



**Современный
Гуманитарный
Университет**

Дистанционное образование

Рабочий учебник

Фамилия, имя, отчество _____

Факультет _____

Номер контракта _____

ПСИХОЛОГИЧЕСКИЙ ПРАКТИКУМ

ЮНИТА 7

**ЭКСПЕРИМЕНТ: ВНУТРЕННЯЯ И ВНЕШНЯЯ
ВАЛИДНОСТЬ**

МОСКВА 1999

Разработано А.В.Мощенко, д-ром психологич. наук

Рекомендовано Министерством общего и профессионального образования Российской Федерации в качестве учебного пособия для студентов высших учебных заведений

КУРС: ПСИХОЛОГИЧЕСКИЙ ПРАКТИКУМ

Юнита 1. Общая характеристика психологического исследования.
Юнита 2. Процедуры получения и описания эмпирических данных.
Юнита 3. Метод наблюдения в психологической практике.
Юнита 4. Стандартные способы обработки и представления данных.
Юнита 5. Психометрия: методы одномерного и многомерного шкалирования.
Юнита 6. Корреляционные исследования и квазиэкспериментальные планы.
Юнита 7. Эксперимент: внутренняя и внешняя валидность.
Юнита 8. Психодиагностика: виды надежности и валидности процедур.

ЮНИТА 7

Рассмотрены разновидности эксперимента, внутренняя и внешняя валидность, способы контроля влияния артефактов.

Для студентов Современного Гуманитарного Университета

Юнита соответствует образовательной профессиональной программе №1

ОГЛАВЛЕНИЕ

ДИДАКТИЧЕСКИЙ ПЛАН	4
ЛИТЕРАТУРА	5
ПЕРЕЧЕНЬ УМЕНИЙ	6
НАУЧНЫЙ ОБЗОР	7
1. Эксперимент и его разновидности	7
1.1. Отличия экспериментов в предметной области	7
1.2. Отличия в формах осуществления эксперимента	8
1.3. Идеальный эксперимент	8
1.4. Реальный эксперимент	10
1.5. Критический эксперимент	12
1.6. Демонстрационный эксперимент	13
1.7. Организация и проведение экспериментального исследования	14
1.8. Экспериментальная выборка	22
1.9. Экспериментальные переменные	26
2. Внутренняя и внешняя валидность	39
2.1. Виды валидности при оценке эксперимента	39
2.2. Внутренняя валидность	40
2.3. Внешняя валидность	41
2.4. Проверка каузальных отношений	41
2.5. Факторы, угрожающие внутренней и внешней валидности	44
2.6. План с предварительным и итоговым тестированием и контрольной группой	45
3. Способы контроля влияния артефактов	52
3.1. Формы экспериментального контроля	52
3.2. Контроль за выводом в экспериментальном исследовании	56
3.3. Вывод о гипотезе на основе оценки валидности эксперимента	57
3.4. Вывод о подкреплённости теории экспериментальными фактами	58
3.5. Проблема возникновения новых гипотез	59
3.6. Проблема артефактных выводов	60
ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ	65
ТРЕНИНГ УМЕНИЙ	68
ГЛОССАРИЙ*	

* Глоссарий расположен в середине учебного пособия и предназначен для самостоятельного заучивания новых понятий.

ДИДАКТИЧЕСКИЙ ПЛАН

Эксперимент и его разновидности. Отличия экспериментов в предметной области. Отличия в формах осуществления эксперимента. Идеальный эксперимент. Реальный эксперимент. Критический эксперимент. Демонстрационный эксперимент. Организация и проведение экспериментального исследования. Экспериментальная выборка. Экспериментальные переменные.

Внутренняя и внешняя валидность. Виды валидности при оценке эксперимента. Внутренняя валидность. Внешняя валидность. Проверка каузальных отношений. Факторы, угрожающие внутренней и внешней валидности. План с предварительным и итоговым тестированием и контрольной группой.

Способы контроля влияния артефактов. Формы экспериментального контроля. Контроль за выводом в экспериментальном исследовании. Вывод о гипотезе на основе оценки валидности эксперимента. Вывод о подкреплённости теории экспериментальными фактами. Проблема возникновения новых гипотез. Проблема артефактных выводов.

ЛИТЕРАТУРА

Базовая

- * 1. Дружинин В.Н. Экспериментальная психология: Учебное пособие. М.: Инфра-М, 1997. 256 с.
- * 2. Кэмпбелл Д. Модели экспериментов в социальной психологии и прикладных исследованиях. СПб.: Социально-психологический центр, 1996. 392 с.

Дополнительная

- * 3. Корнилова Т.В. Введение в психологический эксперимент: Учебник. М.: МГУ, 1997. 256 с.
- 4. Крылов В.Ю. и др. Математические методы в современной психологии. М.: ИП РАН, 1995. 342 с.
- 5. Логвиненко А.Д. Измерения в психологии: математические основы. М., 1993.
- 6. Паповян С.С. Математические методы в социальной психологии. М.: Наука, 1983.
- * 7. Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии: Пер. с англ. М.: Прогресс, 1976. 495 с.
- * 8. Суходольский Г.В. Основы математической статистики для психологов. Л.: Изд-во ЛГУ, 1972. 429 с.
- * 9. Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психологической диагностике / Отв. ред. С.Б.Крымский. Киев: Наукова думка, 1989. 200 с.
- 10. Еремеев Б.А. Статистические процедуры при психологическом изучении текста: Учеб. пособие. СПб.: Образование, 1996. 54 с.
- * 11. Немов Р.С. Психология: В 3-х кн. Кн.3. Психодиагностика. Введение в научное психологическое исследование с элементами математической статистики. М.: Владос, 1998. 632 с.
- 12. Татарова Г.Г. Методология анализа данных в социологии (введение): Учебное пособие для вузов. М.: Изд. дом "Стратегия", 1998. 224 с.

Примечание. Знаком (*) отмечены работы, использованные при составлении научного обзора.

ПЕРЕЧЕНЬ УМЕНИЙ

№ п/п	Умение	Алгоритмы
1	Контроль внешней переменной.	1. Определение влияния внешней переменной на зависимую переменную. 2. Выявление возможности контролировать внешнюю переменную. 3. Применение техники контроля внешней переменной.

1. ЭКСПЕРИМЕНТ И ЕГО РАЗНОВИДНОСТИ

1.1. Отличия экспериментов в предметной области

Экспериментальный метод — метод, который предполагает разработку плана управления независимой переменной.

Классификации типов экспериментов базируются на разных основаниях. Одним из этих оснований, или критериев классификации, является принадлежность исходной гипотезы, условий проведения опытов или выборки испытуемых к специальным областям психологической реальности, т.е. их отнесенность к определенной **предметной области**.

Сложившиеся в той или иной области психологического знания пути перехода от “мира эмпирии” к “миру теорий” включают ориентировку исследователей на обоснованные в рамках конкретных теорий принципы интерпретации психологических закономерностей. Например, проверка гипотезы о роли движений глаз в возникновении иллюзии Мюллера—Лайера традиционно прописана в области психологии восприятия, в то время как анализ факторов, влияющих на идентификацию эмоциональных состояний по фотографиям лиц, — в психологии эмоций. Эксперимент Ла Пьера, показавший разницу проявления этнических предубеждений на вербальном и поведенческом уровнях, традиционно относится к области социальной психологии. Сначала были разосланы в гостиницы запросы о том, будут ли приняты в них китайцы, а потом исследователь реально посетил с китайцами эти гостиницы.

Исследования влияния личностных предпосылок на успешность предпринимательской деятельности должны быть отнесены к развивающейся области - психологии предпринимательства. Но с таким же правом их можно отнести к психологии личности, если рассматривать специфичные условия профессиональной деятельности испытуемых только как конкретный фон проявления личностных факторов регуляции деятельности.

При таком подходе возникает возможность ошибиться в трактовке психологических механизмов, проявившихся в реализованном эксперименте. Классическим примером является так называемый Хоторнский эксперимент, который по его целям следовало бы отнести к психологии труда. Напомним, что в нем изменялись многочисленные факторы внешней среды и оценивалось влияние этих изменений на производительность труда. Как оказалось, любое изменение, независимо от его вида, приводило к улучшению показателей работающих. В качестве адекватной была принята интерпретация, что люди, зная о целях эксперимента, позитивно откликнулись на сам факт попытки улучшить условия их труда. Сейчас этот эксперимент представлен чаще в разделах, посвященных “эффектам экспериментатора”, а не психологии труда, так как проявившаяся в нем закономерность оказалась более общим и проблематичным механизмом.

В то же время забывать о предметной отнесенности психологических гипотез к той или иной области психологических знаний нельзя, поскольку в каждой из этих областей сложились свои специальные нормативы для проверки гипотез, связанные с типом построения самой психологической теории и возможностями сбора эмпирического материала. **Психологические гипотезы** — предположения, сформулированные относительно психологической реальности в рамках использования тех или иных психологических представлений.

* Жирным шрифтом выделены новые понятия, которые необходимо усвоить. Знание этих понятий будет проверяться при тестировании.

Профессионал, занимающийся исследованиями восприятия, будет отличаться от профессионала, посвятившего себя исследованиям в области социальной психологии, именно в освоении форм экспериментального контроля, поскольку эксперимент в каждой из этих областей реализует разные подходы к способам управления переменными, разное понимание причинности и учет иных типов обоснований в организации содержательных выводов.

1.2. Отличия в формах осуществления эксперимента

Эксперимент, проводимый с целью получить эмпирические доводы в пользу или против предполагаемого в гипотезе понимания психологической закономерности, называют **реальным**, или реализованным. Ему противопоставляется **мысленный** эксперимент, т.е. практически не осуществленный. Мысленный эксперимент может быть безупречным и бесконечным.

Безупречный эксперимент — мысленный образец психологического эксперимента, в котором отсутствуют угрозы достоверному выводу.

Бесконечный эксперимент — мысленный образец для оценки реально проводимых экспериментов, в котором число проб или число испытуемых оценивается с точки зрения охвата всех источников варибельности переменных.

Взаимосвязи содержательного и формального планирования в этих типах экспериментов являются общими во всех тех контекстах, которые не учитывают проблемы контроля смешений независимых переменных с побочными. В реализуемых экспериментах нельзя обойтись предположением об отсутствии смешений, а необходимо обеспечить их контроль, чтобы иметь возможность учесть их влияние на полученные экспериментальные эффекты. В мысленных экспериментах можно предположить любые идеальные условия и рассматривать базисные процессы в любых показателях их проявлений, даже если нет пока реальных методик для операционализации переменных. При оценке психологической гипотезы с точки зрения подразумеваемых в мысленном эксперименте результатов нет необходимости соотносить результат действия независимой переменной с оценкой внутренней и операциональной валидности эксперимента.

1.3. Идеальный эксперимент

Идеальный эксперимент — эксперимент, в котором изменению подвергается только независимая переменная, а зависимая переменная контролируется. Другие условия эксперимента остаются неизменными. Идеальный эксперимент предполагает эквивалентность испытуемых, неизменность их характеристик во времени, отсутствие самого физического времени (как это ни парадоксально звучит), возможность проводить эксперимент бесконечно. Следствием этого является проведение всех экспериментальных воздействий одновременно.

Идеальный эксперимент противостоит реальному, в котором изменяются не только интересующие исследователя переменные, но и ряд других условий. Соответствие идеального эксперимента реальному выражается в такой его характеристике, как **внутренняя валидность** — достоверность результатов, которую обеспечивает реальный эксперимент по сравнению с идеальным. Внутренняя валидность характеризует меру влияния на изменение зависимой переменной тех условий (независимой переменной), которые варьирует экспериментатор.

Чем больше влияние на изменение зависимой переменной неконтролируемых исследователем условий, тем ниже внутренняя валидность эксперимента. Следовательно, больше вероятность того, что факты, обнаруженные в эксперименте, являются артефактами.

Планирование эксперимента необходимо для повышения влияния независимой переменной на зависимую за счет контроля других переменных. Переменные, являющиеся источником артефактов, либо устраняются, либо их влияние усредняется (за счет смешения и погашения изменчивости условий эксперимента). Высокая внутренняя валидность — главный признак хорошего эксперимента.

Не все переменные, влияющие на результат исследования, можно учесть или исключить. Те из них, которые нарушают внутреннюю валидность, называются побочными. К числу побочных, полностью неустранимых переменных относят влияние фактора времени, фактора задачи, индивидуальных различий.

Рассмотрим более подробно центральное понятие “валидность”. Обратимся к системе отношений “эксперимент — теория — реальность”. На основе теории выдвигается гипотеза, которая в конечном счете и проверяется в эксперименте. **Гипотеза** — научное предположение в виде высказывания, истинность или ложность которого неизвестна, но может быть проверена опытным путем (эмпирически).

Методики и план эксперимента должны соответствовать проверяемой гипотезе — степень этого соответствия и характеризует **операциональную валидность**. Например, в эксперименте по научению резкий звук зуммера используется как эквивалент “наказания” за ошибку, поэтому возникает сомнение в правильности такой интерпретации. В самом эксперименте необходимо максимально учесть, устранить и т.д. влияние побочных переменных на зависимую переменную. Внутренняя валидность характеризует меру влияния независимой переменной на зависимую по отношению к другим факторам. Иными словами, внутренняя валидность тем выше, чем больше вероятность того, что экспериментальный эффект (изменение зависимой переменной) вызван изменением независимой переменной.

Эксперимент должен воспроизводить внешнюю реальность. Эксперимент, который полностью воспроизводит внешнюю реальность, называется экспериментом полного соответствия. Разумеется, в реальности полное соответствие недостижимо. Мера соответствия экспериментальной процедуры реальности характеризует внешнюю валидность эксперимента.

Дополнительные переменные, которые требуют учета в эксперименте, влияют на внешнюю валидность. Если от внутренней валидности зависит достоверность экспериментальных результатов, то от внешней — переносимость результатов из лабораторных условий на реальные процессы и обобщение их на другие сферы реальности.

Наконец, связь теории и реальности отражается в адекватности теории реальности и прогностичности ее предсказаний. Кэмпбелл ввел еще одно важное понятие, характеризующее валидность эксперимента, а именно — конструктивную валидность. **Конструктивная валидность** выражает адекватность метода интерпретации экспериментальных данных теории, т.е. в структуру следует ввести четвертую составляющую — интерпретацию: теория — эксперимент — интерпретация — реальность.

Конструктивная валидность, по мнению Кэмпбелла, характеризует правильность обозначения (интерпретации) причины и экспериментального эффекта с помощью абстрактных терминов из обыденного языка или формальной теории.

Таким образом, внутренняя валидность определяется достоверностью интерпретации экспериментального эффекта как связи изучаемой причины и следствия (отношение эксперимент — интерпретация), а конструктивная валидность — правильностью употребления терминов той или иной теории при интерпретации данных эксперимента.

Внутренняя валидность является непереносимым условием любого эксперимента. Кэмпбелл выделил восемь основных факторов, нарушающих

внутреннюю валидность эксперимента. Первую группу можно назвать **факторами выборки**:

1) селекция — неэквивалентность групп по составу, которая вызывает систематическую ошибку в результатах;

2) статистическая регрессия — частный случай ошибки селекции, когда группы отбирались на основе “крайних” показателей (иначе — корреляция из-за неоднородности группы);

3) экспериментальный отсев — неравномерное выбывание испытуемых из сравниваемых групп, приводящее к неэквивалентности групп по составу;

4) естественное развитие — изменение испытуемых, являющееся следствием течения времени, без связи с конкретными событиями: изменение состояния (голод, усталость, болезнь и др.), свойств индивида (возрастные перемены, накопление опыта и др.);

Вторая группа — **побочные переменные**, влияние которых приводит к следующим эффектам:

1) эффект “истории” — конкретные события, происходящие в период между начальным и итоговым тестированием, помимо экспериментального воздействия;

2) эффект тестирования — влияние предварительного тестирования на результат итогового;

3) инструментальная погрешность — определяется надежностью метода фиксации поведения испытуемого, т.е. надежностью теста; именно надежность влияет на валидность, по утверждению Кэмпбелла, а не наоборот;

4) взаимодействие факторов: отбора; естественного развития; истории (разные истории экспериментальных групп) и др.

Позже Кэмпбелл описал еще ряд источников нарушения внутренней валидности. Наиболее существенные относятся к экспериментальной процедуре, а именно: компенсаторное сопоставление эффектов различных воздействий, имитация воздействия, когда его на самом деле не происходит, и др.

