень корреляции. По отношению к первым валидизируемая измерительная процедура будет иметь конвергентную валидность, а по отношению ко вторым – дискриминантную.

Алгоритм определения концептуальной валидности состоит в следующем:

- изложить теоретическую концепцию диагностической процедуры;
- выдвинуть ряд гипотез относительно возможных результатов диагностики и их соотношения с другими эмпирическими данными;
- провести эмпирическую проверку выдвинутых гипотез.

Чем больше гипотез получает эмпирическое подтверждение, и чем они убедительнее, тем с большей уверенностью можно говорить о концептуальной валидности диагностической процедуры по отношению к той теоретической концепции, на основе которой она была разработана. Чис-

ленное значение концептуальной валидности диагностической процедуры определяется долей гипотез, получивших подтверждение.

Эмпирическая валидность – валидность, определяемая по отношению к некоторым эмпирическим критериям, в качестве которых, в частности, могут выступать:

- результаты экспертных оценок;
- результаты эмпирического исследования, проведенного с использованием других измерительных процедур, валидность которых уже доказана и не вызывает сомнения.

При использовании в качестве критерия валидации экспертных оценок следует убедиться в достаточной компетентности экспертов, в том, насколько адекватно они понимают оцениваемые показатели, способны ли они дифференцировать проявления оцениваемых показателей на поведенческом уровне, есть ли у них навыки работы с оценочными шкалами и т. д. И в том, и в другом случаях практическое определение эмпирической валидности сводится к определению коэффициентов корреляции совокупности значений измеряемого показателя, полученной при использовании валидируемой диагностической процедуры, и совокупности значений этого же показателя, полученной на той же выборке испытуемых, но другими измерительными процедурами.

Кроме вышеперечисленных, можно отметить еще два вида валидности:

- прогностическую,
- инкрементную.

Прогностическая валидность определяется соотношением результатов измерения с некоторым критерием, характеризующим измеряемое качество, но в более позднее время. Например, если тест интеллекта позволяет оценить, сделать прогноз успеваемости учащегося на более поздний период, значит, этот тест, обладая актуальной валидностью по отношению к интеллекту, обладает еще и прогностической валидностью по отношению к успеваемости. Основная трудность определения прогностической валидности состоит в выборе соответствующего критерия валидации.

Инкрементная валидность характеризует психодиагностическую процедуру с точки зрения значимости ее вклада в определение некоторого общего показателя. Если, например, при определении профессиональной пригодности один из тестов, который входит в батарею используемых тес-

тов, имеет низкую корреляцию с критерием пригодности, но при этом измеряет такое профессионально важное качество, которое другими тестами не измеряется, значит, данный тест обладает инкрементной валидностью.

Поскольку, так или иначе, валидность психодиагностической процедуры характеризуется мерой связи результатов психодиагностики с некоторым критерием, который рассматривается как критерий валидации, при определении валидности широко используются различные процедуры корреляционного анализа. Показателем степени валидности измерительной процедуры обычно выступает соответствующее значение коэффициента корреляции. Однако в общем случае, валидность, в отличие, например, от надежности или дискриминативности, не может быть представлена или охарактеризована каким-либо одним коэффициентом. Поэтому при характеристике валидности психодиагностической процедуры всегда необходимо указывать, о какой именно валидности идет речь и как она определялась.

3.3. Надежность

Понятие надежности Надежность психодиагностической процедуры непосредственно связана с точностью измерения. Если неоднократные измерения каждый раз дают абсолютно разные значения диагностируемого показателя, значит, диагностическая процедура не является надежной.

Как известно, результат любого измерения складывается из двух компонент:

$$X_i = X_{уст} + e_i \quad (3.1)$$

где: $X_{уст}$ – истинное значение показателя,

e_i – ошибка измерения.

Среднее значение измеряемого показателя для k измерений будет равно:

$$X_{cp} = X_{уст} + \frac{\sum_{i=1}^k e_i}{k} \quad (3.2)$$

Увеличивая количество измерений, можно уменьшить погрешность измерения и, соответственно, увеличить его точность и надежность.

При проведении диагностики на группе испытуемых истинные значения $x_{ист}$ каждого испытуемого будут разные, и их дисперсия составит $S_{x_{ист}}^2$.

Общая дисперсия результатов диагностики показателя x на группе испытуемых равна:

$$S_x^2 = S_{x_{ист}}^2 + S_e^2 \quad (3.3)$$

Соотношение (3.3) показывает, что чем меньше дисперсия ошибки S_e^2 , тем в большей степени вариация измеренных значений x соответствует вариации истинных значений $x_{ист}$. Данное соотношение используется для определения надежности диагностической процедуры.

Показателем надежности служит коэффициент $R^2(x, x_{ист})$, определяемый как отношение дисперсии истинных значений $x_{ист}$ измеряемого показателя к дисперсии фактически полученных значений x

$$R^2(x, x_{ист}) = \frac{S_{x_{ист}}^2}{S_x^2} \quad (3.4)$$

Очевидно, что численное значение коэффициента надежности всегда меньше единицы. Величина $R(x, x_{ист})$ называется индексом надежности и представляет собой коэффициент корреляции результатов измерения с истинными значениями диагностируемого показателя.

Используя соотношение (3.3), коэффициент надежности можно представить несколько по-другому:

$$R^2(x, x_{ист}) = \frac{S_x^2 - S_e^2}{S_x^2} = 1 - \frac{S_e^2}{S_x^2} \quad (3.5)$$

Полученное выражение (3.5) носит название *формулы Рюлона*. Ценность этой формулы в том, что она позволяет при известной надежности теста определить стандартную ошибку измерения:

$$S_e = S_x \sqrt{1 - R_{x, x_{uct}}^2} \quad (3.6)$$

Очевидно, что при увеличении надежности теста уменьшается разброс в ошибках измерения.

Знание величины ошибки S_e позволяет построить доверительный интервал для истинного значения показателя:

$$x_{uct} = x_{cp} \pm t \cdot S_x \cdot \sqrt{1 - R_{x, x_{uct}}^2} \quad (3.7)$$

где t – критерий Стьюдента.

Методы определения надежности Для практического определения показателей надежности психодиагностических процедур используются три основных метода:

- метод *ретеста* (метод повторной диагностики),
- метод *параллельного тестирования*,
- метод *расщепления*.

Метод ретеста Надежность психодиагностической процедуры, измеренная методом ретеста, называется ретестовой надежностью. Схема определения ретестовой надежности имеет следующий вид:

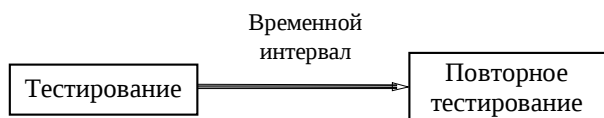


Рис. 3.3. Схема определения ретестовой надежности психодиагностической процедуры.

Повторное тестирование (или ретестирование) осуществляется через определенный временной интервал на той же выборке испытуемых с использованием той же измерительной процедуры и при одинаковых условиях.

Очень многое в методе ретеста зависит от величины временного интервала. Если временной интервал небольшой, испытуемые, уже знакомые с процедурой измерения, помнят результаты предыдущих измерений и

ориентируются на них при повторных измерениях. При этом возможно проявление двух негативных эффектов:

- подгонка под первоначальные результаты,
- стремление к демонстрации "новых" результатов.

И в том, и в другом случаях происходит искажение результатов повторной диагностики и ошибка в определении ретестовой надежности психодиагностической процедуры.

Если временной интервал между измерениями будет слишком большим, то, возможно, испытуемые и забудут данные первичного тестирования, но в самих измеряемых качествах могут произойти настолько значительные изменения, что при повторном тестировании диагностика будет осуществляться, по существу, уже другого, нового качества.

В связи этими обстоятельствами метод ретеста, как метод определения надежности психодиагностической процедуры, практически не используется.

Метод параллельного тестирования Определение надежности психодиагностических процедур методом параллельного тестирования осуществляется с использованием, так называемых, *параллельных форм* или эквивалентных вариантов соответствующих диагностических процедур. Две диагностические процедуры называются параллельными (эквивалентными), если они диагностируют одно и то же свойство и на одной и той же выборке позволяют получить:

- одинаковые средние значения $X_{cp1} = X_{cp2}$,
- одинаковые значения дисперсий $S_{x_1}^2 = S_{x_2}^2$,
- одинаковые ошибки измерения $S_{e_1}^2 = S_{e_2}^2$.

Схема определения надежности диагностической процедуры методом параллельного тестирования имеет следующий вид:

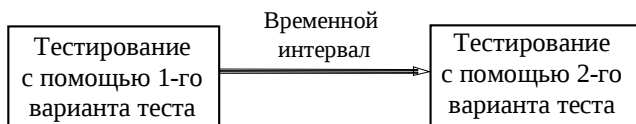


Рис. 3.4. Схема определения надежности диагностической процедуры методом параллельного тестирования.

Как и в методе ретеста, здесь осуществляется повторное тестирование одной и той же выборки через определенный временной интервал, но уже с помощью другого "параллельного" или эквивалентного варианта диагностической процедуры. В связи с этим, надежность психодиагностической процедуры, определенная методом параллельного тестирования, называется *эквивалентной надежностью*.

Высокие значения коэффициента корреляции $R_{x_1x_2}$ результатов измерений по параллельным (эквивалентным) формам диагностической процедуры указывают не только на надежность самой диагностической процедуры, но и на эквивалентность психологического содержания используемых форм этой процедуры.

При разработке параллельных форм психодиагностических процедур важно, чтобы они были одинаковы или эквивалентны:

- по количеству заданий,
- по содержанию заданий и форме их представления,
- по диапазону и уровню сложности (трудности) заданий,
- по характеру инструкции, поясняющим примерам, форме бланков и т. д.

Для многих задач психодиагностики параллельные формы психодиагностических процедур отсутствуют, поскольку соблюсти все эти требования очень трудно. Тем не менее, существуют, например, параллельные формы теста Айзенка для диагностики интеллекта, опросника Кеттелла для диагностики свойств личности.

Метод расщепления Метод расщепления, который используется для определения надежности психодиагностических процедур, является, по существу, развитием метода параллельного тестирования с помощью эквивалентных форм. Он базируется на допущении о том, что не только вся диагностическая процедура в целом может иметь параллельную (эквивалентную) форму, но и отдельные задания внутри диагностической процедуры могут рассматриваться как отдельные варианты или параллельные формы теста. В связи с этим любой тест, содержащий множество заданий, можно "расщепить", например, на две части и каждую из них рассматривать как параллельную форму по от-

ношению друг к другу. Однако скомпоновать две части можно разными способами, и вероятнее всего каждый способ будет давать разные численные оценки надежности теста.

В принципе тест с множеством заданий можно расщепить на три и более частей. Если все части теста согласуются между собой, имея высокие коэффициенты взаимной корреляции, это свидетельствует о его высокой надежности. С другой стороны, внутренняя согласованность отдельных частей психодиагностической процедуры свидетельствует о высокой гомогенности (однородности) ее содержания, т. е. о том, что каждая ее часть (каждое задание) устойчиво измеряет одно и то же психологическое свойство или качество.