1.4. Реальный эксперимент

Реальный эксперимент отличается как от эксперимента идеального, так и от самой реальности, которую он призван моделировать. Это отличие фиксируется таким понятием, как “внешняя валидность” психологического эксперимента. В настоящее время факторами, влияющими на внешнюю валидность, считают неустраняемые особенности эксперимента, отличающие его от реальной ситуации.

Внешняя валидность определяет, в какой мере результаты, полученные в эксперименте, будут соответствовать жизненной ситуации, послужившей “первообразом” для эксперимента. Кроме того, внешняя валидность характеризует возможность обобщения, переноса результатов, полученных в эксперименте, на весь класс жизненных ситуаций.

Внешняя валидность иногда трактуется как характеристика эксперимента, определяющая возможность переноса (обобщения) результатов на различные времена, места, условия и группы людей. Однако возможность переноса является следствием двух причин: 1) соответствия условий эксперимента его “первообразной” жизненной ситуации (“репрезентативность” эксперимента); 2) типичности самой “первообразной” ситуации для реальности (“репрезентативность” ситуации). Выбранная для моделирования в эксперименте ситуация может быть совершенно нерепрезентативной с точки зрения жизни той группы испытуемых, которая участвует в эксперименте, или являться редкой и нетипичной.

Внешняя валидность, как определяет ее Готтсданкер, влияет прежде всего на достоверность выводов, которую дают результаты реального эксперимента, по сравнению с экспериментом полного соответствия. Для достижения высокой

внешней валидности нужно, чтобы уровни дополнительных экспериментальных переменных соответствовали их уровням в реальности. Эксперимент, который не имеет внешней валидности, считается неверным. Добавим, что он неверен, если источником гипотезы является реальность, обыденное знание, а не теория. Эксперимент, не соответствующий реальности, может обладать совершенной внутренней и операциональной валидностью. Другое дело, что прямой перенос его результатов в реальность невозможен без учета влияния на зависимую переменную помимо независимой и дополнительных переменных.

Очевидно, что достижение полной внешней валидности невозможно в принципе, поэтому любое “чистое” аналитическое исследование является внешне невалидным. Вместе с тем рекомендуется максимально учитывать влияние дополнительных переменных на экспериментальный эффект, так как неизвестно, когда будет построена теория для их объяснения, а данные, возможно, придется использовать на практике.

Заботу о внешней валидности эксперимента особо проявляют исследователи, работающие в прикладных областях: клинической психологии, педагогической и организационной психологии. Это понятно, ибо для решения своих повседневных задач им чаще приходится прибегать к постановке экспериментов, имитирующих реальность. По сути, историческая дискуссия сторонников лабораторного эксперимента и “естественного эксперимента” была отражением разного методического подхода специалистов, занимавшихся фундаментальной или прикладной психологией. **Лабораторный эксперимент** — эксперимент в специально созданных условиях, позволяющих выделить независимую переменную путем контроля всех других условий, с которыми может смешиваться ее влияние.

Кэмпбелл ставит знак равенства между внешней валидностью, репрезентативностью эксперимента и возможностью генерализации его результатов. Он относит к факторам, угрожающим внешней валидности, в первую очередь эффекты, связанные с особенностями объекта исследования: обучаемость, наличие памяти, способность эмоционально реагировать на ситуации. Кэмпбелл называет главные причины нарушения внешней валидности:

1. Эффект тестирования — уменьшение или увеличение восприимчивости испытуемых к экспериментальному воздействию под влиянием тестирования. Например, предварительный контроль знаний учеников может повысить их интерес к новому учебному материалу. Поскольку генеральная совокупность предварительному тестированию не подвергается, то результаты для нее могут быть нерепрезентативными.

2. Условия проведения исследования. Они вызывают реакцию испытуемого на эксперимент. Следовательно, его данные нельзя переносить на лиц, не принимавших участия в эксперименте; этими лицами является вся генеральная совокупность, кроме экспериментальной выборки.

3. Взаимодействие факторов отбора и содержания экспериментального воздействия. Их следствия — артефакты (в экспериментах с добровольцами или испытуемыми, участвующими по принуждению).

4. Интерференция экспериментальных воздействий. Испытуемые обладают памятью и обучаемостью. Если эксперимент состоит из нескольких серий, то первые воздействия не проходят для них бесследно и сказываются на появлении эффектов от последующих воздействий.

Большинство причин нарушения внешней валидности связано с особенностями психологического эксперимента, проводимого с участием человека, которые отличают психологическое исследование от эксперимента, осуществляемого специалистами в области других естественных наук.

На взаимодействие процедуры тестирования и содержания экспериментального воздействия первым обратил внимание в 1949 г Р.Л.Соломон

при проведении исследования школьников: предварительное тестирование снижало эффективность обучения. Исследование социальных установок показало, что предварительное тестирование оказывало влияние на установки личности и ее подверженность убеждению, а в экспериментах Ховленда, наоборот, оно ослабляло убеждающее воздействие фильмов.

Эффект тем больше, чем необычнее процедура тестирования и чем более сходно по содержанию экспериментальное воздействие с тестом. Для того чтобы избежать эффекта предварительного тестирования, Кэмпбелл рекомендует применять экспериментальные планы с предварительно не тестируемыми группами.

Как уже отмечалось, “взаимодействие состава группы и воздействия” связано с неслучайным участием испытуемых в эксперименте. Реакция может быть двух типов: готовность добровольцев “подвергнуться воздействию” и отказ, негативная реакция тех, кто принудительно привлечен к участию в эксперименте. В исследованиях по научению могут согласиться участвовать только интеллектуально одаренные люди. Выбывание испытуемых в ходе эксперимента может вызываться экспериментальным воздействием. К примеру, лица, потерпевшие неудачу при выполнении заданий на мотивацию достижения, могут отказаться от участия в последующих сериях.

Естественно, практически невозможно исключить лишь фактор “реакции на эксперимент”. Еще раз отметим, что проблема внутренней валидности разрешима в принципе, так как можно подобрать соответствующие процедуры планирования эксперимента и математической обработки результатов, чтобы обеспечить заданный уровень их достоверности.

Согласно Кэмпбеллу, проблема внешней валидности как репрезентативности эксперимента по отношению к реальности неразрешима, так как индукция, т.е. обобщение, никогда не может быть полностью объективна.

Проблема внешней валидности как адекватности ситуации эксперимента его “первообразной” жизненной ситуации также неразрешима логическими и математическими средствами. Она требует привлечения всей совокупности научных психологических знаний для описания ситуации в целом.

1.5. Критический эксперимент

По отношению к эксперименту с проверкой одного теоретического предположения, где эмпирические данные “за” и “против” рассматриваются в рамках одной и той же объяснительной схемы, можно выделить **критический эксперимент**. В последнем случае предполагается возможность выбора на основании полученных опытных данных между разными объяснительными схемами, т.е. возможность предпочтения одной из сравниваемых конкурирующих между собой теоретических интерпретаций. В этом случае за экспериментальной и противоположной ей контргипотезой стоят разные психологические объяснения, т.е. предполагается проверка двух теоретических гипотез.

Критический эксперимент проводится для того, чтобы одновременно проверить все возможные гипотезы. Подтверждение одной из них ведет к опровержению всех других возможных альтернатив. Для критического эксперимента положительный и отрицательный исходы связываются с разными интерпретационными схемами, что позволяет делать выбор между разными теоретическими объяснениями. Постановка критического эксперимента в психологии требует не только тщательного его планирования, но и высокого уровня разработки научной теории. Поскольку в нашей науке преобладают не дедуктивные модели, а эмпирические обобщения, исследователи крайне редко проводят критический эксперимент.

Обычно в эксперименте сопоставляются две гипотезы, предполагающие

положительный и отрицательный исходы в рамках одной и той же интерпретации. Контргипотеза звучит как отрицание связи, постулируемой в экспериментальной гипотезе, но не как ее объяснение в рамках другой интерпретационной схемы. Такой эксперимент обычно называется **контрольным**, поскольку в нем создаются условия, при которых равновероятным было бы получение данных как в пользу, так и против экспериментальной гипотезы. То есть контролируются условия, при которых именно действенность предполагавшейся причинно-следственной зависимости определяет изменение показателей зависимой переменной в ожидаемом направлении.

1.6. Демонстрационный эксперимент

Традиционному контрольному эксперименту, направленному на проверку причинно-следственной гипотезы, могут быть противопоставлены так называемые **демонстрационные эксперименты**, которые лучше было бы называть демонстрационными опытами, а не экспериментами, поскольку в них нельзя при получении отрицательных результатов перейти к оценке контргипотезы. Условия для получения данных в направлении утверждений, подразумеваемых в контргипотезах, здесь отсутствуют. То есть нет этапа принятия решения об экспериментальном факте, поскольку нет выбора в интерпретации между альтернативами “за” и “против”. Наиболее известны и типичны два вида демонстрационных опытов. Первый представлен инициированием эффектов, которые практически однозначно воспроизводимы при аналогичной ситуации (или аналогичной организации воздействий) любым человеком, выполняющим инструкции по проведению, т.е. ожидаемый эффект наблюдается практически в 100% случаев. Отличным примером здесь являются перцептивные феномены. Так, при условии наличия белого экрана и установки, позволяющей проецировать на него монохроматические цвета, а также знании экспериментатором законов смешения цветов любой исследователь получает однозначные результаты. Конечно, если испытуемые не являются цветоаномалами.

В таких опытах не могут быть получены данные против той или иной гипотезы; эти данные однозначно свидетельствуют в пользу демонстрируемых законов и решают скорее диагностическую задачу — проявления подразумеваемой психологической реальности. Воспроизводимость этих опытов означает возможность переноса решения этой диагностической задачи в любом будущем промежутке времени, по отношению к которому можно утверждать отсутствие изменений в самом базисном процессе, подвергающемся “проявлению”. **Базисный процесс** — реконструируемый в теории или в гипотетических конструктах процесс, определяющий показатели зависимой переменной в психологическом эксперименте.

Второй вид демонстрационных опытов характеризуется тем, что однозначно может демонстрировать правильность или адекватность психологической гипотезы только по отношению к уже состоявшемуся событию. Обычно это демонстрация феноменов, связанных с мотивационной или волевой регуляцией действий человека, с групповой динамикой и с феноменами в других областях психологической реальности, которые возникают “здесь и сейчас”. Их повторно нельзя воспроизвести в том смысле, что их актуал-генез зависит от активности познающего или действующего субъекта, от личностных, а не только стимульных факторов. Например, в опытах, демонстрирующих особенности просоциального поведения, испытуемый подвергается такому “воздействию”, как крики из соседней комнаты. Экспериментатор может наблюдать, бросится ли человек на помощь. Предполагается, что он не догадывается о том, что крики мнимые, записаны ранее и воспроизводятся с помощью магнитофона. Но заранее исследователь не может предполагать, как поступит каждый конкретный

испытуемый, т.е. он не может продемонстрировать ожидаемый эффект наверняка. Однако если ожидаемое поведение все же имело место, то обратная интерпретация будет однозначной — имел место акт просоциального поведения.

1.7. Организация и проведение экспериментального исследования

Прежде всего обратимся к множеству психологических эмпирических методов. Рассмотрим их с позиций общенаучных критериев, предъявляемых к организации исследования.

Все методы, применяемые для получения эмпирического материала, можно условно разделить на активные и пассивные. К **активным методам** относятся лабораторный эксперимент и его различные модификации, квазиэксперимент, к **пассивным методам** — наблюдение, клинический метод, метод анализа продуктов деятельности, измерение и корреляционные исследования и т.д. Применяя методы первой группы, исследователь активно вызывает явление или процесс, воздействуя на объект. Используя же методы второй группы, он лишь довольствуется регистрацией естественного процесса.

Однако регистрировать поведение исследователь может непосредственно или применяя некоторый инструмент (опросник, тест и т.д.), так же как и воздействовать на испытуемого он может непосредственно в ходе беседы либо организуя деятельность обследуемого в лабораторном эксперименте с помощью исследовательских средств (приборов, заданий и т.д.). Это разделение психологических методов основано на критерии измерения “инструментальность— непосредственность”.

Наблюдение тоже может быть инструментальным, например, при использовании видеозаписи, магнитофонной записи и другой аппаратуры. Однако, в отличие от измерительного корреляционного исследования, при этом отсутствует взаимодействие испытуемого с инструментом. В эксперименте и в ходе измерения испытуемый активно выполняет задания, а при наблюдении никакой задачи ему не ставится, он ведет себя естественным образом.

Существует еще один важный критерий, который разделяет методы организации исследования на две большие группы. Это критерий соответствия метода признакам идеального исследования, как он понимается в современной методологии науки. По этому критерию различаются следующие методы организации исследования:

1) экспериментальное исследование, систематическое наблюдение или корреляционное исследование. Особенность их состоит в том, что исследователь пытается установить причинную или корреляционные связи между основными переменными, контролируя внешние переменные. Для этого он целенаправленно отбирает группы испытуемых или наблюдаемых индивидов, планирует определенным образом последовательность своих действий;

2) естественные эксперимент и наблюдение, беседа, клинический метод, метод описания частных случаев и др. Они применяются для выявления особенностей поведения человека, служат источником для эмпирических обобщений и выдвижения индуктивных гипотез, которые в дальнейшем могут стать материалом для теоретических рассуждений и проверяться в критических экспериментах. Способы контроля переменных (независимой, зависимой, внешних) систематически не применяются, хотя возможно пользование сложными техниками фиксации данных (картами наблюдения, аудио- и видеоаппаратурой, тестами и др.);

3) квазиэксперимент. “Промежуточный” между естественными методами проведения исследования и методами, где применяется строгий контроль переменных. Воздействие может быть выделено в природе как независимое от

исследователя, естественно происходящее. В этом случае мы получаем исследовательский метод, занимающий именно промежуточное положение между экспериментом и наблюдением. Под квазиэкспериментом принято понимать такой метод, при котором не удастся полностью реализовать схему, предписываемую идеальным исследованием, но эти отношения частично компенсируются использованием особых квазиэкспериментальных планов (см. юниту 6).

Согласно традиции, экспериментальное исследование противопоставляется всем неэкспериментальным методам.

Так, Ф.Дж. МакГиган противопоставляет экспериментальный метод следующим:

- 1) классическому клиническому методу;
- 2) естественному наблюдению;
- 3) опросу;
- 4) архивному исследованию;
- 5) установлению корреляционной связи;
- 6) квазиэксперименту.

М.В. Мэтлин приводит иной список методов, противопоставляемых эксперименту:

- 1) естественное наблюдение, к которому примыкает полевое исследование и полевой (естественный) эксперимент;
- 2) опрос;
- 3) корреляционное исследование;
- 4) архивное исследование;
- 5) изучение отдельных случаев.

П.К. Козби противопоставляет эксперименту корреляционный метод, а также доэкспериментальные методы: опрос, полевое наблюдение и полевой эксперимент. Отдельно он рассматривает методы измерения поведения, относя к ним архивный метод, самооценку (самоопрос) и тестирование.

Экспериментальное исследование в психологии отличается от других методов тем, что экспериментатор активно манипулирует независимой переменной, тогда как при прочих методах возможны лишь варианты отбора уровней независимых переменных. Нормальным вариантом экспериментального исследования является наличие основной и контрольных групп испытуемых. В неэкспериментальных исследованиях, как правило, все группы равноценны, поэтому проводится их сравнение.

По формальным основаниям выделяются несколько типов экспериментального исследования.

Различают исследовательский (поисковый) и подтверждающий эксперимент. Различие их обусловлено уровнем разработанности проблемы и наличием знаний о связи зависимой и независимой переменных.

Поисковый эксперимент проводится тогда, когда неизвестно, существует ли причинная связь между независимой и зависимой переменными. Поэтому поисковое исследование направлено на проверку гипотезы о наличии или отсутствии причинной зависимости между переменными А и В.

Если существует информация о качественной связи между двумя переменными, выдвигается гипотеза о виде этой связи. Тогда исследователь проводит **подтверждающий эксперимент**, в котором выявляется вид функциональной количественной связи между независимой и зависимой переменными.

Алгоритм исследования в целом выглядит так:

1. Выдвигается гипотеза о качественной причинной связи А и В.
2. Проводится поисковый эксперимент.
3. В случае неподтверждения гипотезы выдвигается другая качественная

гипотеза и проводится новый поисковый эксперимент; если же качественная гипотеза подтверждается, выдвигается количественная функциональная гипотеза.

4. Проводится подтверждающий эксперимент.