В случае негомогенности содержания психодиагностической процедуры разные задания направлены на диагностику разных аспектов диагностируемого качества, и одинаковые результаты диагностики у разных испытуемых допускают различные толкования. Рассмотрим, например, случай, когда при диагностике уровня интеллектуального развития двое испытуемых показывают один и тот же результат – 50 % выполненных заданий. Значит ли это, что у них одинаковый уровень интеллекта? Если бы содержание теста интеллекта было гомогенным, т. е. все задания имели бы примерно одинаковую степень сложности, то ответ был бы положительным. Однако, если тест негомогенный, возможен вариант, когда один испытуемый выполнил более сложные задания, а другой – более легкие. В этом случае равное количество выполненных заданий еще не говорит о равенстве уровней интеллекта.

Очевидно, что психодиагностическая процедура, содержащая большее количество однородных заданий, является более надежной, так как при этом уменьшается дисперсия случайной составляющей измерения

$S_{e_i}^2$, а при ее уменьшении, согласно соотношению (3.5), коэффициент надежности $R^2(x, x_{ист})$ повышается.

На величину коэффициента надежности диагностической процедуры оказывает существенное влияние увеличение или уменьшение количества

заданий, входящих в состав этой процедуры. Изменение надежности в этом случае описывается формулой Спирмена-Брауна:

$$r = \frac{a \cdot r'}{1 + (a - 1) \cdot r'} \quad (3.8)$$

где r' - исходное значение коэффициента надежности;

r - значение коэффициента надежности после изменения количества заданий;

a - степень изменения количества заданий теста, определяемая отношением $a = k/k'$, в котором k и k' - число заданий, соответственно, до и после изменений.

При расщеплении диагностической процедуры с множеством заданий на две части переход от расщепленной части к полной форме означает удвоение числа заданий. Соответственно, формула для оценки надежности диагностической процедуры методом расщепления на две части принимает вид:

$$r = \frac{2r'}{1 + r'} \quad (3.9)$$

П.Ж. Рюлон предложил другой способ оценки надежности психодиагностической процедуры методом расщепления, исходя из дисперсии S_d^2 разностей между результатами каждого испытуемого по обоим расщепленным частям диагностической процедуры и дисперсии S_x^2 суммарных результатов.

Любую разницу между результатами испытуемого по двум расщепленным частям диагностической процедуры можно рассматривать как некоторую ошибку. Дисперсия таких разностей, поделенная на дисперсию общих результатов, дает долю дисперсии ошибки. Вычитая долю дисперсии ошибки из единицы, получаем долю истинной дисперсии, которая и равна коэффициенту надежности. Формула Рюлона для данного случая имеет вид:

$$R^2(x, x_{ucm}) = 1 - \frac{S_d^2}{S_x^2} \quad (3.10)$$

Расщепление психодиагностической процедуры на две части наиболее распространенный вариант при определении ее надежности. Однако, в

принципе, если диагностическая процедура содержит множество заданий, количество расщепленных частей может быть равно количеству заданий. В этом случае для оценки надежности психодиагностической процедуры определяют дисперсию результатов по каждому заданию и далее вычисляют коэффициент надежности по формуле Кронбаха:

$$R^2(x, x_{uct}) = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k S_i^2}{S_x^2} \right) \quad (3.11)$$

где: S_i^2 – дисперсия результатов по отдельному i -ому заданию;

k – число заданий.

В связи с тем, что вычисление надежности по формуле Кронбаха предполагает вычисление дисперсий S_i^2 , а наличие дисперсий возможно лишь для измерительных шкал интервального типа, данная формула применима только для тех диагностических процедур, где выполнение каждого отдельного задания оценивается в интервальных шкалах. Если выполнение заданий оценивается в дихотомических шкалах (например, "0 – 1", "верно – неверно"), коэффициент надежности психодиагностической процедуры вычисляется по формуле Кьюдера-Ричардсона, которая имеет вид:

$$R^2(x, x_{uct}) = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k p_i q_i}{S_x^2} \right) \quad (3.12)$$

где: p_i и q_i – доли испытуемых, выполнивших и, соответственно, не выполнивших задание (или выполнивших их по-разному);

k – число заданий;

S_x^2 – дисперсия общих результатов диагностики.

Для каждого коэффициента надежности необходимо давать полное описание типа и особенностей выборки, на которой он определялся. Это связано с тем, что приводимый коэффициент надежности применим только к тем выборкам, которые подобны (эквивалентны) выборке, на которой

он был определен. Для других выборок рассчитываются свои значения коэффициентов надежности.

Низкий коэффициент надежности психодиагностической процедуры ограничивает возможности определения корреляционных связей между измеряемыми показателями.

Допустим, имеется два распределения значений измеряемого показателя: а) эмпирическое распределение $x_{эм}$ и б) распределение истинных значений показателя $x_{ист}$. По определению, коэффициент корреляции этих распределений характеризует индекс надежности измерительной процедуры, с помощью которой было получено эмпирическое распределение, т. е. значение $R_{x, x_{ист}}$. Возьмем далее эмпирическое распределение значений любого другого показателя, например y . Очевидно, что распределение значений y будет отличаться от распределения истинных значений $x_{ист}$. Следовательно, при определении корреляции показателя x с любым другим показателем, например y , мы не можем получить коэффициент корреляции r_{xy} выше, чем значение $R_{x, x_{ист}}$. Другими словами, коэффициент корреляции r_{xy} ограничен степенью надежности измерительных процедур, которые используются для измерения значений коррелируемых показателей:

$$r_{xy} \leq r_{x_{ист}y} \cdot \sqrt{R_{x, x_{ист}}^2} \quad (3.13)$$

Если, например, коэффициент надежности измерительной процедуры равен 0,8 (обычно надежность измерительных процедур в психологических исследованиях бывает меньше), то предельное значение r_{xy} не может превышать значение 0,894. При $R_{x, x_{ист}} = 0,7$ коэффициент корреляции r_{xy} не превышает значения 0,837.

Допустим теперь, что, используя измерительные процедуры с неизвестными значениями коэффициента надежности, мы провели измерение показателей памяти и интеллекта на достаточно большой выборке испытуемых и вычислили коэффициент корреляции между значениями этих показателей, который оказался равным, например, 0,26 при достоверности на уровне $\alpha = 0,01$. Какой вывод на основании этих результатов можно

сделать? Первым напрашивается вывод о том, что между измеренными показателями существует весьма слабая связь. Однако, возможен и другой вывод – что при проведении соответствующих измерений были использованы психодиагностические процедуры с очень низкими показателями надежности. Этот пример показывает, что для того, чтобы иметь возможность делать достоверные выводы, необходимо использовать не только валидные, но и достаточно надежные тесты.

3.4. Дискриминативность

В психодиагностике понятие дискриминативности вводится для характеристики способности психодиагностической процедуры дифференцировать (различать) объекты по измеряемым показателям. Если, например, на выборке из N испытуемых производится диагностика объема памяти по шкале, имеющей всего три градации – низкий, средний, высокий, то, скорее всего, в каждой группе испытуемых, несмотря на то, что они отнесены к одной группе, найдутся испытуемые все-таки отличающиеся между собой по измеряемому показателю. Следовательно, данная диагностическая процедура с измерительной шкалой с тремя градациями недостаточно четко различает испытуемых по показателю объема памяти, или, другими словами, эта процедура обладает низкой дискриминативностью.

Понятие дискриминативности непосредственно соотносится с понятием чувствительности психодиагностической процедуры и с уровнем вариативности значений измеряемого показателя. Чем выше вариативность значений измеряемого показателя, тем выше дискриминативность психодиагностической процедуры. С другой стороны, высокая степень вариативности значений показателя является предпосылкой низкой надежности психодиагностической процедуры. Следовательно, дискриминативность психодиагностической процедуры в какой-то степени противостоит ее надежности, что приводит к возникновению, так называемого, психометрического парадокса. Психометрический парадокс наиболее ярко проявляется при использовании личностных опросников и других измерительных процедур, характеризующихся множественностью заданий. Суть его состоит в том, что вопросы (утверждения, задания), имеющие высокий показатель дискриминативности, являются неустойчивыми по отношению к повторному обследованию, и, наоборот, в ответах на вопросы (утвержде-

ния, задания), обладающие низкой дискриминативностью, наблюдается высокая стабильность. Вариация в ответах испытуемых, характеризующая дискриминативность измерительной процедуры, может обуславливаться тремя факторами:

- индивидуальными различиями в степени выраженности измеряемого качества,
- индивидуальными различиями в понимании значения вопроса,
- индивидуальными различиями в степени трудности (легкости) принятия решения об ответе.

Для измерения дискриминативности вводится понятие отношения между индивидуальными значениями измеряемого показателя любых двух испытуемых: эти значения либо равны, либо различны. Общее количество возможных отношений между парами испытуемых в выборке из N испытуемых равно $\frac{N \cdot (N-1)}{2}$.

Общее количество отношений "равенства" среди этих пар составляет:

$$K_{=} = \frac{\sum f_i^2 - \sum f_i}{2} \quad (3.14)$$

где f_i - частота встречаемости совпадающих значений показателя, $\sum f_i = N$.

Общее количество отношений "различия" составляет:

$$K_{\neq} = \frac{(\sum f_i)^2 - \sum f_i^2}{2} = \frac{N^2 - \sum f_i^2}{2} \quad (3.15)$$

Наибольшее количество отношений "различия" достигается тогда, когда все значения показателя имеют одинаковую частоту встречаемости $f_i = \frac{N}{k}$, где k – количество вопросов (утверждений, заданий) или количество

граждений измеряемого показателя, и составляет:

$$K_{\max} = \frac{N^2 - N^2 \cdot \frac{1}{k}}{2} \quad (3.16)$$

Дискриминативность психодиагностической процедуры определяется показателем дельта Фергюсона δ , который равен:

$$d = \frac{K_{\neq}}{K_{\max}} = \frac{N^2 - \sum f_i^2}{N^2 - \frac{N^2}{k}} = \frac{k}{k-1} \left(1 - \frac{\sum f_i^2}{N^2} \right) \quad (3.17)$$

При максимально возможной вариативности индивидуальных значений измеряемого показателя показатель дискриминативности измерительной процедуры дельта Фергюсона равен 1. Если все индивидуальные значения показателя равны между собой, показатель дельта Фергюсона равен 0, что указывает на низкую (нулевую) дискриминативность измерительной процедуры.

В качестве примера использования соотношения (3.17) вычислим дискриминативность психодиагностической процедуры, с помощью которой производилась диагностика уровня тревожности в выборках из 15 мальчиков и 15 девочек (см. таблица 2.13). Измерения проводились методом экспертных оценок по 10-ти балльной шкале. Результаты измерений можно представить в виде следующей таблицы:

*Частоты встречаемости значений уровня тревожности
в выборках из мальчиков и девочек*

Таблица 3.1

Выборка	Градации значений уровня тревожности									
	1	2	3	4	5	6	7	8	9	10
Мальчики	-	-	4	6	5	-	-	-	-	-
Девочки	-	-	-	-	3	5	5	2	-	-

Для выборки из мальчиков имеем $\sum f_i = 4^2 + 6^2 + 5^2 = 77$, для выборки из девочек $\sum f_i = 3^2 + 5^2 + 5^2 + 2^2 = 63$, для двух выборок $\sum f_i = 4^2 + 6^2 + 8^2 + 5^2 + 5^2 + 2^2 = 170$.