5. Принимается (или отвергается) и уточняется гипотеза о виде связи между переменными.

В психологической исследовательской практике для характеристики различных видов экспериментального исследования используются также понятия “пилотажное исследование”, или “пилотажный эксперимент”, “полевое исследование”, или “естественный эксперимент”.

Термин “**пилотажное исследование**” применяется для обозначения пробного, первого, эксперимента или серии экспериментов, в которых апробируются основная гипотеза, подходы к исследованию, план и т.д. Обычно пилотаж проводят перед “большим”, трудоемким экспериментальным исследованием, чтобы потом не тратить деньги и время попусту. Пилотажное исследование проводится на меньшей выборке испытуемых, по сокращенному плану и без строгого контроля внешних переменных. Надежность данных, получаемых в результате пилотажа, невелика, но его проведение позволяет устранить грубые ошибки, связанные с выдвиганием гипотезы, планированием исследования, контролем переменных и т.д. Кроме того, в ходе пилотажа можно сузить “зону поиска”, конкретизировать гипотезу и уточнить методику проведения “большого” исследования.

Полевое исследование проводится для изучения связи между реальными переменными в повседневной жизни, например, между статусом ребенка в группе и количеством его контактов в игре со сверстниками или занимаемой им территории в игровой комнате. По своей сути полевое исследование (или полевой эксперимент) относится к квазиэкспериментам, так как при его проведении нет возможности строго контролировать внешние переменные, отбирать группы и распределять внутри их испытуемых, управлять независимой переменной и точно регистрировать зависимую переменную. Но в некоторых случаях полевой, или естественный, эксперимент — единственно возможный способ получения научной информации (в психологии развития, социальной психологии, клинической психологии или психологии труда и т.д.). Сторонники естественного эксперимента утверждают, что лабораторный эксперимент является искусственной процедурой, дает невалидные результаты, поскольку “вырывает” испытуемого из контекста повседневной жизни. Но в полевых исследованиях ошибок, помех, влияющих на точность и надежность данных, неизмеримо больше, чем в лабораторном исследовании. Поэтому психологи стремятся планировать естественный эксперимент максимально близко к схеме проведения лабораторного эксперимента и перепроверять результаты, полученные “в поле”, более строгими процедурами.

Экспериментальное исследование проводится по определенной схеме: формулировка проблемы и выдвигание гипотезы, конструирование методики и подбор аппаратуры, отбор испытуемых, создание плана для контроля переменных, проведение эксперимента, обработка и интерпретация результатов, подготовка научного отчета. Некоторые из этапов являются обязательными, некоторые иногда могут отсутствовать, но последовательность шагов необходимо запомнить, чтобы не делать элементарных ошибок.

Рассмотрим основные **этапы психологического экспериментального исследования** и кратко опишем их содержание.

1. Любое исследование начинается с **определения** его **темы**. Тема ограничивает область исследований, круг проблем, выбор предмета, объекта и метода. Однако первым этапом собственно самого исследования является первичная **постановка проблемы**. Исследователь должен уяснить себе, чем он неудовлетворен в современном психологическом знании, где он ощущает пробелы, какие факты и закономерности не поддаются объяснению, какие теории

дают противоречащие друг другу объяснения поведения человека и т.д. Эмпирическое исследование проводится в трех основных случаях:

- проверка гипотезы о существовании явления;
- проверка гипотезы о существовании связи явлений;
- проверка гипотезы о причинной зависимости явления А от В.

Собственно эксперимент применяется только для обнаружения причинной связи явлений. Что является основной причиной повышенной невротизации (синдром школьной дезадаптации) детей 6-7 лет в первые месяцы учебы в школе? Возможные варианты ответов: низкий уровень готовности к учебной деятельности (несформированность первичных учебных навыков), недостаточное развитие произвольности, низкий уровень интеллекта, социально-психологические проблемы вхождения в классный коллектив и т.д. Каждая из этих гипотез требует проверки, хотя все они кажутся обоснованными и правдоподобными.

2. После первичной постановки проблемы наступает **этап работы с научной литературой**. Исследователь должен ознакомиться с экспериментальными данными, полученными другими психологами, и попытками объяснения причин заинтересовавшего его явления. К услугам современного исследователя компьютерные базы данных, сеть Internet, библиотеки, специализированные журналы.

Первичная работа начинается с поиска определений базовых понятий, которые содержатся в психологических словарях, а также в словарях и энциклопедиях по смежным дисциплинам. Там же имеются и ссылки на основные публикации по проблеме. Следующий шаг — составление библиографии по тематике исследования с помощью библиотечных систематических каталогов. Предварительное знакомство с публикациями на тему исследования можно осуществить посредством реферативных журналов. В нашей стране единственный журнал такого рода, издаваемый ВИНИТИ, — “0.4. Биология. Раздел 0.4.И. Психология”. Из зарубежных изданий наиболее авторитетным является “Psychological Abstract”, выпускаемый Американской психологической ассоциацией. В нем содержатся краткие аннотации большинства работ, публикуемых в англоязычных психологических журналах. Более подробную информацию об исследованиях, относящихся к выделенной проблеме, следует искать в самих публикациях: статьях научных журналов, сборниках и монографиях. Наиболее авторитетные психологические российские научные журналы — это “Психологический журнал”, издаваемый Российской академией наук; “Вопросы психологии” — орган Российской академии образования; “Вестник МГУ, серия Психология”, выпускаемый издательством МГУ; “Психологическое обозрение” — орган Российского психологического общества; “Иностранная психология”, издаваемый Институтом психологии РАН

Результат работы над литературным обзором — это уточнение проблемы, возникновение новой гипотезы и идеи плана экспериментального исследования. Возможно, что психолог и откажется от исследования, так как проблема может показаться неразрешимой или, наоборот, настолько исследованной, что ничего нового к имеющимся результатам добавить уже нельзя.

3. На следующем этапе происходит **уточнение гипотезы и определение переменных**. Первичная постановка проблемы уже скрыто предполагает варианты ответа на нее. Например, вопрос о том, что в большей мере — наследственность или среда — влияет на уровень развития общего интеллекта, ограничивает множество общих теоретических предположений. Аналогично, если поставить проблему — какой анализатор выполняет функцию интегратора сенсорной информации на ранних стадиях развития психики ребенка, — ответ будет ограничен списком анализаторных систем и “нулевым” ответом (“никакой”).

Экспериментальная гипотеза, в отличие от теоретической, должна быть сформулирована в виде высказывания: “Если ... то ...”. Кроме того, она должна

быть конкретизирована и операционализирована. Это означает, что входящие в высказывание “если А, то В” переменные А и В должны контролироваться в эксперименте: А — управляться экспериментатором, а В — регистрироваться непосредственно или с помощью аппаратуры. Определение переменных в терминах экспериментальной процедуры и их операционализация завершают этап уточнения гипотезы. Тем самым уточняется предмет экспериментального исследования: та сторона психики, на которую направлено экспериментальное воздействие и которая регулирует поведение, регистрируемое в ходе эксперимента. Психическая реальность всегда выступает в эксперименте “промежуточной переменной”.

Психолог управляет не психической реальностью, а внешними параметрами ситуации, воздействующими на психику испытуемого. Регистрируя независимую переменную, он исходит из того, что между промежуточной переменной и параметрами поведения существует функциональная (психорегулятивная) связь. Это основная общая гипотеза — предпосылка любого психологического эксперимента.

Помимо независимой, зависимой и промежуточной переменных должны быть определены и операционализированы внешние переменные, которые могут влиять на зависимую переменную.

4. Исследователь должен **выбрать экспериментальный инструмент**, который позволял бы ему: а) управлять независимой переменной; б) регистрировать зависимую переменную. Речь идет о конкретной методике и аппаратуре психологического эксперимента. Кроме того, условия эксперимента (помещение, ситуация, время и др.) должны сохранять константность величины воздействия внешних переменных на зависимую переменную.

Характер используемой аппаратуры определяется тем, какую методику выберет или сконструирует экспериментатор. Исследование феномена полезависимости—полнезависимости может проводиться с помощью различных конкретных методик: а) теста “Спрятанные фигуры”; б) методики “Стержень — рамка” или “Крест — рамка”; в) методики “Стержень — наклонная комната” и т.п. В первом случае исследователь использует в работе карандаш, бланк теста Виткина — Готшалльда и секундомер. Во втором случае лучше проводить автоматизированный эксперимент с помощью компьютера, предъявление стимульного материала реализуя на дисплее, а действия испытуемого регистрируя посредством джойстика. Наконец, третий вариант требует специального оборудования экспериментальной камеры.

В психологическом эксперименте может применяться самая разнообразная аппаратура, в том числе психофизиологическая. Следует лишь избегать перегрузки испытуемого избыточными для исследования тестами.

К сожалению, в России не налажен выпуск аппаратуры для проведения психологических экспериментальных исследований. Нет стандарта на оборудование экспериментальных лабораторий. Выпуск тестовых методик также не удовлетворяет потребности исследователей и практиков. Поэтому основная аппаратура либо изготавливается самостоятельно, кустарным способом, либо используется медицинское оборудование и аппаратура для биофизических и психофизиологических исследований.

5. **Планирование экспериментального исследования** является центральным этапом всей процедуры. В первую очередь речь идет о выделении внешних переменных, которые могут влиять на зависимую переменную. Планирование необходимо для обеспечения внешней и внутренней валидности эксперимента.

Следующим шагом является выбор экспериментального плана. Какой план предпочтительнее? Ответ на этот вопрос зависит от того, какова экспериментальная гипотеза, какое число внешних переменных вы должны контролировать в эксперименте, какие возможности предоставляет ситуация для

проведения исследований и т.д. При ограниченности времени и ресурсов (в том числе финансовых) выбирают максимально простые экспериментальные планы. Для проверки сложных гипотез, требующих управления несколькими независимыми переменными и\или учета многих дополнительных переменных, используют соответствующие планы.

Исследователь может проводить эксперимент при участии одного испытуемого. В этом случае он применяет какой-либо из планов исследования для одного испытуемого. Если исследователь работает с группой, то он может выбрать ряд планов с использованием экспериментальной и контрольных групп. Простейшими являются планы для двух групп (основная и контрольная). Если необходим более усложненный контроль, применяются планы для нескольких групп.

Другой вариант, часто используемый в психологии, — факторные планы. Они используются, если требуется выявить влияние двух и более независимых переменных на одну зависимую. При этом независимые переменные могут иметь несколько уровней интенсивности. Простейшие факторные планы типа 2x2 или 2x2x2 предполагают использование двух и соответственно трех независимых переменных с двумя уровнями градации. Существуют и более сложные экспериментальные планы.

Процессуальная классификация планов для исследования связи двух переменных создана Д. Кэмпбеллом. Основными являются: простой план для двух групп с предварительным тестированием (тест — воздействие — ретест); план для двух рандомизированных групп без предварительного тестирования (рандомизация — воздействие — тест); план Соломона для четырех групп, объединяющий оба этих плана. Они называются планами истинных экспериментов.

В случае, если план истинного эксперимента реализовать невозможно или не нужно, исследователь применяет один из квазиэкспериментальных планов (см. юниту 6).

6. Отбор и распределение испытуемых по группам проводятся в соответствии с принятым экспериментальным планом. Всю совокупность потенциальных испытуемых, которые могут быть объектами данного психологического исследования, обозначают как популяцию, или генеральную совокупность. Множество людей, принимающих участие в исследовании, называют выборкой. Состав экспериментальной выборки должен моделировать (представлять) генеральную совокупность, поскольку выводы, получаемые в эксперименте, распространяются на всех членов популяции, а не только на представителей этой выборки.

Выбор популяции зависит от целей исследования. Специалист по психогенетике выбирает в качестве испытуемых моно- и дизиготных близнецов, а также их родителей, дедушек и бабушек. Медицинский психолог при изучении агрессивного поведения у лиц с синдромом посттравматического стресса исследует выборку ветеранов афганской войны.

Все потенциальные испытуемые характеризуются разным полом, возрастом, социальным положением, уровнем образования, состоянием здоровья и т.д. Кроме того, они обладают различными индивидуально-психологическими особенностями, например разными уровнями интеллекта, нейротизма, агрессивности. Для того чтобы выборка представляла генеральную совокупность, потенциальным испытуемым должны быть предоставлены равные шансы стать реальными участниками исследования.

Техника рандомизации состоит в том, что всем представителям совокупности присваивается индекс, а затем производится случайный отбор в группу необходимой численности для участия в эксперименте. В этом случае мы имеем три группы: 1) всю генеральную совокупность; 2) группу рандомизации, из которой производится отбор; 3) экспериментальную рандомизированную выборку.

Одно из требований к выборке — репрезентативность. Выборка должна качественно и количественно представлять генеральную совокупность, основные типы потенциальных испытуемых, существующие в популяции. Испытуемые должны быть правильно распределены по экспериментальной и контрольным группам, чтобы все группы были эквивалентными.

Кроме того, исследователь распределяет группы относительно разных условий эксперимента так, чтобы контролировать или учитывать возможные эффекты последовательности, дифференцированного переноса и др.

7. Проведение эксперимента является, очевидно, наиболее ответственной частью исследования, требующей от человека не только знаний и навыков, но и способностей к экспериментированию. Любой самый лучший замысел можно испортить небрежным проведением эксперимента.

Подготовка эксперимента. Исследователь готовит экспериментальное помещение и оборудование. Если это необходимо, проводится несколько пробных опытов для отладки процедуры эксперимента. Важнейшим моментом является разработка и уточнение инструкции. Она должна состоять из кратких предложений, каждое из которых включает не более 11 слов. В инструкции с помощью абзацев выделяются смысловые блоки. Ее проверяют на понятность и простоту, проводя предварительный опыт на 5—10 испытуемых.

Инструктирование и мотивирование испытуемых. Инструкция должна включать в себя мотивационные компоненты. Испытуемому необходимо знать, какие возможности предоставляет ему участие в эксперименте. Это может быть денежная оплата (характерно для американской и бывшей советской психологии), информация о своих способностях и личностных чертах, помощь в решении личных проблем и т.д. Поскольку ситуация эксперимента для большинства испытуемых непривычна, они испытывают тревогу. Кроме того, скорость понимания инструкции зависит от индивидуальных когнитивных способностей, особенностей темперамента, знания языка и т.д. Поэтому следует проверить, правильно ли испытуемые поняли инструкцию, и повторить ее при необходимости, избегая, однако, дополнительных развернутых комментариев.

Экспериментирование. Вначале следует убедиться в дееспособности испытуемого, в том, что он здоров и желает участвовать в эксперименте. В ходе эксперимента исследователь организует процесс взаимодействия с испытуемым, зачитывает инструкцию, проводит, если это необходимо, обучающую серию. Он варьирует независимую переменную (задачи, внешние условия и др.), проводит сам или с помощью ассистента регистрацию поведения испытуемого. Перед экспериментатором должна лежать инструкция, в которой зафиксирован порядок его действий в ходе исследования.

Обычно в эксперименте принимает участие и ассистент. Он берет на себя вспомогательные задачи. Чаще всего именно ассистент ведет протокол, в котором фиксируются ответы испытуемого. Кроме того, ассистент ведет общее наблюдение за поведением испытуемого и его состоянием, а также за всеми отклонениями от стандартной процедуры эксперимента. Он же следит за работой аппаратуры. Эксперимент в зависимости от целей исследования может быть частично или полностью автоматизированным. Проблеме автоматизации психологического исследования и применения компьютеров в эксперименте на человеке посвящено множество монографий и сборников научных работ. Протоколировать ответы испытуемого лучше сразу посредством ввода данных в электронную таблицу, которая создается в компьютере (пакет EXEL).

Рекомендуется регистрировать дополнительные признаки поведения испытуемого, его эмоциональные реакции по ходу эксперимента. Необходимым завершающим этапом является постэкспериментальное интервью. По завершении эксперимента следует провести беседу с испытуемым и поблагодарить его за участие в исследовании.

8. Выбор методов статистической обработки и ее проведение — следующий этап исследования.

Обычно методы обработки данных выбираются на стадии планирования эксперимента или же еще раньше — при выдвижении экспериментальной гипотезы. Экспериментальная гипотеза преобразуется в статистическую. Возможных типов статистических гипотез в экспериментальном исследовании немного: а) о сходстве или различии двух и более групп; б) о взаимодействии независимых переменных; в) о статистической связи независимых и зависимых переменных; г) о структуре латентных переменных (относится к корреляционному исследованию).

Статистические оценки дают информацию не о наличии, а о достоверности сходств и различий результатов контрольных и экспериментальных групп.

Существуют “привязки” определенных методов обработки результатов к экспериментальным планам. Для оценки различий данных, полученных при применении плана для двух групп, используют критерии t , X^2 и F . Факторные планы требуют применения дисперсионного анализа для оценки влияния независимых переменных на зависимую, а также определения меры их взаимодействия друг с другом.

Существуют стандартные пакеты программ для математической обработки данных. Наиболее известными и доступными являются “Statistica”, “Stadia”, “Statgraphics”. Все пакеты делятся на виды: специализированные пакеты и пакеты общего назначения. Для исследователей рекомендуются пакеты общего назначения. Западные статистические пакеты требуют хорошей подготовки пользователя на уровне знания университетского курса математической статистики и многомерного анализа данных. Каждая программа снабжается документацией. Отечественные пакеты более приближены к возможностям нашего пользователя. Сопутствующая информация (справочник, интерпретатор выводов и др.) включается в программную систему. Примерами являются отечественные статистические пакеты “Stadia”, “Мезозавр”, “Эвриста”.

9. Выводы и интерпретация результатов завершают исследовательский цикл. Итогом экспериментального исследования является подтверждение или опровержение гипотезы о причинной зависимости между переменными: “Если А, то В”.

Подтверждение статистических гипотез (о различиях, связи и пр.) — решающий аргумент в пользу принятия экспериментальной гипотезы. Исследователь сопоставляет свои выводы с выводами других авторов, высказывает гипотезы о причинах сходства или различия между собственными данными и результатами предшественников. И наконец, он интерпретирует свои выводы в терминах теоретической гипотезы. Он должен ответить на вопрос: можно ли считать подтверждение или опровержение эмпирической гипотезы подтверждением или же опровержением той или иной теории. Возможно, что ни одна теория не может объяснить полученные в эксперименте результаты. Тогда экспериментатор, если он склонен к теоретизированию, пытается сам теоретически объяснить результаты. Кроме того, он высказывает предположения о возможности обобщения и переноса своих данных на другие ситуации, популяции и т.д.

10. Конечным продуктом исследования являются научный отчет, рукопись статьи, монография. Существуют определенные требования к оформлению рукописной научной работы, наглядному представлению результатов и структуре изложения (см. юниту 4).

Исследование считается завершенным, если экспериментальная гипотеза опровергнута или не опровергнута с заданной надежностью, а результаты исследования в соответствующей форме представлены на суд научной общественности.

1.8. Экспериментальная выборка

Выбор объекта исследования — следующая задача, которую предстоит решать психологу. В прикладных исследованиях свобода психолога ограничена тем, что объект известен с самого начала. В других случаях психолог волен выбирать тот объект, который больше соответствует задачам исследования.

Идеальным объектом психологического исследования может выступать либо отдельный индивид, либо группа. В первом случае мы говорим об общепсихологическом эксперименте, во втором — о социально-психологическом. Но в конкретном эксперименте не только реальный объект должен соответствовать по своим характеристикам идеальному объекту, но и результаты, полученные при его участии, должны применяться к другим объектам. Если бы все люди были похожи друг на друга, а еще лучше — оказались совершенно одинаковыми, то не возникло бы никаких проблем. Эксперимент можно провести с участием одного испытуемого, а полученные результаты применить для объяснения поведения всех других людей. Но люди различаются по полу, возрасту, расе, национальности, принадлежности к той или иной культуре или религии, социальному и экономическому положению и т.д. Следовательно, простая генерализация (обобщение) данных, полученных при исследовании одного испытуемого, невозможна. Однако на практике, например в лабораторных экспериментах по исследованию сенсорных процессов, памяти, внимания и т.д., этими различиями пренебрегают, считая, что наш испытуемый может представлять любого индивида.

Для экспериментов с одним испытуемым разработаны особые экспериментальные планы. Эксперимент с одним испытуемым проводится тогда, когда: 1) индивидуальными различиями можно пренебречь, исследование чрезвычайно велико по объему и включает множество экспериментальных проб; 2) испытуемый — уникальный объект, например гениальный музыкант или творчески одаренный шахматист; 3) от испытуемого требуется особая компетентность при проведении исследования (эксперимент с обученными испытуемыми); 4) повторение данного эксперимента с участием других испытуемых невозможно.

Чаще исследование проводится с экспериментальной группой, в которой все испытуемые объективно различны, но отобраны и распределены по подгруппам с помощью той или иной стратегии. Различают четыре основных вида конструирования экспериментальных групп.

При **первом варианте** исследование проводится с двумя различными группами: экспериментальной и контрольной, которые ставятся в разные условия. Это наиболее распространенный способ.

Второй вариант предполагает исследование одной группы: ее поведение изучается и в экспериментальных и в контрольных условиях. Он применяется, когда имеется только экспериментальная группа и нет возможности сформировать контрольную. Но этот план никак не контролирует “эффект последовательности” и используется лишь в тех редких случаях, когда эффектом последовательности можно пренебречь.

Третий вариант — использование конструирования групп методом “парного дизайна” — состоит в следующем. Для каждого субъекта группы подбирается эквивалентный (или похожий) ему, и они распределяются по разным группам. Соответственно контрольная и экспериментальная группы становятся похожими по составу испытуемых. Конечно, в этом случае невозможно соблюсти полную эквивалентность групп в обоих условиях эксперимента, но данный способ значительно лучше, чем эксперимент с участием одной группы в разных условиях.

Наконец, **четвертый план** является смешанным: все группы ставятся в разные условия. При этом образуются несколько групп. Этот способ применяется при

факторном планировании эксперимента.

Формирование выборки испытуемых — экспериментальной группы — должно подчиняться ряду правил.

1. **Содержательный критерий** (критерий операциональной валидности). Напомним, что операциональная валидность определяется соответствием экспериментального метода проверяемой гипотезе.

Подбор экспериментальной группы должен определяться предметом и гипотезой исследования. Бесмысленно проверять степень развития произвольного запоминания у годовалых и двухлетних детей или выявлять у группы московских бомжей, в какой мере уровень их интеллекта влияет на выбор того или иного кандидата в Государственную Думу (ибо они голосовать не имеют права).

Таким образом, экспериментатор должен создать модель идеального объекта экспериментального исследования для своего частного случая, по возможности его описать и следовать этому описанию при формировании экспериментальной группы. Характеристики реальной экспериментальной группы должны минимально отклоняться от характеристик идеальной экспериментальной группы.

2. **Критерий эквивалентности испытуемых** (критерий внутренней валидности). Процедура подбора эквивалентных групп и эквивалентных испытуемых называется **рандомизацией**.

Результаты, полученные при исследовании экспериментальной выборки, должны распространяться на каждого ее члена. То есть необходимо учесть все значимые характеристики объекта исследования, различия в выраженности которых могут существенно повлиять на зависимую переменную. Допустим, надо проверить влияние ситуативной тревожности детей на скорость овладения школьными навыками. В этом случае состав экспериментальной группы должен быть подобран так, чтобы в нее входили дети с одинаковым уровнем развития интеллекта. Если же это не удастся сделать, то при обработке данных используется нормировка результатов на величину значимого параметра.

3. **Критерий репрезентативности** (критерий внешней валидности). Существуют теоретические статистические критерии репрезентативности (представленности) выборки испытуемых. Группа лиц, участвующих в эксперименте, должна представлять всю часть популяции, по отношению к которой можно применять данные, полученные в эксперименте. Величина экспериментальной выборки определяется видом статистических мер и выбранной точностью (достоверностью) принятия или отвержения экспериментальной гипотезы. Она может быть равна множеству индивидов, чье поведение нас интересует. Экспериментальная выборка может представлять лишь часть интересующего нас множества. Главная проблема состоит в том, на какие другие интересующие нас группы можно распространить результаты проводимого исследования.

Подбор экспериментальной группы осуществляется с помощью различных стратегий. Напомним, для чего нужна **стратегия отбора групп**. Задача сводится, во-первых, к устранению «эффекта смешения». Под этим термином понимается влияние индивидуальных различий между испытуемыми на связь независимой и зависимой переменных. Например, на публичное поощрение за хорошую работу люди разного темперамента реагируют неодинаково. Тем самым контролируется влияние побочной переменной на внутреннюю валидность. Во-вторых, экспериментальная группа должна представлять изучаемую популяцию, т.е. обеспечивать внешнюю валидность эксперимента.

Привлечение реальных групп порождает систематическое смешение независимой переменной с индивидуальными свойствами испытуемых.

Для нас уже очевидно, что экспериментальная выборка представляет собой модель популяции в целом или той ее части, поведение которой нас интересует.

Наиболее простой вариант был уже рассмотрен — **составление репрезентативной группы** испытуемых, характеристики которой соответствуют характеристикам интересующей нас популяции.

Иногда невозможно найти способ создания репрезентативной группы. Тогда используется **метод приближенного моделирования**. В частности, исследование может проводиться при участии студентов 2-го курса университета, а данные приписываются всем людям, или “людям в возрасте от 17 до 21 года”, или “людям со средним образованием в возрасте от 17 до 21 года” и т.д. Чем меньше генерализация, т.е. чем точнее набор критериев, описывающих популяцию, на которую распространяются выводы о характеристиках экспериментальной выборки, тем выше внешняя валидность эксперимента.

При моделировании популяции **методом случайного выбора**, или рандомизации, экспериментальную выборку составляют так, что каждой личности предоставляется равный шанс для участия в эксперименте. Каждому индивиду присваивается номер; с помощью таблицы случайных чисел производится формирование экспериментальной выборки. Процедура очень трудно осуществима, поскольку каждый представитель интересующей нас популяции должен быть учтен. На практике прибегают к более простым способам случайного отбора. Отбирают любую группу испытуемых, затем измеряют у них значимое для эксперимента индивидуальное свойство. После этого испытуемых распределяют по группам методом Монте-Карло так, что вероятность попасть в группу для каждого испытуемого одинакова.

Наконец, существует **стратометрический способ моделирования выборки**. Генеральная совокупность рассматривается как совокупность групп, обладающих определенными характеристиками. В экспериментальную выборку отбираются испытуемые с соответствующими характеристиками — так, чтобы в ней были равнопредставлены лица из каждой страты. Чаще всего используются следующие характеристики: пол, возраст, политические предпочтения, образование и уровень доходов. Эту стратегию применяют психодиагносты при разработке тестов; педагогические психологи; в основном же ею пользуются социологи и социальные психологи при опросах общественного мнения, исследовании социальных установок и т.д.

Ряд авторов выделяют **стратегию попарного отбора**. При этом экспериментальная и контрольная группы составляются из индивидов, эквивалентных по значимым для эксперимента побочным параметрам. Идеальный вариант — использование близнецовых пар (моно- и дизиготных). Разновидностью этой стратегии является подбор однородных подгрупп, в которых испытуемые уравниваются по всем характеристикам, кроме интересующих исследователя дополнительных переменных. Другой вариант — выделение значимой дополнительной переменной. Все испытуемые тестируются, ранжируются по уровню выраженности переменной. При формировании групп необходимо учитывать, чтобы испытуемые, обладающие одинаковыми или близкими значениями переменной, попали в разные группы.

Итак, существуют **шесть стратегий построения групп**: 1) рандомизация, 2) попарный отбор, 3) рандомизация с выделением страт (стратометрический отбор), 4) приближенное моделирование, 5) репрезентативное моделирование, 6) привлечение реальных групп.

Различают **два основных типа привлечения испытуемых в группу**: отбор и распределение. Отбор проводят при рандомизации, рандомизации с выделением страт, при репрезентативном и приближенном моделировании. Распределение осуществляется при способе составления групп из эквивалентных пар и исследованиях с участием реальных групп.

Считается, что наилучшая внешняя и внутренняя валидность достигается при стратегии подбора эквивалентных пар и стратометрической рандомизации:

индивидуальные особенности испытуемых с помощью этих стратегий контролируются максимально. В остальных же случаях нет никаких гарантий эквивалентности испытуемых, контролируемости индивидуальных различий и представительности группы.

Привлечение добровольцев или принудительное участие в эксперименте нарушает репрезентативность выборки. Отметим лишь, что стратегии попарного моделирования, приближенного моделирования и стратометрической рандомизации, в отличие от стратегии рандомизации (“случайно отобранных групп”), предполагают, что нам известен дополнительный параметр — индивидуальная особенность, которая может оказать значимое влияние на результат эксперимента. Ошибка в выделении этого параметра и/или недоучет других параметров приводят экспериментатора к неудаче. Очевидно, что применение таких стратегий формирования групп должно опираться на серьезные знания в области дифференциальной психологии (психологии индивидуальных различий).

Многие авторы, в частности Кэмпбелл, считали подбор эквивалентных пар менее надежным методом, чем рандомизация. Рандомизацию можно считать наиболее надежной стратегией и с точки зрения репрезентации в эксперименте изучаемой популяции, и с точки зрения контроля дополнительных переменных. Наиболее существенная проблема при рандомизации: в какой мере первичная выборка, из которой формируются экспериментальная и контрольные группы, действительно представляет генеральную совокупность? Решение этой проблемы — дело искусства и интуиции экспериментатора.

Отдельная проблема — численность экспериментальной выборки. В зависимости от целей и возможностей она может варьироваться от одного испытуемого до нескольких тысяч человек. Количество испытуемых в отдельной группе (экспериментальной или контрольной) в большинстве экспериментальных исследований варьируется от 1 до 100. Рекомендуется, чтобы численность сравниваемых групп была не менее 30—35 человек из соображений статистических: коэффициенты корреляции выше 0,35 при таком количестве испытуемых значимы при $\alpha = 0,05$.

Если же для обработки данных используется факторный анализ, то существует простое правило: надежные факторные решения можно получить лишь в том случае, когда количество испытуемых не менее чем в 3 раза превышает число регистрируемых параметров. Кроме того, целесообразно увеличивать количество испытуемых по крайней мере на 5—10% больше требуемого, поскольку часть из них будет “отбракована” в ходе эксперимента или при анализе экспериментальных протоколов (не поняли инструкцию, не приняли задачу, дали “девиантные” результаты и т.д.).

Что касается состава по полу и возрасту, то рекомендуется (кроме специальных случаев) разбивать общую группу на подгруппы мужчин и женщин и обрабатывать данные отдельно для каждой подгруппы. Возрастной состав определяется исходя из целей исследования. Для ориентировки приведем наиболее распространенную периодизацию возрастов:

- 1) новорожденный — 0-10 дней;
- 2) грудной — 10 дней-1 год;
- 3) раннее детство — 1-2 года;
- 4) первый период детства — 3-7 лет;
- 5) второй период детства — 8-12 лет (мальчики), 8-11 лет (девочки);
- 6) подростковый — 13-16 лет (мальчики), 12-15 лет (девочки);
- 7) юношеский — 17-21 год (мужчины), 16-20 лет (женщины);
- 8) средний:
первый период — 22-35 лет (мужчины), 21-35 лет (женщины);
второй период — 36-60 лет (мужчины), 36-55 лет (женщины);
- 9) пожилой возраст — 61-75 (мужчины), 55-75 лет (женщины);

- 10) старческий возраст — 75-90 лет;
- 11) долгожители — 90 лет и старше.

1.9. Экспериментальные переменные

Суть эксперимента состоит в том, что экспериментатор варьирует независимую переменную, регистрирует изменение зависимой переменной и контролирует внешние переменные.

Экспериментатор проверяет гипотезу о причинной связи двух явлений, А и В. Понятие “причинность” является одним из наиболее сложных в науке. Существует ряд эмпирических признаков причинной связи между двумя явлениями.

Первый признак — разделенность причины и следствия во времени и предшествование причины следствию. Если исследователь обнаруживает изменения в объекте после экспериментального воздействия, по сравнению с аналогичным объектом, который таковому не подвергался, у него есть повод говорить о том, что экспериментальное воздействие стало причиной изменения состояния объекта. Наличие воздействия и сравнение объектов являются необходимыми условиями такого вывода, ибо не всегда предшествующее событие — причина последующего. Отлет гусей на юг отнюдь не является причиной выпадения снега через месяц.

Второй признак — наличие статистической связи между двумя переменными (причиной и следствием). Изменение величины одной из переменных должно сопровождаться изменением другой. Иначе говоря, между переменными должна наблюдаться либо линейная корреляция, как между уровнем вербального интеллекта и школьной успеваемостью, либо нелинейная корреляция, как между уровнем активации и степенью эффективности научения (закон Йеркса—Додсона).