В соответствии с выражением (3.17) находим показатель дискриминативности экспертных оценок по 10-ти балльной шкале:

$$\text{для выборки из мальчиков } d = \frac{10}{9} \left(1 - \frac{77}{15^2} \right) = 0,73$$

$$\text{для выборки из девочек } d = \frac{10}{9} \left(1 - \frac{63}{15^2} \right) = 0,80$$

$$\text{для двух выборок } d = \frac{10}{9} \left(1 - \frac{170}{30^2} \right) = 0,90.$$

Как показывают расчеты, дискриминативность психодиагностической процедуры, использованной при измерении уровня тревожности, достаточно высокая, причем для выборки из девочек несколько выше, чем для выборки из мальчиков.

Наличие различий в показателях дискриминативности психодиагностической процедуры (в данном случае процедуры экспертных оценок), полученных на разных выборках испытуемых, может обуславливаться разными факторами и, соответственно, может объясняться по-разному. Например, это можно объяснить наличием действительно более высокой вариативности истинных значений уровня тревожности у девочек по сравнению с мальчиками, что и выявляется при диагностике. Однако различия могут быть обусловлены и более пристрастным отношением эксперта к оценке уровня тревожности девочек на фоне более индифферентного отношения к мальчикам. В этом случае различия в показателях дискриминативности психодиагностической процедуры могут объясняться высоким уровнем субъективности данной процедуры.

Рассмотренный нами случай является типичным примером, свидетельствующим о том, что для получения достоверных психодиагностических данных необходимо не просто выбрать подходящую психодиагностическую процедуру, удовлетворяющую требованиям валидности и надежности, но и учитывать возможность влияния на результаты психодиагностики множества сопутствующих факторов. В свою очередь, это требует наличия у психодиагноста не только навыков работы с различными методиками и тестами, но и глубоких знаний в области теоретической психологии.

ЛИТЕРАТУРА

1. *Анастаси А.* Психологическое тестирование: Книга 1 /Пер. с англ.. – М.: Педагогика, 1982. –320 с.
2. *Батаршев А.В.* Тестирование: Основной инструментарий практического психолога: Учебное пособие. – 2-е изд. –М.: Дело, 2001. – 240 с.
3. *Бородкин Л.И., Промахина И.М.* Методы многомерного статистического анализа //Количественные методы в исторических исследованиях. –М.: Высш. шк., 1984. - С. 267-275.
4. *Витцлак Г.* Основы психодиагностики //Психодиагностика: Теория и практика /Пер. с нем. –М.: Прогресс, 1986. –С. 27-122.
5. *Гайда В.К., Захаров В.П.* Психологическое тестирование: учебное пособие. - Л.: Изд-во ЛГУ, 1982. –101 с.
6. *Гласс Дж., Стэнли Дж.* Статистическим методы в педагогике и психологии. –М.: Прогресс, 1976. –496 с.
7. *Гусев А.Н., Измайлов Ч.А., Михалевская М.Б.* Измерение в психологии: общий психологический практикум. М.: Смысл, 1987. – 287 с.
8. *Дружинин В.Н.* Экспериментальная психология: Учебное пособие. – М.: ИНФРА-М, 1997. –256 с.
9. *Дюк В.А.* Компьютерная психодиагностика. –СПб.: Изд-во "Братство", 1994. –364 с.
10. *Кендэлл М.* Ранговые корреляции. –М.: Статистика, 1975. – 216 с.
11. *Клайн П.* Справочное руководство по конструированию тестов /Пер. с англ. –Киев, 1994. –284 с.
12. *Количественные методы в исторических исследованиях* //Под ред. И.Д. Ковальченко. –М.: Высшая школа, 1984. –384 с.
13. *Кулагин Б.В.* Основы профессиональной психодиагностики. –Л.: Медицина, 1984. – 216 с.
14. *Лидерс А.Г.* Особенности позиции практического психолога по отношению к клиенту //Журнал практ. психолога. –1996. – № 3. –С.15.
15. *Общая психодиагностика* /Под ред. А.А. Бодалева, В.В. Столина. –М.: Изд-во МГУ, 1987.
16. *Психодиагностические методы* (в комплексном лонгитюдном исследовании студентов. –Изд-во ЛГУ, 1976. –248 с.

17. *Психологическая диагностика: Учебное пособие* /Под ред. К.М. Гуревича, Е.М. Борисовой. – М.: Изд-во УРАО, 2000. –304 с.
18. *Садыкова С.З. Взаимосвязь доминирующего цвета и типа темперамента в изобразительной деятельности детей: дисс. канд. психол. наук.* – Казань, 2002.
19. Сидоренко Е.В. Методы математической обработки в психологии. – СПб.: ООО "Речь", 2002. –350 с.
20. *Стивенс С.С. Математика, измерение и психофизика* //Экспериментальная психология /Под ред. С.С. Стивенса. –М.: Иностран. лит-ра, 1960. –Т.1. – С.19-92.