Наличие корреляции — недостаточное условие для вывода о причинно-следственной связи, так как связь может быть случайной или обусловленной третьей переменной.

Третий признак — причинно-следственная связь регистрируется, если экспериментальная процедура исключает иные возможности объяснения связей А и В, кроме как причинной, и все другие альтернативные причины возникновения явления В исключены.

Проверка экспериментальной гипотезы о причинной связи двух явлений производится следующим образом. Экспериментатор моделирует предполагаемую причину: она выступает в качестве экспериментального воздействия, а следствие — изменение состояния объекта — регистрируется с помощью какого-либо измерительного инструмента.

Экспериментальное воздействие служит для изменения независимой переменной, которая является непосредственной причиной изменения зависимой переменной. Так, экспериментатор, предъявляя испытуемому сигналы различной околороговой громкости, изменяет его психическое состояние — испытуемый либо слышит, либо не слышит сигнал, что приводит к различным моторным или вербальным ответам (“да” — “нет”, “слышу” — “не слышу”).

Какие независимые, зависимые и внешние переменные встречаются при проведении психологического эксперимента?

1.9.1. Независимая переменная

Независимая переменная — управляемое экспериментальное воздействие (фактор), т.е. активно изменяемая и контролируемая исследователем переменная, представленная на двух или более уровнях (качественных или количественных).

Функциональный контроль независимой переменной — основное условие для заключения о влиянии управляемых экспериментатором воздействий на измеряемые зависимые переменные.

Исследователь должен стремиться оперировать в эксперименте только независимой переменной. Эксперимент, где это условие соблюдается, называют чистым экспериментом. Но чаще всего в ходе эксперимента, варьируя одну переменную, экспериментатор изменяет вместе с тем ряд других. Это изменение может быть вызвано действием экспериментатора и обусловлено связью двух переменных. Например, в эксперименте по выработке простого двигательного навыка он наказывает испытуемого за неудачи электрическим током. Размер наказания может выступать в качестве независимой переменной, а скорость выработки навыка — в качестве зависимой переменной. Наказание не только закрепляет у испытуемого соответствующие реакции, но и порождает у него ситуативную тревогу, которая влияет на результаты — увеличивает число ошибок и уменьшает скорость выработки навыка.

Центральная проблема при проведении экспериментального исследования — выделение независимой переменной и ее изоляция от других переменных. Исследователи различают разные виды независимой переменной: качественную (“есть подсказка” — “нет подсказки”), количественную (уровень денежного вознаграждения).

В качестве независимых переменных в психологическом эксперименте могут выступать:

1) **характеристики задания** — то, чем может манипулировать экспериментатор более или менее свободно. По традиции, идущей от бихевиоризма, считается, что экспериментатор варьирует только характеристиками стимулов, но в его распоряжении гораздо больше возможностей. Экспериментатор может варьировать стимулы или материал задания, изменять тип ответа испытуемого (вербальный или невербальный ответ), менять шкалу оценивания и т.д. Он может варьировать инструкцию, меняя цели, которые должен достичь испытуемый в ходе выполнения задания. Экспериментатор может варьировать средства, которые имеет испытуемый для решения задачи, и ставить перед ним препятствия, изменять систему поощрений и наказаний в ходе выполнения задания и т.д.;

2) **особенности ситуации** (внешние условия). К ним следует отнести те переменные, которые непосредственно не входят в структуру экспериментального задания, выполняемого испытуемым. Это может быть температура в помещении, обстановка, наличие внешнего наблюдателя и т.д.

Чем может варьировать экспериментатор? Во-первых, это физические параметры ситуации: расположение аппаратуры, внешний вид помещения, освещенность, звуки и шумы, температура, размещение мебели, окраска стен, время проведения эксперимента (время суток, длительность и т.д.), т.е. все физические параметры ситуации, не являющиеся стимулами. Во-вторых, это социально-психологические параметры: изоляция — работа в присутствии экспериментатора; работа в одиночку — работа с группой и т.д. В-третьих, это особенности общения и взаимодействия испытуемого (испытуемых) и экспериментатора.

Эксперименты по выявлению эффекта социальной фасилитации (усиления) проводились по следующей схеме: испытуемому давалась какая-либо сенсомоторная или интеллектуальная задача. Он сначала выполнял ее в одиночку, а затем в присутствии другого человека или нескольких людей (последовательность, разумеется, менялась в разных группах). Оценивалось изменение продуктивности испытуемых. В этом случае задача испытуемого оставалась неизменной, изменялись лишь внешние условия эксперимента.

Судя по публикациям в научных журналах, за последние годы резко возросло количество экспериментальных исследований, в которых применяется варьирование внешних условий. **Дополнительное варьирование** — способ

повышения уровня обобщения выводов из экспериментального исследования, который предполагает систематическое варьирование дополнительных условий, важных для распространения каузальной зависимости на множество других ситуаций;

3) **управляемые особенности** (состояния) испытуемого. К ним относятся психические процессы, состояния, образования и свойства.

Иногда выделяют четвертый вид переменных — **константные характеристики** испытуемого (интеллект, пол, возраст и т.д.), но они относятся к **дополнительным переменным**, поскольку на них нельзя воздействовать, а можно лишь учесть их уровень при формировании экспериментальных и контрольных групп.

Как правило, в современном экспериментальном исследовании дифференциально-психологические особенности индивидов, такие как интеллект, пол, возраст, социальное положение (статус) и т.д., учитываются в качестве дополнительных переменных, которые контролируются экспериментатором в общепсихологическом эксперименте.

Влияние дифференциально-психологических, демографических и прочих константных параметров на поведение индивида изучают в **корреляционных исследованиях**. Но эти переменные могут превращаться во “вторую основную переменную” в дифференциально-психологическом исследовании, и тогда используется факторный план.

1.9.2. Зависимая переменная

Зависимая переменная — измеряемая в эксперименте переменная, изменения которой причинно обусловлены действием независимой переменной.

Психологи имеют дело с поведением испытуемого, поэтому в качестве зависимой переменной выбираются параметры вербального и невербального поведения. К ним относятся: число ошибок, которое совершил оператор; время, которое затратил испытуемый при решении задачи; изменения мимики его лица при просмотре эротического фильма; время двигательной реакции на звуковой сигнал и т.д.

Среди зависимых переменных выделяются базисные. **Базисная переменная** — единственная зависимая переменная, на которую оказывает влияние независимая переменная.

Выбор поведенческого параметра определяется исходной экспериментальной гипотезой. Исследователь должен ее максимально конкретизировать, т.е. добиться того, чтобы зависимая переменная была операционализирована — поддавалась регистрации в ходе эксперимента.

Параметры поведения условно можно разделить на формально-динамические и содержательные. **Формально-динамические (или пространственно-временные) параметры** достаточно легко поддаются аппаратурной регистрации. Приведем примеры этих параметров.

1. **Точность**. Это наиболее часто регистрируемый параметр. Поскольку большинство заданий, предъявляемых испытуемому в психологических экспериментах, являются задачами на достижения, то точность или ошибочность действий будет главным регистрируемым параметром поведения.

2. **Латентность**. Психические процессы протекают скрытно от внешнего наблюдателя. Время от момента предъявления сигнала до выбора ответа называется латентным временем. В некоторых случаях латентное время является важнейшей характеристикой процесса, например, при решении мыслительных задач.

3. **Длительность, или скорость, выполнения** является характеристикой исполнительного действия. Время между выбором действия и окончанием его

выполнения называют скоростью действия (в отличие от латентного времени).

4. **Темп, или частота, действий.** Это важнейшая характеристика, особенно при исследовании простейших форм поведения.

5. **Продуктивность** — отношение числа ошибок или качества выполнения действий ко времени выполнения. Служит важнейшей характеристикой при исследовании научения, познавательных процессов, процессов принятия решения и т.д.

Содержательные параметры поведения предполагают категоризацию формы поведения либо в терминах обыденного языка, либо в терминах той теории, предположения которой проверяются в данном эксперименте.

Распознавание различных форм поведения — дело специально обученных экспертов или наблюдателей. Требуется немалый опыт, чтобы безошибочно различать разные уровни агрессии или удивления, характеризовать один поступок как проявление покорности, а другой — как проявление подобострастия.

Проблема фиксации качественных особенностей поведения решается посредством: а) обучения наблюдателей и разработки карт наблюдения; б) измерения формально-динамических характеристик поведения с помощью тестов.

Зависимая переменная должна быть валидной и надежной. Надежность переменной проявляется в устойчивости ее регистрируемости при изменении условий эксперимента в течение времени. Валидность зависимой переменной определена только в конкретных условиях эксперимента и применительно к определенной гипотезе.

Можно выделить три типа зависимых переменных: 1) одномерную; 2) многомерную; 3) фундаментальную. В первом случае регистрируется лишь один параметр, и именно он считается проявлением зависимой переменной (между ними существует функциональная линейная связь), как, например, при изучении времени простой сенсомоторной реакции. Во втором случае зависимая переменная многомерна. Например, уровень интеллектуальной продуктивности проявляется во времени решения задачи, его качестве, трудности решенной задачи. Эти параметры могут фиксироваться независимо. В третьем случае, когда известно отношение между отдельными параметрами многомерной зависимой переменной, параметры рассматриваются в качестве аргументов, а сама зависимая переменная — в качестве функции.

Существует еще одно важное свойство зависимой переменной, а именно — сензитивность (чувствительность) зависимой переменной к изменениям независимой. Суть в том, что манипуляция независимой переменной влияет на изменение зависимой. Если же мы манипулируем независимой переменной, а зависимая не изменяется, то зависимая переменная несензитивна по отношению к независимой. Два варианта проявления несензитивности зависимой переменной получили названия “эффект потолка” и “эффект пола”. Первый случай встречается тогда, когда предъявляемая задача так проста, что уровень ее выполнения много выше всех уровней независимой переменной. Второй эффект, напротив, возникает тогда, когда задание настолько сложно, что уровень его выполнения оказывается ниже всех уровней независимой переменной.

Итак, как и прочие компоненты психологического исследования, зависимая переменная должна быть валидна, надежна, обладать чувствительностью к изменению уровня независимой переменной.

Существуют два основных приема фиксации изменений зависимой переменной. Первый прием применяется наиболее часто в экспериментах с участием одного испытуемого. Изменение зависимой переменной регистрируется во время эксперимента вслед за изменением уровня независимой переменной. Примером является фиксация результатов в экспериментах по научению. Кривая научения представляет собой классический вариант тренда — изменения успешности выполнения заданий от числа проб (времени проведения экспери-

мента). Для обработки таких данных применяется статистический аппарат анализа трендов.

Второй прием фиксации изменения уровня независимой переменной называется отсроченным измерением. Между воздействием и эффектом проходит определенный промежуток времени, его длительность устанавливается по времени отдаленности следствия от причины. Например, прием дозы алкоголя увеличивает время сенсомоторной реакции не сразу, а по прошествии определенного времени. То же самое можно сказать о влиянии заучивания конкретного количества иностранных слов на успешность перевода текста на редкий язык: эффект проявляется не сразу (если проявляется).

1.9.3. Отношения между переменными

В основе построения современной экспериментальной психологии лежит формула — поведение (В) есть функция (f) личности (Р) и ситуации (S):

$$B = f(P;S).$$

Большинство специалистов по теории психологического эксперимента считают, что в психологии существуют два типа законов: 1) “стимул — ответ”; 2) “организм — поведение”.

Первый тип законов обнаруживается в ходе экспериментального исследования, когда стимул (задача, ситуация) — это независимая переменная, а зависимая переменная — ответ испытуемого.

Второй тип законов является продуктом метода систематического наблюдения и измерения, поскольку свойствами организма управлять с помощью психологических средств нельзя.

Существуют ли “пересечения”? Разумеется. Ведь в психологическом эксперименте зачастую учитывается влияние так называемых дополнительных переменных, большинство из которых является дифференциально-психологическими характеристиками. Следовательно, есть смысл добавить в список и **системные законы**, описывающие влияние ситуации на поведение личности, обладающей определенными свойствами. В психофизиологических и психофармакологических экспериментах можно воздействовать на состояние организма, а в ходе формирующего эксперимента — целенаправленно и необратимо изменять те или иные свойства личности.

В классическом психологическом поведенческом эксперименте устанавливается функциональная зависимость вида:

$$R = f(S),$$

где R — ответ, а S — ситуация (стимул, задача). Переменная S систематически варьируется, а детерминированные ею изменения ответа испытуемого фиксируются. В ходе изучения проявляются условия, при которых испытуемый ведет себя тем или иным образом. Результат фиксируется в форме линейной или нелинейной зависимости.

Другой тип зависимостей символизируется как зависимость поведения от личностных свойств или состояний организма испытуемого:

$$R = f(O) \text{ или } R = f(P),$$

где R - ответ, O - организм, P - личность. Исследуется зависимость поведения испытуемого от того или иного состояния организма (болезни, усталости, уровня активации, фрустрации потребностей и т.д.) или от личностных свойств

(тревожности, мотивации и т.д.). Исследования проводятся с участием групп людей, различающихся по данному признаку: свойству или актуальному состоянию.

Естественно, эти две строгие зависимости являются простейшими формами отношений между переменными. Возможны более сложные зависимости, устанавливаемые в конкретном эксперименте, в частности, факторные планы позволяют выявить зависимости вида $R = f(S1, S2)$, когда ответ испытуемого зависит от двух варьируемых параметров ситуации, а поведение является функцией состояния организма и среды.

Рассмотрим основные возможные **варианты отношений между независимой и зависимой переменными**. Существует как минимум шесть видов связи переменных. Первый, он же простейший, — **отсутствие зависимости**. Графически он выражается в форме прямой, параллельной оси абсцисс на графике, где по оси абсцисс (X) отложены уровни независимой переменной, а по оси ординат (Y) — зависимой переменной (рис.1.1). Зависимая переменная не чувствительна к изменению независимой.

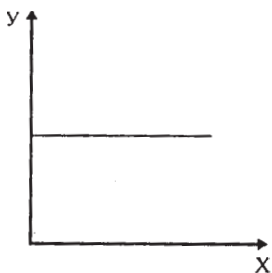


Рис.1.1

Монотонно возрастающая зависимость наблюдается тогда, когда увеличению значений независимой переменной соответствует изменение зависимой переменной (рис.1.2).



Рис.1.2

Монотонно убывающая зависимость наблюдается, если увеличению значений независимой переменной соответствует уменьшение уровня независимой переменной (рис.1.3).



Рис. 1.3

Нелинейная зависимость U-образного типа обнаруживается в большинстве экспериментов, в которых выявляются особенности психической регуляции поведения (рис. 1.4).



Рис. 1.4

Инвертированная U-образная зависимость получается в многочисленных экспериментальных и корреляционных исследованиях как в психологии личности, мотивации, так и в социальной психологии (рис. 1.5).

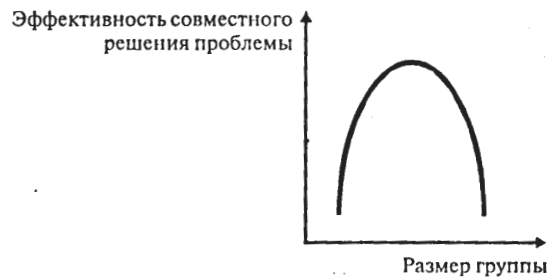


Рис. 1.5

Последний вариант зависимости обнаруживается не так часто, как предыдущие, — **сложная квазипериодическая зависимость** уровня зависимой переменной от уровня независимой (рис.1.6).



Рис.1.6

При выборе способа описания работает «принцип экономии»: простое описание лучше, чем все другие эквивалентные по успешности описания. Любое простое описание лучше, чем комплексное. Поэтому аргументы, распространенные в отечественных научных дискуссиях, типа «все гораздо сложнее на самом деле, чем представляет автор» по меньшей мере бессмысленны. Тем более, что никто не знает, как «на самом деле».

1.9.4. Внешние переменные

Внешние переменные экспериментальной ситуации экспериментатор должен контролировать. Среди внешних переменных выделяют: 1) **побочные переменные**, которые порождают систематическое смешение, ведущее к появлению ненадежных данных (фактор времени, фактор задачи, индивидуальные особенности испытуемых); 2) **дополнительную переменную**, которая существенна для изучаемой связи между причиной и следствием. При проверке частной гипотезы уровень дополнительной переменной должен соответствовать ее уровню в изучаемой реальности. Например, при изучении связи уровня развития непосредственного и опосредованного запоминания дети должны быть одного возраста. Возраст в этом случае — дополнительная переменная. Если же проверяется общая гипотеза, то эксперимент проводится при разных уровнях дополнительной переменной, т.е. с участием групп детей разного возраста.