*Таблица перевода процентных соотношений (P)
в стандартные отклонения (z)*

P (%)	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0		-3,09	-2,88	-2,75	-2,65	-2,58	-2,51	-2,46	-2,41	-2,37
1	-2,33	-2,29	-2,26	-2,22	-2,20	-2,17	-2,14	-2,12	-2,10	-2,07
2	-2,05	-2,03	-2,01	-2,00	-1,98	-1,96	-1,94	-1,93	-1,91	-1,90
3	-1,88	-1,87	-1,85	-1,84	-1,83	-1,81	-1,80	-1,79	-1,77	-1,76
4	-1,75	-1,74	-1,73	-1,72	-1,71	-1,70	-1,68	-1,67	-1,66	-1,65
5	-1,64	-1,64	-1,63	-1,62	-1,61	-1,60	-1,59	-1,58	-1,57	-1,56
6	-1,55	-1,55	-1,54	-1,53	-1,52	-1,51	-1,50	-1,50	-1,49	-1,48
7	-1,48	-1,47	-1,46	-1,45	-1,45	-1,44	-1,43	-1,43	-1,42	-1,41
8	-1,41	-1,40	-1,39	-1,38	-1,38	-1,37	-1,37	-1,36	-1,35	-1,35
9	-1,34	-1,33	-1,33	-1,32	-1,32	-1,31	-1,30	-1,30	-1,29	-1,29
10	-1,28	-1,28	-1,27	-1,26	-1,26	-1,25	-1,25	-1,24	-1,23	-1,23
11	-1,22	-1,22	-1,22	-1,21	-1,21	-1,20	-1,20	-1,19	-1,19	-1,18
12	-1,18	-1,17	-1,17	-1,16	-1,16	-1,15	-1,15	-1,14	-1,14	-1,13
13	-1,13	-1,12	-1,12	-1,11	-1,11	-1,10	-1,10	-1,09	-1,09	-1,08
14	-1,08	-1,08	-1,07	-1,07	-1,06	-1,06	-1,05	-1,05	-1,05	-1,04
15	-1,04	-1,03	-1,03	-1,02	-1,02	-1,01	-1,01	-1,00	-1,00	-1,00
16	-0,99	-0,99	-0,99	-0,98	-0,98	-0,97	-0,97	-0,97	-0,96	-0,96
17	-0,95	-0,95	-0,95	-0,94	-0,94	-0,93	-0,93	-0,93	-0,92	-0,92
18	-0,91	-0,91	-0,91	-0,90	-0,90	-0,90	-0,89	-0,89	-0,89	-0,88
19	-0,88	-0,87	-0,87	-0,87	-0,86	-0,86	-0,86	-0,85	-0,85	-0,84
20	-0,84	-0,84	-0,83	-0,83	-0,83	-0,82	-0,82	-0,82	-0,81	-0,81
21	-0,80	-0,80	-0,80	-0,79	-0,79	-0,79	-0,78	-0,78	-0,78	-0,77
22	-0,77	-0,77	-0,77	-0,76	-0,76	-0,76	-0,75	-0,75	-0,75	-0,74
23	-0,74	-0,74	-0,73	-0,73	-0,73	-0,72	-0,72	-0,72	-0,71	-0,71
24	-0,71	-0,70	-0,70	-0,70	-0,69	-0,69	-0,69	-0,68	-0,68	-0,68
25	-0,67	-0,67	-0,67	-0,66	-0,66	-0,66	-0,66	-0,65	-0,65	-0,65
26	-0,64	-0,64	-0,64	-0,63	-0,63	-0,63	-0,62	-0,62	-0,62	-0,61
27	-0,61	-0,61	-0,61	-0,60	-0,60	-0,60	-0,59	-0,59	-0,59	-0,59
28	-0,58	-0,58	-0,58	-0,57	-0,57	-0,57	-0,56	-0,56	-0,56	-0,55
29	-0,55	-0,55	-0,55	-0,54	-0,54	-0,54	-0,54	-0,53	-0,53	-0,53
30	-0,52	-0,52	-0,52	-0,52	-0,51	-0,51	-0,51	-0,50	-0,50	-0,50
31	-0,49	-0,49	-0,49	-0,49	-0,48	-0,48	-0,48	-0,48	-0,47	-0,47
32	-0,47	-0,47	-0,46	-0,46	-0,46	-0,45	-0,45	-0,45	-0,44	-0,44
33	-0,44	-0,44	-0,43	-0,43	-0,43	-0,43	-0,42	-0,42	-0,42	-0,41
34	-0,41	-0,41	-0,41	-0,40	-0,40	-0,40	-0,40	-0,39	-0,39	-0,39
35	-0,39	-0,38	-0,38	-0,38	-0,37	-0,37	-0,37	-0,37	-0,36	-0,36

Приложение 1 (продолжение)

P (%)	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
36	-0,36	-0,36	-0,35	-0,35	-0,35	-0,35	-0,34	-0,34	-0,34	-0,33
37	-0,33	-0,33	-0,33	-0,32	-0,32	-0,32	-0,32	-0,31	-0,31	-0,31
38	-0,31	-0,30	-0,30	-0,30	-0,29	-0,29	-0,29	-0,29	-0,28	-0,28
39	-0,28	-0,28	-0,27	-0,27	-0,27	-0,27	-0,26	-0,26	-0,26	-0,25
40	-0,25	-0,25	-0,25	-0,24	-0,24	-0,24	-0,24	-0,23	-0,23	-0,23
41	-0,23	-0,22	-0,22	-0,22	-0,22	-0,21	-0,21	-0,21	-0,21	-0,20
42	-0,20	-0,20	-0,20	-0,19	-0,19	-0,19	-0,19	-0,18	-0,18	-0,18
43	-0,18	-0,17	-0,17	-0,17	-0,16	-0,16	-0,16	-0,16	-0,15	-0,15
44	-0,15	-0,15	-0,14	-0,14	-0,14	-0,14	-0,13	-0,13	-0,13	-0,13
45	-0,12	-0,12	-0,12	-0,12	-0,11	-0,11	-0,11	-0,11	-0,10	-0,10
46	-0,10	-0,10	-0,09	-0,09	-0,09	-0,09	-0,08	-0,08	-0,08	-0,08
47	-0,07	-0,07	-0,07	-0,07	-0,06	-0,06	-0,06	-0,06	-0,05	-0,05
48	-0,05	-0,05	-0,04	-0,04	-0,04	-0,04	-0,04	-0,03	-0,03	-0,03
49	-0,03	-0,02	-0,02	-0,02	-0,01	-0,01	-0,01	-0,01	0,00	0,00
50	0,00	0,00	0,00	0,01	0,01	0,01	0,01	0,02	0,02	0,02
51	0,03	0,03	0,03	0,03	0,04	0,04	0,04	0,04	0,04	0,05
52	0,05	0,05	0,05	0,06	0,06	0,06	0,06	0,07	0,07	0,07
53	0,07	0,08	0,08	0,08	0,08	0,09	0,09	0,09	0,09	0,10
54	0,10	0,10	0,10	0,11	0,11	0,11	0,11	0,12	0,12	0,12
55	0,12	0,13	0,13	0,13	0,13	0,14	0,14	0,14	0,14	0,15
56	0,15	0,15	0,15	0,16	0,16	0,16	0,16	0,17	0,17	0,17
57	0,18	0,18	0,18	0,18	0,19	0,19	0,19	0,19	0,20	0,20
58	0,20	0,20	0,21	0,21	0,21	0,21	0,22	0,22	0,22	0,22
59	0,23	0,23	0,23	0,23	0,24	0,24	0,24	0,24	0,25	0,25
60	0,25	0,25	0,26	0,26	0,26	0,27	0,27	0,27	0,27	0,28
61	0,28	0,28	0,28	0,29	0,29	0,29	0,29	0,30	0,30	0,30
62	0,31	0,31	0,31	0,31	0,32	0,32	0,32	0,32	0,33	0,33
63	0,33	0,33	0,34	0,34	0,34	0,35	0,35	0,35	0,35	0,36
64	0,36	0,36	0,36	0,37	0,37	0,37	0,37	0,38	0,38	0,38
65	0,39	0,39	0,39	0,39	0,40	0,40	0,40	0,40	0,41	0,41
66	0,41	0,41	0,42	0,42	0,42	0,43	0,43	0,43	0,43	0,44
67	0,44	0,44	0,44	0,45	0,45	0,45	0,46	0,46	0,46	0,47
68	0,47	0,47	0,47	0,48	0,48	0,48	0,48	0,49	0,49	0,49
69	0,49	0,50	0,50	0,50	0,51	0,51	0,51	0,52	0,52	0,52
70	0,52	0,53	0,53	0,53	0,54	0,54	0,54	0,54	0,55	0,55

Приложение 1 (продолжение)

P (%)	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
71	0,55	0,55	0,56	0,56	0,56	0,57	0,57	0,57	0,58	0,58
72	0,58	0,59	0,59	0,59	0,59	0,60	0,60	0,60	0,61	0,61
73	0,61	0,61	0,62	0,62	0,62	0,63	0,63	0,63	0,64	0,64
74	0,64	0,65	0,65	0,65	0,66	0,66	0,66	0,66	0,67	0,67
75	0,67	0,68	0,68	0,68	0,69	0,69	0,69	0,70	0,70	0,70
76	0,71	0,71	0,71	0,72	0,72	0,72	0,73	0,73	0,73	0,74
77	0,74	0,74	0,75	0,75	0,75	0,76	0,76	0,76	0,77	0,77
78	0,77	0,77	0,78	0,78	0,78	0,79	0,79	0,79	0,80	0,80
79	0,80	0,81	0,81	0,82	0,82	0,82	0,83	0,83	0,83	0,84
80	0,84	0,84	0,85	0,85	0,86	0,86	0,86	0,87	0,87	0,87
81	0,88	0,88	0,89	0,89	0,89	0,90	0,90	0,90	0,91	0,91
82	0,91	0,92	0,92	0,93	0,93	0,93	0,94	0,94	0,95	0,95
83	0,95	0,96	0,96	0,97	0,97	0,97	0,98	0,98	0,99	0,99
84	0,99	1,00	1,00	1,01	1,01	1,01	1,02	1,02	1,03	1,03
85	1,03	1,04	1,04	1,05	1,05	1,05	1,06	1,06	1,08	1,08
86	1,08	1,09	1,09	1,10	1,10	1,10	1,11	1,11	1,12	1,12
87	1,13	1,13	1,14	1,14	1,15	1,15	1,16	1,16	1,17	1,17
88	1,18	1,18	1,19	1,19	1,20	1,20	1,21	1,21	1,22	1,22
89	1,22	1,23	1,23	1,24	1,25	1,25	1,26	1,26	1,27	1,28
90	1,28	1,29	1,29	1,30	1,30	1,31	1,32	1,32	1,33	1,33
91	1,34	1,35	1,35	1,36	1,37	1,37	1,38	1,38	1,39	1,40
92	1,41	1,41	1,42	1,43	1,43	1,44	1,45	1,45	1,46	1,47
93	1,48	1,48	1,49	1,50	1,50	1,51	1,52	1,53	1,54	1,55
94	1,55	1,56	1,57	1,58	1,59	1,60	1,61	1,62	1,63	1,64
95	1,64	1,65	1,66	1,67	1,68	1,70	1,71	1,72	1,73	1,74
96	1,75	1,76	1,77	1,79	1,80	1,81	1,83	1,84	1,85	1,87
97	1,88	1,90	1,91	1,93	1,94	1,96	1,98	2,00	2,01	2,03
98	2,05	2,07	2,10	2,12	2,14	2,17	2,20	2,22	2,26	2,29
99	2,33	2,37	2,41	2,46	2,51	2,58	2,65	2,75	2,88	3,09