Ключевой переменной называется дополнительная переменная, особо значимая для эксперимента, **контрольной переменной** — дополнительная переменная, которая в факторном эксперименте становится второй основной.

1.9.5. Контроль переменных

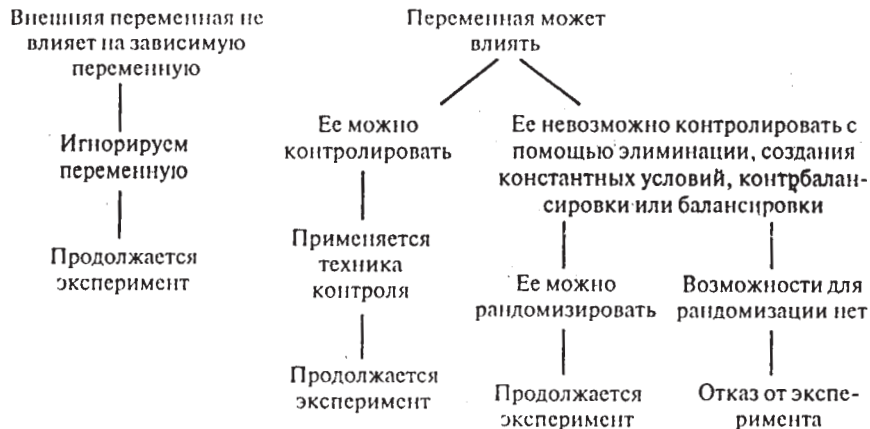
Следует различать контроль независимой переменной и контроль внешних (побочных и дополнительных) переменных. Контроль независимой переменной состоит в ее активном варьировании или знании закономерности ее изменения. **Латинский квадрат** — план организации многоуровневого эксперимента, в котором вторым экспериментальным фактором выступает порядок предъявления разных условий независимой переменной. Второй смысл понятия «контроль» — управление внешними переменными эксперимента. Влияние внешних переменных сводится к эффекту смешения.

Различают два основных **способа контроля независимой переменной**. Эти способы лежат в основе двух типов эмпирического исследования: активного и пассивного. Напомним, что в психологии к активным относятся деятельностный метод (эксперимент) и коммуникативный (беседа), а к пассивным — наблюдение и измерение. Пассивные методы называют также методами систематизированной регистрации, или систематизированного наблюдения (включая в него и процедуру измерения).

В эксперименте контроль независимой переменной производится с помощью активного манипулирования, варьирования. При систематизированном наблюдении (а также измерении) контроль осуществляется за счет отбора (селекции) требуемых значений независимой переменной из числа уже существующих переменных. Примером активного контроля является изменение громкости сигнала, подаваемого экспериментатором в наушники. Примером пассивного контроля может служить разбиение группы учеников на неуспевающих, среднеуспевающих и высокоуспешных при исследовании влияния уровня успешности обучения на статус личности в учебной группе.

При планировании исследования следует иметь в виду, что принципы, предъявляемые к формированию плана для активного и пассивного исследований, одни и те же, за исключением контроля эффектов, связанных с экспериментальным воздействием.

Обратите особое внимание на таблицу, в которой отображен предложенный МакГиганом (1993) **алгоритм пошагового контроля влияния внешних переменных на зависимую переменную**:



Рассмотрим последовательно различные **способы контроля внешних переменных**.

1. **Элиминация (исключение)**. Наиболее простой по сути, но не по возможностям осуществления способ контроля. Экспериментальную ситуацию конструируют таким образом, чтобы исключить какое-либо присутствие в ней внешней переменной. Например, в психофизических лабораториях часто создаются экспериментальные камеры, изолирующие испытуемого от внешних звуков, шумов, вибрационного воздействия и электромагнитных полей. Но зачастую элиминировать влияние внешних переменных невозможно. Например,

трудно представить себе, как можно исключить влияние таких переменных, как пол, возраст или интеллект.

2. Создание константных условий. Если внешние переменные не удается исключить из экспериментальной ситуации, то исследователю приходится делать их неизменными. При этом влияние внешней переменной остается неизменным на всех испытуемых, при всех значениях независимой переменной и на протяжении всего эксперимента. Однако эта стратегия не позволяет полностью избежать эффекта смешения: данные, полученные при константных значениях внешних переменных, можно переносить только на те реальные ситуации, в которых значения внешних переменных такие же, какими они были при исследовании. Исследователь стремится сделать неизменными внешние пространственно-временные условия проведения эксперимента. В частности, экспериментальные пробы или наблюдение за поведением проводятся со всеми испытуемыми в одно и то же время суток и в один и тот же день недели, например в понедельник в 9 часов утра. Однако это не гарантирует от эффекта смешения. Допустим, мы тестируем уровень достижений школьников при решении простых арифметических задач. Школьники-«совы», у которых уровень работоспособности приходится на вторую половину дня, будут в менее благоприятном состоянии, чем школьники-«жаворонки». Если они преобладают в группе, то их результаты будут смещены по сравнению с результатами, которые могли бы получиться на генеральной совокупности.

Следует стандартизировать технику проведения исследования и оборудование экспериментальных помещений (звуки, ароматы, окраску стен, вид мебели, расположение мебели и т.д.).

Исследователь стремится сделать константными дополнительные переменные — уравнивать группы испытуемых по основным значимым для исследования индивидуальным характеристикам (уровню образования, полу, возрасту).

Экспериментатор должен предъявлять инструкцию одинаково всем испытуемым (разумеется, исключая те случаи, когда она изменяется в соответствии с планом эксперимента). Он должен стремиться сохранять неизменными интонацию и силу голоса. Рекомендуется записывать инструкцию на магнитофон и предъявлять запись (кроме особых случаев).

3. Балансировка. В тех случаях, когда отсутствует возможность создать константные условия проведения эксперимента или константности условий недостаточно, применяют технику балансировки эффекта от действия внешних переменных. Балансировка применяется в двух ситуациях: 1) если невозможно идентифицировать внешнюю переменную; 2) если ее можно идентифицировать и использовать специальный алгоритм для контроля этой переменной.

Рассмотрим способ балансировки влияния неспецифических внешних переменных. Он состоит в том, что в дополнение к экспериментальной группе в план эксперимента включается контрольная группа. Экспериментальное исследование контрольной группы проводится в тех же условиях, что и экспериментальной. Отличие в том, что экспериментальное воздействие осуществляется только на испытуемых, включенных в экспериментальную группу. Тем самым изменение зависимой переменной в контрольной группе обусловлено лишь внешними переменными, а в экспериментальной — совместным действием внешних и независимой переменных.

Разумеется, при этом нельзя выделить специфическое влияние каждой внешней переменной и особенности такого влияния независимой переменной из-за эффекта взаимодействия переменных.

1. Способ балансировки с применением контрольной группы:



2. Способ балансировки с выделением эффекта внешней переменной:



Для того чтобы определить, как влияет на зависимую переменную та или иная внешняя переменная, используют план, включающий более чем одну контрольную группу. В общем случае число контрольных групп в экспериментальном плане должно быть $N = n + 1$, где n — число внешних переменных. Вторая контрольная группа помещается в экспериментальные условия, где исключено действие одной из внешних переменных, влияющих на зависимую переменную экспериментальной и первой контрольной групп. Различие в результатах 1-й и 2-й контрольных групп позволяет выделить специфическое влияние одной из внешних переменных.

Несколько отличается процедура балансировки при контроле известных внешних переменных. Типичный пример учета такой переменной — выявление уровня влияния принадлежности испытуемых к тому или иному полу на результаты эксперимента, поскольку известно, что многие данные, полученные на выборке мужчин, невозможно перенести на женскую выборку. Пол — это дополнительная переменная, поэтому планирование сводится к выявлению эффекта действия независимой переменной на зависимую в каждой из двух экспериментальных групп.

Аналогично строится эксперимент по сравнению эффекта различных аппаратных методик в зависимости от возраста испытуемых и др.

В более сложных экспериментах применяется балансировка нескольких переменных одновременно. Примером может служить учет влияния пола экспериментатора на поведение испытуемых при тестировании интеллекта. У нас две группы испытуемых, мужчин и женщин, и два экспериментатора (мужчина и женщина). План эксперимента может выглядеть следующим образом:

Группа 1 (эксперимент):

Мужчины — экспериментатор мужчина
 Мужчины — экспериментатор женщина
 Женщины — экспериментатор мужчина
 Женщины — экспериментатор женщина

Группа 2 (контроль):

Мужчины — экспериментатор мужчина
 Мужчины — экспериментатор женщина
 Женщины — экспериментатор мужчина
 Женщины — экспериментатор женщина

4. Контрбалансировка. Этот прием контроля дополнительной переменной чаще всего применяют тогда, когда эксперимент включает в себя несколько серий. Испытуемый оказывается в разных условиях последовательно, и предыдущие условия могут изменять эффект воздействия последующих условий. К примеру, при исследовании дифференциальной слуховой чувствительности не безразлично, какой звук, громкий или более тихий, предъявлялся испытуемому первым, а какой — вторым. Также при выполнении тестов интеллекта важен порядок предъявления испытуемому задач: от простой к сложной или от сложной к простой. В первом случае более развитые испытуемые больше утомляются и теряют мотивацию, так как вынуждены решать большее количество задач, чем менее развитые. При втором варианте предъявления заданий менее развитые испытуемые испытывают стресс неуспеха и вынуждены решать больше задач, чем их более развитые коллеги. В этих случаях для ликвидации эффектов последовательности и последствия используют контрбалансировку. Смысл ее состоит в том, что порядок предъявления разных задач, стимулов, воздействий в одной из групп компенсируется иным порядком предъявления заданий в другой группе.

Приведем пример плана контроля за внешней переменной для двух условий.

Группа	Экспериментальная серия (порядок)	
	1-я	2-я
1-я	Громкий звук	Тихий звук
2-я	Тихий звук	Громкий звук

Для трех независимых переменных применяется такой план контрбалансировки, например, как при предъявлении трех цветов — красного, желтого, зеленого:

Группа	Экспериментальная серия (порядок)		
	1-я	2-я	3-я
1-я	Кр.	Ж	З
2-я	Кр.	З	Ж
3-я	Ж	Кр.	З
4-я	Ж	З	Кр.
5-я	З	Кр.	Ж
6-я	З	Ж	Кр.

Контрбалансировка применяется в тех случаях, когда есть возможность провести несколько серий. Следует лишь учитывать, что большое число опытов может вызвать утомление у испытуемого. Но этот план позволяет контролировать эффект последовательности. Упрощение же плана контрбалансировки приводит к появлению эффекта последовательности. Однако контрбалансировка не позволяет полностью исключить еще один эффект, а именно влияние изменения порядка предъявления заданий на значение зависимой переменной. Он называется дифференцированным переносом: переход от ситуации 1 (когда она создается первой) к ситуации 2 отличается от перехода от ситуации 2 (когда она идет первой) к ситуации 1. Этот эффект приводит к тому, что реальные различия между двумя разными экспериментальными ситуациями при регистрации преувеличиваются.

Итак, техника контрбалансировки заключается в том, что каждый испытуемый получает более чем один вариант воздействия (AB или BA) и эффект последовательности целенаправленно распределяется на все экспериментальные условия.

При балансировке каждый испытуемый получает лишь одно экспериментальное воздействие — внешняя переменная балансируется за счет выявления эффекта ее действия на членов экспериментальной группы, по сравнению с эффектом, полученным при исследовании контрольной группы. Испытуемый может оказаться только в экспериментальной или же только в контрольной группе и получить воздействие какой-нибудь внешней переменной в обеих группах. Балансировка используется при исследовании независимых групп, тогда как контрбалансировка применяется в исследованиях с повторяющимися воздействиями.

5. Рандомизация. Рандомизацией называется процедура, которая гарантирует равную возможность каждому члену популяции стать участником эксперимента. Каждому представителю выборки присваивается порядковый номер, а выбор испытуемых в экспериментальную и контрольную группы проводится с помощью таблицы случайных чисел. Рандомизация является способом, позволяющим исключить влияние индивидуальных особенностей испытуемых на результат эксперимента.

Рандомизация применяется в двух случаях: 1) когда известно, как управлять внешними переменными в экспериментальной ситуации, однако у нас нет возможности использовать одну из предшествующих техник контроля; 2) когда мы предполагаем оперировать какой-либо внешней переменной в экспериментальной ситуации, однако не можем ее специфицировать и применить другие техники.

Если предположить, что значение дополнительной переменной подчиняется вероятностным законам (например, описывается нормальным распределением), то в состав экспериментальной и контрольных групп войдет выборка, которая имеет те же уровни дополнительных переменных, что и генеральная совокупность.

По мнению многих специалистов, в том числе Кэмпбелла, уравнивание групп посредством процедуры рандомизации является единственно надежным способом элиминации влияния внешних (побочных и дополнительных) переменных на зависимую. Кэмпбелл определяет рандомизацию как универсальный способ уравнивания групп перед экспериментальным воздействием. Другие способы, например метод попарного сравнения, характеризуются им как малонадежные и ведущие к невалидным выводам.

Эти приемы, разумеется, не позволяют полностью избежать воздействий со стороны внешних переменных, однако их выполнение является своеобразной профилактической процедурой: мытье рук перед едой не дает 100%-ной гарантии от заболевания дизентерией, но существенно снижает вероятность заболевания.

2. ВНУТРЕННЯЯ И ВНЕШНЯЯ ВАЛИДНОСТЬ

2.1. Виды валидности при оценке эксперимента

Валидность исследования — оценка проведенного эмпирического исследования с точки зрения правильности его организации и возможности считать полученные результаты и сделанные выводы достоверными.

Проблема перехода от теории к уровню экспериментальной гипотезы — это вопрос **конструктивной валидности** эксперимента. Решение проблем операционализации переменных в методических процедурах их измерения (и управления) — вопросы **операциональной валидности**. Если обобщение зависимости включает перенос ее на “жизненные”, т.е. реальные, виды ситуаций и деятельности человека — это решение проблем соответствия переменных с точки зрения **внешней валидности**, или, как ее иногда называют, “экологической” валидности.

Конструктивная валидность обсуждается при любых типах экспериментов как адекватность перехода от научной гипотезы к экспериментальной или от “рабочей” к интерпретационной теоретической гипотезе. Понятно, что вопросы операциональной валидности охватывают этап перехода от уже сформулированных экспериментальных и контргипотез к процедурам их воплощения. Р. Готтсданкер рассматривает понятие операциональной валидности только применительно к такому типу эксперимента, как лабораторный. В этом случае не различаются два названных этапа разработки методических процедур, поскольку предполагается представленность психологического конструкта в самом способе его измерения. Однако с этим можно спорить, учитывая несводимость обобщенного понятия к его эмпирическому наполнению.

Оценка валидности связана как с оценкой реализации выбранных форм экспериментального контроля, так и с оценкой системы умозаключений при организации исследования с точки зрения следования нормативам экспериментального вывода. Нормативы, связанные с возможными обобщениями из психологических экспериментов, предполагают дифференциацию разных видов валидности.

Психологическая гипотеза, эксплицируемая, или выводимая, как следствие из той или иной теоретической концепции, при повторных ее проверках может несколько видоизменяться, если изменяются средства операционализации входящих в нее переменных. В том случае, когда предполагается выявление одной и той же теоретически предполагаемой зависимости, но психологические конструкты конкретизируются разными методическими средствами, это называется **концептуальными репликациями**.

Так, при множественных измерениях зависимой переменной разными методиками повышается обоснованность утверждений о характере изучаемого базисного процесса. Изменения типов задания независимых переменных также расширяют диапазон обобщений устанавливаемой психологической закономерности. Существенно, что при этом имеется в виду однозначность интерпретации самого психологического конструкта (“мотивация достижения”, “функциональные состояния”, “гибкость-ригидность контроля” и т.д.).

На самом деле в психологии часто при использовании одного и того же понятия изучаются разноуровневые базисные процессы, то есть в экспериментах с терминологически сходными формулировками гипотез представлены достаточно отличающиеся друг от друга типы переменных.