Критические значения *t*-критерия Стьюдента

Число степеней свободы	Уровень достоверности		
	0,90	0,95	0,99
3	2,35	3,18	5,84
4	2,13	2,78	4,60
5	2,02	2,57	4,03
6	1,94	2,45	3,70
7	1,90	2,37	3,50
8	1,86	2,30	3,36
9	1,83	2,26	3,25
10	1,81	2,23	3,17
11	1,80	2,20	3,11
12	1,78	2,18	3,06
13	1,77	2,16	3,02
14	1,76	2,14	2,98
15	1,75	2,13	2,95
16	1,74	2,12	2,92
17	1,74	2,11	2,90
18	1,73	2,10	2,88
19	1,73	2,09	2,86
20	1,72	2,09	2,85
21	1,72	2,08	2,83
22	1,72	2,07	2,82
23	1,71	2,07	2,81
24	1,71	2,06	2,80
25	1,71	2,06	2,79
26	1,71	2,06	2,78
27	1,70	2,05	2,77
28	1,70	2,05	2,76
29	1,70	2,05	2,75
30	1,70	2,04	2,74
40	1,68	2,02	2,70
60	1,67	2,00	2,67
80	1,66	1,99	2,64
100	1,66	1,98	2,62
200	1,65	1,97	2,60

Критические значения $\chi^2_{кр}$

Число степеней свободы	Уровень достоверности		
	0,90	0,95	0,99
1	2,71	3,84	6,63
2	4,61	5,99	9,21
3	6,25	7,81	11,3
4	7,78	9,49	13,3
5	9,24	11,1	15,1
6	10,6	12,6	16,8
7	12,0	14,1	18,5
8	13,4	15,5	20,1
9	14,7	16,9	21,7
10	16,0	18,3	23,2
11	17,3	19,7	24,7
12	18,5	21,0	26,2
13	19,8	22,4	27,7
14	21,1	23,7	29,1
15	22,3	25,0	30,6
16	23,5	26,3	32,0
17	24,8	27,6	33,4
18	26,0	28,9	34,8
19	27,2	30,1	36,2
20	28,4	31,4	37,6
21	29,6	32,7	38,9
22	30,8	33,9	40,3
23	32,0	35,2	41,6
24	33,2	36,4	43,0
25	34,4	37,7	44,3
26	35,6	38,9	45,6
27	36,7	40,2	47,0
28	37,9	41,3	48,3
29	39,1	42,6	49,6
30	40,3	43,8	50,9

СТРУКТУРА ПЕДАГОГИЧЕСКИХ УСТАНОВОК ЛИЧНОСТИ (СПУЛ)

Тест предназначен для выявления индивидуальной структуры педагогических установок личности на основе ранжирования 18 качеств, характеризующих "хорошего" учителя.

Процедура тестирования предполагает ранжирование качеств по степени их необходимости для учителя (педагога), чтобы его педагогическая деятельность была достаточно эффективна с точки зрения обследуемого.

Результаты тестирования позволяют определить соотношение и уровень доминирования у обследуемого следующих типов и видов педагогических установок:

Типы установок:

1. Авторитарная
2. Демократическая
3. Либеральная

Виды установок:

1. На ученика
2. На учебно-воспитательный процесс
3. На личность учителя

Преобладающий вид установки указывает, на что обращает внимание обследуемый в большей степени, когда речь заходит о педагогической деятельности, что в системе его представлений о педагогической деятельности находится на переднем плане, а что уходит на задний план.

С другой стороны, к составляющим учебно-воспитательного процесса обследуемый может предъявлять различные требования и выражать различное отношение.

Соответственно выделяются три типа педагогических установок:

- авторитарный;
- демократический;
- либеральный.

При конструировании теста для каждого вида и типа педагогических установок были подобраны соответствующие качества личности учителя - по два качества на каждое сочетание вида и типа установок. Всего 18 качеств (см. Перечень качеств личности "хорошего" учителя).

Распределение качеств по видам и типам установок представлено в таблице, которая используется в качестве "ключевой" при обработке результатов тестирования.

"КЛЮЧЕВАЯ" ТАБЛИЦА
к обработке результатов тестирования по тесту СПУЛ

Тип установки	В и д у с т а н о в к и			Сумма рангов
	1. На ученика	2. На УВП	3. На учителя	
1. Авторитарная	1, 10	4, 13	7, 16	
2. Демократическая	8, 17	2, 11	5, 14	
3. Либеральная	6, 15	9, 18	3, 12	
Сумма рангов				171

ПЕРЕЧЕНЬ КАЧЕСТВ ЛИЧНОСТИ
"хорошего" учителя

1. Высокая требовательность к успеваемости и нормам поведения учащихся.
2. Выбор учебного материала с учетом способностей и интересов учащихся.
3. Увлеченность своим предметом, склонность к научно-исследовательской работе.
4. Стремление дать весь учебный материал, предусмотренный программой.
5. Способность прислушиваться к мнению коллег и администрации.
6. Стремление к тесным, доверительным взаимоотношениям с учащимися.
7. Наличие устойчивых педагогических взглядов, принципов и убеждений.
8. Признание права учащихся на самостоятельное решение и собственное мнение.
9. Предоставление учащимся возможности самостоятельного выбора учебных предметов.
10. Умение поддерживать дисциплину и порядок в классе.
11. Широкое использование в учебном процессе групповых форм учебной деятельности.
12. Стремление к самообразованию, расширению своего кругозора.
13. Соблюдение в учебном процессе требований государственного стандарта образования.
14. Стремление к совершенствованию стиля педагогической деятельности.
15. Не применение в отношениях с учащимися каких-либо форм принуждения.
16. Стремление к совершенствованию педагогического мастерства, владению различными средствами педагогического воздействия.
17. Предоставление учащимся большей свободы и самостоятельности на уроках.
18. Склонность к инновациям, внедрению современных педагогических технологий.