Специальным случаем является обобщение, связанное с организацией в эксперименте “очищенных” условий для проверки так называемых **точных гипотез**. Обычно это функция лабораторного эксперимента: получить данные в условиях, отвечающих причинно-следственным зависимостям в специально

сформулированной модели, чтобы эмпирически оценить объяснительную силу этой теоретической модели. Однако обобщения теоретического характера связаны не только с таким ходом рассуждения, как “эксперимент — модель — теория”. Оценка силы тех или иных обобщений связана с анализом свойств самой теории уже применительно к “миру теорий” (а не к миру “психологической реальности”). Независимо от того, моделируется ли в эксперименте ситуация, отражающая “мир теорий” (т.е. теоретическую модель) или “внешний мир”, оценивается как объяснительная сила высказывания, сформулированного в виде экспериментальной гипотезы, так и возможность перехода от той или иной теории как системы объяснений к этому эмпирически нагруженному высказыванию. Экспериментальная гипотеза как эмпирическая гипотеза будет уже включать переменные, которые можно не только наблюдать, но и измерять по той или иной шкале. Однако не все теоретические гипотезы смогут найти свой эмпирический базис.

Итак, в целом о валидности эксперимента говорят, подразумевая все формы контроля, т.е. все виды валидности. Результаты валидного эксперимента могут служить основанием **достоверного вывода**, если в целом реализована логически обоснованная система умозаключений, включающая взаимопереходы между разными уровнями гипотез, проверяемых в эксперименте. Остается подчеркнуть, что ошибки выводов, или **недостоверные выводы**, возможны как в результате неверных обобщений, так и в результате проведения невалидных экспериментов.

При понимании экспериментирования как чувственной и предметной деятельности ученого не следует, таким образом, ограничиваться описанием и оценкой того, как операционализируются переменные или как осуществляется их регистрация. Формулирование системы гипотез, оценка валидности эксперимента и реализация выводов входят как нормативы, или нормативные регуляторы, этой деятельности.

Пока мы не затронули проблему, в какой степени выбор конкретной формы эксперимента связан с разработкой того или иного экспериментального плана (как схемы сбора данных при разных условиях независимой переменной) и с другими типами контроля (выбор типа эксперимента, осуществление “первичного” контроля, “дополнительного варьирования” переменных, введение расширенной переменной и т.д.). Это станет возможным после введения в проблемы планирования и описания оснований классификаций типов экспериментов.

2.2. Внутренняя валидность

Внутренняя валидность — критерий оценки планирования и проведения исследования, позволяющий считать достоверными выводы относительно представленной в гипотезе зависимости.

Внутренняя валидность эксперимента обеспечивает вывод об исследуемой зависимости, т.е. об отношении между независимой и зависимой переменными. Если полученные данные характеризуются ненадежностью или наличием смещений (систематических, несистематических, сопутствующих), то нельзя считать обоснованным утверждение, формулируемое в экспериментальной гипотезе.

Кэмпбелл отмечает, что установление внутренней валидности требует отбраковывания альтернативных объяснений связи между зависимой и независимой переменными, а установление конструктивной валидности — альтернативных интерпретаций соотношения причины и следствия с понятиями, взятыми из той или иной теории.

С точки зрения Кэмпбелла, **хороший эксперимент** должен: 1) выявлять временную последовательность предполагаемых причины и следствия; 2) показывать, что вероятные причины и эффект взаимосвязаны; 3) исключать

влияние побочных переменных, которыми можно было бы объяснить экспериментальный эффект; 4) исключать альтернативные гипотезы о теоретических конструктах, объясняющих эту связь.

2.3. Внешняя валидность

Внешняя валидность — критерий оценки исследования с точки зрения решения проблем соответствия сформулированных в гипотезе переменных и реализованных методически условий тем, на которые осуществляется обобщение полученных результатов.

Внешняя валидность предполагает решение проблем соответствия (независимой переменной, зависимой переменной, внешних переменных), обеспечивает возможность обобщения на те виды ситуаций или на те виды деятельности, которым соответствуют экспериментально контролируемые переменные. Если речь идет о теоретических обобщениях, то оценка внешней валидности уступает место оценке конструктивной валидности. Решение вопросов о конструктивной и внешней валидности приводит исследователя к выбору типа эксперимента: “дублирующего” реальный мир, искусственного или лабораторного.

Внешняя валидность имеет особое значение на эмпирической стадии развития науки. В принципе возможны такие эксперименты, которые не соответствуют никаким реальным жизненным ситуациям, а служат лишь для проверки гипотез, источником которых является развитая теория. В развитых науках стремятся избегать “прямого замыкания” экспериментальный результат - реальность, так как понятно, что эксперимент строится исходя из требований проверяемой теории, а не из требований соответствия реальности. Моделирование некоторых условий, например, в экспериментах по сенсорной депривации или выработке классических условных рефлексов не соответствует никакой жизненной реальности, если под реальностью понимать то, что было, а не то, что потенциально может быть.

2.4. Проверка каузальных отношений

Эксперименты являются средством проверки каузальных гипотез. **Гипотеза каузальная** — причинно-следственная гипотеза, которая включает содержательную интерпретацию взаимосвязи между независимой и зависимой переменными, формулируется как высказывание о влиянии причинно действующих условий на изучаемый базисный процесс.

По традиции предполагается, что существуют три необходимых условия для утверждения с определенной долей уверенности о том, что отношение между двумя переменными является каузальным и что направление причинности от А к В.

Первое условие относится к временной последовательности и формулируется следующим образом: причина должна предшествовать по времени эффекту. Обычно это условие соблюсти нетрудно, если исследователь знает, когда испытуемые подвергались экспериментальному воздействию. В квазиэкспериментах исследователь может сочетать такого рода знание со своим знанием результатов испытуемых до и после экспериментального воздействия и может связать воздействие с некоторыми показателями изменения зависимой переменной. Исследователь, проводящий истинный эксперимент, знает, что соответствующая рандомизация обеспечивает, в вероятностном плане, предварительную (до введения экспериментального воздействия) эквивалентность различных экспериментальных групп. Поэтому если после экспериментального воздействия обнаружатся различия, то он полагает, что такие различия, по всей вероятности, вызваны этим воздействием.

Второе условие для обоснованного вывода о причинной зависимости В от А состоит в том, что воздействие или воздействия должны быть статистически связаны с эффектом, так как, если возможная причина и эффект не связаны друг с другом, одно не может быть причиной другого. Для проверки существования такой ковариации используются статистические показатели; разработаны критерии для принятия решения о существовании “действительной” ковариации в результатах (например, $p < 0,05$). Таким образом, статистические показатели действуют как фильтры. К сожалению, они подвержены ошибкам даже в тех случаях, когда используются должным образом. С их помощью не всегда удастся выявить как действительно существующую, так и ложную композицию ковариации. Поскольку использование статистических показателей ведет к весьма важным решениям, разумно было бы показать те причины, которые порождают неверные выводы о ковариации. Обозначим их как **угрозы для валидности статистического вывода**.

Третье условие каузального вывода состоит в том, что не должно быть правдоподобного альтернативного объяснения появления В помимо А. Это условие труднее всего соблюсти в связи с двумя обстоятельствами. Первое, относящееся в основном к квазиэкспериментам, связано с валидностью альтернативных интерпретаций, которые предполагают, что кажущаяся каузальная связь А и В фактически может быть обусловлена действием третьей переменной, которая и вызывает изменения В. К примеру, можно установить на фабрике новую машину и отмечать, связано ли это с ростом производительности труда. Если это так, то можно приписать этот рост нововведению. Однако всегда существует вероятность того, что он никак не связан с новой машиной и может быть обусловлен сезонным ростом производительности, который происходит в такое время каждый год. И это только один из примеров такой третьей переменной. В дальнейшем мы представим систематизированный перечень таких дополнительных переменных, обозначая их как **угрозы для внутренней валидности** (см. подразд. 2.5).

Существование угроз для внутренней валидности наводит на мысль, что предполагаемая причинно-следственная связь может быть лишь видимой. Это отличается от второго значения “альтернативной интерпретации”, когда А, которым манипулируют, и В, которое измеряется, и в самом деле каузально связаны, но существуют сомнения, действительно ли эмпирические операции осуществляются с объектами, представленными конструктами А и В, которые исследователь пытался применить для обозначения А и В. Большинство теоретических контroversий в психологии именно такого рода — например, контroversий о том, можно ли объяснить связь между увеличением оплаты и более высокой производительностью тем, что высокая зарплата формирует чувство неравенства или разрушает ожидания или “я-концепцию” и т.д. В таких исследованиях речь идет не о внутренней валидности, то есть не о том, являются ли изменения в оплате причиной изменений в производительности. Скорее, это вопрос о том, как оплата должна быть выражена в терминах соответствующей теории, доступных обобщению.

Мы рассматриваем опасности для конструктивной валидности, которые следует понимать как угрозы для правильного обозначения причины и эффекта с помощью абстрактных терминов, взятых из обыденного языка или из формальной теории. В действительности проблема конструктивной валидности несколько шире и, очевидно, имеет отношение к попыткам обозначения любых аспектов эксперимента, включая характер обстановки, в которой проводится эксперимент, особенности участвующих лиц и т.д.

Стоит отметить, что термин “внутренняя валидность” в прошлом употреблялся не вполне правильно, потому что с его помощью выражали как сомнения в правильности вывода о причинных связях между А и В, так и сомнения

в том, как следует обозначать причину и следствия. Такое смешение понятий может возникнуть вследствие того, что альтернативные интерпретации должны быть отброшены в процессе установления как внутренней, так и конструктивной валидности.

Однако установление внутренней валидности предполагает отбрасывание альтернативных интерпретаций вероятных каузальных связей между А, которым манипулируют, и В, которое измеряется, а установление конструктивной валидности предполагает отбрасывание альтернативных интерпретаций того, как А и В соотносятся с гипотетическими понятиями. Поскольку в экспериментах крайне важно выяснить, является ли связь между двумя переменными каузальной, необходимым условием для выведения заключения о каузальности является отбрасывание альтернативных интерпретаций связи между переменными (то есть устранение угроз для внутренней валидности), а не отбрасывание альтернативных интерпретаций операций с причиной и следствием (то есть угроз для конструктивной валидности).

Заключение о каузальной связи в какой-то промежуток времени, в одной обстановке и при одной выборке испытуемых дает мало уверенности в том, что обнаруживаемая каузальная связь является устойчивой. Вопрос о возможности обобщения результатов для различных периодов времени, различных условий и групп мы рассматриваем как вопрос о внешней валидности.

К сожалению, социальным наукам не посчастливилось иметь точные теории, надежные измерения или повторяющийся циклический порядок в результатах наблюдений. Представьте себе, что обнаружено различие в производительности труда работника до и после того, как он прошел специальное обучение. Как мы можем определить, что в этом различии следует отнести за счет собственно обучения, что за счет естественного развития работника, или улучшений в навыке выполнения задания при обследовании, или в результате различных событий в жизни человека, которые оказывают влияние на зависимую переменную между предварительным тестированием и тестированием после экспериментального воздействия, и что объясняется любым сочетанием этих факторов? Более того, даже если бы мы могли определить конкретные численные величины, связанные с каждым из перечисленных выше объяснений, можно ли быть уверенными в том, что нам удастся измерить соответствующие результаты с такой надежностью, которая позволила бы осуществить выбор между теориями? Ответ на эти вопросы зависит, конечно, от величины предсказываемых различий в приросте производительности труда и от конкретного вида проверки. Мы полагаем, однако, что в социальных науках найдется немного неэкспериментальных условий, в которых точные предсказания могли бы быть с успехом использованы для проверки конкурентных гипотез.

Хотя предшествующее изложение наводит на мысль, что эксперименты более пригодны для проверки предположений о причинности, чем неэкспериментальные исследования, из этого не следует делать вывод, что эксперименты являются надежным средством получения ответов на все вопросы, связанные с проверкой гипотез о каузальных связях. Перечень угроз внутренней и внешней валидности, валидности статистического вывода и конструктивной валидности, а также выявление взаимосвязей между этими видами валидности делают очевидным несовершенство экспериментирования. Экспериментальные планы нуждаются в улучшении, могут и должны быть улучшены, чтобы способствовать установлению каузальной зависимости. Но мы заблуждались бы, если бы считали, что один эксперимент или даже целая исследовательская программа, рассчитанная на несколько лет, могли бы дать окончательный ответ на основные вопросы, связанные с надежным установлением каузальной связи, обозначением ее составляющих и определением возможности обобщения вывода.

2.5. Факторы, угрожающие внутренней и внешней валидности

В этом подразделе приводятся 12 факторов, угрожающих валидности различных экспериментальных планов. Каждый фактор будет подробнее описан при обсуждении тех планов, в которых его действие проявляется наиболее отчетливо. Основным при составлении этого списка было различение внутренней и внешней валидности.

Внутренняя валидность — это тот минимум, без которого не может быть интерпретирован ни один эксперимент: действительно ли именно это экспериментальное воздействие привело к изменениям в данном эксперименте? Внешняя валидность относится к возможности обобщения вывода: на какие популяции, ситуации, другие независимые переменные, параметры воздействия и переменные измерения могут быть распространены результаты эксперимента. Несомненно, что оба типа критериев являются важными, хотя часто они находятся в противоречии в том смысле, что увеличение валидности одного типа может угрожать валидности другого типа. Хотя внутренняя валидность является неперенным условием и хотя проблема внешней валидности, как и проблема индуктивного вывода, никогда не может быть полностью решена, очевидно, что нашим идеалом является выбор таких экспериментальных планов, которые обеспечивают оба типа валидности. Это особенно важно для педагогических исследований, где весьма желательным является распространение данных на практические ситуации определенного характера. Различия и взаимоотношения между этими двумя видами валидности станут яснее после рассмотрения примеров при обсуждении конкретных типов эксперимента.

Факторы, ставящие под угрозу внутреннюю валидность:

- 1) **фактор фона** - конкретные события, которые происходят между первым и вторым измерением наряду с экспериментальным воздействием;
- 2) **фактор естественного развития** — изменения испытуемых, являющиеся следствием течения времени, например взросление, усиление голода, усталости;
- 3) **эффект тестирования** — влияние выполнения заданий, применяемых для измерения, на результаты повторного испытания;
- 4) **инструментальная погрешность** - нестабильность измерительного инструмента, при которой изменения в калибровке инструмента могут вызвать изменения в результатах измерения;
- 5) **статистическая регрессия** - фактор, имеющий место тогда, когда группы отбираются на основе крайних показателей и оценок;
- 6) **фактор отбора испытуемых** — неэквивалентность групп по составу, вызывающая появление систематической ошибки в результатах эксперимента;
- 7) **фактор отсева испытуемых** — неравномерность выбывания испытуемых из сравниваемых групп в ходе эксперимента;
- 8) **взаимодействие фактора отбора с естественным развитием** и др., которые в ряде квазиэкспериментальных планов с несколькими группами ошибочно принимаются за эффект экспериментальной переменной;

К факторам, ставящим под угрозу внешнюю валидность, или репрезентативность эксперимента, относятся:

- 9) **реактивный эффект** - возможное уменьшение или увеличение восприимчивости испытуемых к экспериментальному воздействию под влиянием предварительного тестирования. Результаты исследования лиц, прошедших предварительное тестирование, будут нерепрезентативны по отношению к тем, кто не подвергался предварительному тестированию, то есть тем, из кого состоит генеральная совокупность, из которой были отобраны испытуемые;
- 10) **эффекты взаимодействия фактора отбора и экспериментального воздействия;**
- 11) **условия организации эксперимента, вызывающие реакцию испытуемых**

на эксперимент, которая не позволяет распространить полученные данные о влиянии экспериментальной переменной на лиц, подвергающихся такому же воздействию в неэкспериментальных условиях;

12) **взаимная интерференция экспериментальных воздействий**, нередко возникающая, когда одни и те же испытуемые подвергаются нескольким воздействиям, поскольку влияние более ранних воздействий, как правило, не исчезает.

2.6. План с предварительным и итоговым тестированием и контрольной группой

2.6.1. Средства контроля внутренней валидности

Контроль угроз внутренней валидности — способ контроля, который осуществляется благодаря устранению источников несистематической изменчивости при приближении эксперимента к бесконечному или идеальному.

В плане с предварительным и итоговым тестированием и контрольной группой уравнивание групп достигается путем рандомизации. В аналогичном квазиэкспериментальном плане (см. юниту 6) используются реально существующие группы без уверенности в том, что они эквивалентны. План с предварительным и итоговым тестированием и контрольной группой имеет следующий вид:

R	O1	X	O2
R	O3		O4

где R - случайное распределение испытуемых по разным группам, O - процесс наблюдения или измерения, X - независимая переменная, влияние которой подлежит измерению.

В традиционных исследованиях научения к признанию необходимости контрольных групп привело выявление эффектов тестирования. В педагогических исследованиях главной проблемой было естественное развитие, как в детской психологии — проблема “врожденное — приобретенное”. В исследованиях изменения установок, например в ранних исследованиях влияния кино, наиболее важным является необходимость контролировать фон. Во всяком случае полезно, видимо, кратко обсудить здесь, как или при каких условиях осуществляется контроль за этими факторами.

Фон (“история”) контролируется в той мере, в какой общие события, которые могут привести к различию O1 — O2, способны привести также к различию O3 — O4. Однако отметим, что во многих возможных применениях этого плана не контролируются специфические события, имевшие место в группе в ходе обследования (“внутригрупповая история”). Если все испытуемые экспериментальной группы, сформированной на основе рандомизации, подвергаются однократному экспериментальному воздействию и контрольная группа также исследуется однократно, то посторонние частные события, происходящие в каждой из этих групп (шутка, вызвавшая оживление, пожар на улице, предварительные замечания экспериментатора и т.д.), могут стать конкурентными объяснениями отличия O1 — O2 от O3 — O4.

Вполне возможно, что именно внутригрупповые события и есть тот самый источник “значимых”, но посторонних различий, который является дефектом исследовательской практики. В типичном эксперименте контроль внутригрупповых событий обеспечивается тем, что тестирование испытуемых проводится в индивидуальном порядке и путем случайного распределения опытов и испытуемых по экспериментальным или контрольным режимам. Тем не менее

надо отметить, что даже при индивидуальной работе фон может оказаться неконтролируемым, если вся экспериментальная группа проходит тестирование до контрольной и т.д.

План с предварительным и итоговым тестированием и контрольной группой рассчитан на одновременное проведение исследования в экспериментальной и контрольной группах. Если же действительно проводить опыты одновременно, то необходимо использовать различных экспериментаторов, при этом различия между экспериментаторами в свою очередь становятся одним из видов внутригрупповой истории, смешиваемой с X.

Оптимальным решением является рандомизация отдельных сеансов эксперимента с учетом тех ограничений, которые связаны с необходимостью уравнивания таких важных источников смещений, как экспериментатор, время дня, день недели, начало или конец семестра, близость экзаменов и т.п. Все испытуемые, участвующие в одном сеансе, подвергаются воздействию одних и тех же посторонних событий, что может оказаться отличной от X причиной сходства их реакций.

Для некоторых исследований, связанных с групповым тестированием, можно распределить экспериментальные воздействия и предъявлять их отдельным лицам внутри одной группы в случайном порядке, подобно тому как это делается при использовании множественных форм теста, когда изучается влияние порядка трудности отдельных заданий по тесту. В таких случаях особенности "внутригрупповой истории" будут общими для обоих видов процедур и при объяснении полученных различий они не смогут стать правдоподобной конкурентной гипотезой, смешиваемой с X.

Естественное развитие и эффект тестирования контролируются постольку, поскольку они одинаковым образом проявляются в экспериментальной и контрольной группах. В тех случаях, когда созданы условия для контроля "внутригрупповой истории", легко контролировать и инструментальную погрешность, особенно если для O используются ответы испытуемых, полученные с помощью фиксированной методики, такой, как напечатанный тест. Однако при использовании наблюдателей или интервьюеров проблема становится более сложной. В случае, когда число наблюдателей не позволяет осуществлять их рандомизированное распределение по отдельным сеансам, необходимо использовать каждого наблюдателя и в экспериментальных, и в контрольных сеансах; кроме того, для предупреждения искажения результатов наблюдений необходимо, чтобы наблюдатели не знали, каким видам воздействий подвергается каждый из испытуемых. Регистрация группового взаимодействия облегчит контроль инструментального фактора и в исследованиях поведения и взаимодействия в группах учащихся, поскольку отдельные фрагменты записи, полученные при предварительном и итоговом тестировании в экспериментальной и контрольной группах, могут предъявляться в случайном порядке.

Эффект регрессии контролируется, когда речь идет о различиях между средними, независимо от того, насколько экстремальными были показатели группы при предварительном тестировании, если экспериментальная и контрольная группы были составлены из некоторой экстремальной совокупности на основе рандомизации. В этом случае контрольная группа имеет такую же регрессию, что и экспериментальная. Однако даже при использовании этого плана часто возникают ошибки интерпретации из-за регрессионных артефактов. Экспериментатор может использовать контрольную группу для подтверждения группового среднего эффекта X, а затем при определении того, какие предварительные показатели подгрупп испытуемых из экспериментальной группы особенно подвергались влиянию X, оставить результаты контрольной группы без внимания. Если у экспериментальной группы в целом обнаружится улучшение результатов, то экспериментатор столкнется с поразительным явлением,

состоящим в том, что у тех, кто вначале имел самые низкие показатели, это улучшение было наибольшим, в то время как у тех, кто вначале имел наивысшие показатели, не наблюдается никакого прогресса. В этих условиях такие результаты достигаются за счет того, что регрессионный артефакт способствует увеличению показателей у тех, кто имел наихудшие предварительные показатели, и нейтрализует их у тех, чьи предварительные показатели были наилучшими. Если же в целом не наблюдается улучшения результатов, то экспериментатор может ошибочно “обнаружить”, что это произошло вследствие двух взаимно компенсирующих процессов: прироста у “низких” и потери у “высоких”.

Одно из средств предупреждения такой ошибки интерпретации — проведение параллельного анализа экстремальных показателей предварительного тестирования в контрольной группе и использование этих данных при интерпретации увеличения показателей путем сравнения результатов итогового тестирования соответствующих — выделенных по данным предварительного тестирования — экспериментальных и контрольных подгрупп. Отметим, что асимметрия распределений, получаемых при отборе подгрупп, ставит под сомнение пригодность процедур, основанных на нормальной кривой.

Действие **фактора отбора, или состава групп**, в качестве объяснения различий исключается в той степени, в какой рандомизация обеспечивает эквивалентность групп в момент R. Эта степень определяется принятой нами выборочной статистикой. Поэтому уверенность в равенстве выше для большого числа случайных распределений, чем для малого. В пределах, определяемых величиной ошибки, для гипотезы об отсутствии различий это утверждение иногда оказывается неверным. Это означает, что иногда могут обнаруживаться явно “значимые” различия между результатами предварительного тестирования. Поэтому, хотя простая, или стратифицированная, рандомизация и обеспечивает несмещенное распределение испытуемых по группам, она не дает абсолютной уверенности в исходной эквивалентности таких групп. Тем не менее рандомизация — единственный способ решения. Это утверждение столь категорично потому, что в педагогических исследованиях наблюдается весьма распространенное ошибочное предпочтение способа попарного уравнивания групп.

Попарное уравнивание фактически не способствует преодолению исходных групповых различий. Это не исключает применения попарного подбора в качестве дополнения к рандомизации, когда для достижения статистической точности подбирают учащих по парам, а затем в случайном порядке одного члена каждой пары включают в экспериментальную группу, а другого — в контрольную. Но применение попарного уравнивания вместо рандомизации недопустимо даже в квазиэкспериментальных планах, использующих лишь две реально существующие группы: одну — экспериментальную, другую — контрольную. Даже в этих “слабых экспериментах” существуют лучшие, чем попарное уравнивание, способы уменьшить исходные средние значения различия в этих двух выборках.

Данные, которые можно получить с помощью этого плана, позволяют установить, может ли выбывание объяснить прирост O1 — O2. Потери, выбывания и случаи, когда доступны лишь частичные данные, представляют немалые трудности, и обычно о них не упоминают. Как правило, эксперименты с методами обучения длятся в течение нескольких дней, недель или месяцев. Если предварительное и итоговое тестирования проводятся в классах, из которых составляются экспериментальная и контрольная группы, и условия эксперимента требуют присутствия испытуемых на занятиях, в то время как испытуемые контрольной группы эти занятия не посещают, то различная посещаемость на трех этапах — предварительного тестирования, экспериментального воздействия, итогового тестирования — ведет к выбыванию. Этот фактор может вызвать неуловимые смещения выборки. Если исключить из экспериментальной группы тех, кто первоначально был намечен для участия в ней, но не смог посещать

экспериментальные занятия, это вызовет такое селективное сокращение группы, которое сделает ее несопоставимой с контрольной группой. В частности, экспериментальная группа оказывается “более добросовестной”. По-видимому, более предпочтительным было бы — хотя это редко наблюдается — использование результатов всех испытуемых контрольной и экспериментальной групп, которые подвергались предварительному и итоговому тестированию, включая и тех участников экспериментальной группы, которые не подвергались воздействию X. Очевидно, при таком способе анализа эффект X выглядит ослабленным, но зато это позволяет избежать смещения выборки.

Такая процедура основывается на допущении, что нет более простых смещений вследствие выбывания; такое допущение может быть частично проверено путем анализа числа и предварительных показателей тех лиц, кто проходил предварительное тестирование, но отсутствовал при итоговом тестировании. Возможно, что некоторые X влияют скорее на степень выбывания, а не на изменение индивидуальных показателей. Конечно, даже если степень выбывания одна и та же, остается возможность сложных взаимодействий, которые делают различным характер выбывания из экспериментальной и контрольной групп.

2.6.2. Средства контроля внешней валидности

Контроль угроз внешней валидности — способ контроля, который осуществляется благодаря тому, что это эксперимент полного соответствия. Описанные выше факторы внутренней невалидности были факторами, непосредственно влияющими на данные O. И эти факторы сами по себе могли вызвать изменения, которые можно ошибочно принять за результат действия X. Если в план эксперимента включается еще и контрольная группа, эти факторы вызывают эффекты, проявляющиеся и в контрольной группе; в экспериментальной же группе их действие накладывается на действие X. В терминах дисперсионного анализа фон (“история”), естественное развитие, эффект тестирования и т. д. могут быть представлены как главные эффекты и как таковые они могут контролироваться в эксперименте, обеспечивая ему внутреннюю валидность. С другой стороны, угрозы внешней валидности могут быть отнесены к эффектам взаимодействия X с некоторыми другими переменными. Тем самым они представляют собой потенциальные ограничения эффектов X некоторой нежелательно узкой совокупностью условий.

Как известно, наблюдаемые эффекты действия X могут проявляться только в группах, “подготовленных” предварительным тестированием. Логически мы не имеем права распространять эти результаты на ту большую, предварительно не обследованную популяцию, о которой мы хотели бы говорить.

В этом подразделе мы обсудим несколько таких угроз обобщаемости и способы их уменьшения. Так, поскольку существуют валидные планы, в которых не используется предварительное тестирование, и поскольку во многих ситуациях (не обязательно в исследованиях обучения) хотят распространить результаты на группы, не подвергавшиеся предварительному тестированию, такие планы более предпочтительны, так как обеспечивают большую внешнюю валидность или возможность обобщения. В области обучения часто высказываемые сомнения по поводу возможности применения результатов крайне искусственных экспериментов в реальной практике являются суждениями о внешней валидности. Поэтому учет таких соображений при обсуждении оптимальных экспериментальных планов встретит одобрение тех практических работников, которые правильно понимают, что такие соображения безосновательно игнорируются в обычных формальных курсах по экспериментальной методологии. Ниже мы покажем несколько путей повышения внешней валидности экспери-

ментов, создания более подходящих условий для обобщения их результатов и применения их в практике обучения без утраты внутренней валидности.

Прежде чем перейти к обсуждению этих вопросов, необходимо сделать несколько предостережений. Они касаются некоторых острых в учении об индукции проблем. Они являются таковыми вследствие периодически повторяющегося вынужденного признания того, что индукция, или обобщение никогда не могут быть полностью логически обоснованы. В то время как проблемы внутренней валидности разрешимы в рамках логики вероятностной статистики, проблемы внешней валидности логически не разрешимы никаким ясным, убедительным способом.

Если имеется внутренне валидный план, то эффект продемонстрирован только для тех конкретных условий, которые оказались общими для экспериментальной и контрольной групп, то есть только для предварительно обследованных групп, включающих лиц определенного возраста, интеллекта, социально-экономического статуса, географического региона, взятых в определенный исторический момент, при данной ориентации светил, определенном атмосферном давлении, уровне гамма-излучения и т.д.

Логически мы не имеем права выходить за эти пределы, то есть не имеем права вообще обобщать. Но в действительности мы пытаемся делать обобщения, устанавливая закономерности и отбрасывая некоторые из этих обобщений в других столь же конкретных, но отличных условиях. Изучая историю науки, мы узнаем о “подтверждении” обобщения посредством накопления опыта обобщения, но это не логическое обобщение, выводимое из данных исходного эксперимента. Столкнувшись с этим, мы при обобщении высказываем предположения о пока еще не доказанных закономерностях, включая и такие, которые совсем не исследовались. Так, при исследовании процесса преподавания мы полностью принимаем предположение о том, что ориентация в магнитном поле не оказывает влияния на результаты. Но на основе отдельных исследований мы знаем, что предварительное тестирование часто влияет на результаты исследования, и поэтому его как ограничение для обобщения следует устранить. Если же мы изучаем железные бруски, то по опыту знаем, что процедура предварительного взвешивания не оказывает никакого влияния на результаты последующих замеров, но ориентация в магнитном поле, если она систематически не контролируется, может в значительной степени ограничить обобщение результатов наших исследований.

Взаимодействие тестирования и X. Если ограничиваться только работами по планированию экспериментов как таковых, то приоритет в выделении угрозы внешней валидности со стороны предварительного тестирования принадлежит Соломону, хотя еще раньше аналогичные соображения привели отдельных экспериментаторов к использованию плана, в котором нет предварительного тестирования. Особенно отчетливо эта проблема предстает в исследованиях по изменению установок, где тесты для измерения последних содержат много весьма необычного материала (например, редко можно встретить в печатном тексте столько враждебных заявлений, сколько их содержится в типичном тесте о национальных предрассудках), и вполне вероятно, что предварительное тестирование оказывает влияние на установки личности и ее подверженность убеждению.

Если сделать аудиторию сензитивной посредством предварительного тестирования, то может возрасти обучающее влияние X, так как внимание аудитории будет определенным образом сфокусировано. Понятно, что такое воздействие X может быть эффективно лишь для предварительно тестируемой группы.

Снижение внешней валидности, обусловленное влиянием предварительного тестирования на X, конечно, зависит от того, в какой степени повторяющиеся

обследования характерны для той популяции, на которую хотят распространить выводы. Исследовательские интервью и процедуры измерения установок совсем нетипичны для сферы массовой коммуникации. Но когда мы хотим обобщить данные исследований в области педагогики, мы сталкиваемся с условиями, в которых тестирование — обычное явление, особенно если экспериментатор в качестве тестирования использует обычные опросы в классе. Но, возможно, нежелательное взаимодействие тестирования с X не возникает и тогда, когда экспериментальное тестирование O будет сходно с тем, которое всегда применяется в этой области. В тех же случаях, когда тестирование осуществляется с помощью необычных процедур или связано с маскировкой цели исследования, перестройкой восприятия и мышления, удивлением, стрессом и т.д., необходимо применять планы экспериментов с предварительно не тестируемыми группами.

Взаимодействие отбора (состава группы) и X. Так как в плане с предварительным и итоговым тестированием и контрольной группой контролируется влияние отбора только при интерпретации различий между экспериментальной и контрольной группами, остается опасность того, что полученные результаты валидны лишь для той популяции, из которой были составлены экспериментальная и контрольная группы. Такая возможность тем вероятней, чем больше трудностей мы встречаем при поиске испытуемых для эксперимента.

Рассмотрим более подробно, что означает взаимодействие фактора отбора (состава группы) и X. Если мы намерены проводить исследование в рамках одной школы, которая выразила готовность сотрудничать, то при случайном распределении испытуемых в экспериментальную и контрольную группы мы можем не заботиться об «основном эффекте» школы как таковой. Если средние результаты экспериментальной и контрольной групп лишь в равной мере улучшены посредством такого исследования, то этому не стоит придавать значения. Если же окажется, что особенности этой школы способствуют большей (по сравнению с той совокупностью, на которую собираются распространить данные) эффективности экспериментального курса, то это обстоятельство может быть серьезным. Надо убедиться в незначительности взаимодействия характеристик школы с экспериментальным воздействием. Некоторые экспериментальные переменные могут быть довольно чувствительны к характеристикам школы (взаимодействовать с ними), другие — нечувствительны. Такие взаимодействия могут возникать при использовании в исследовании школ со сходными средними IQ и могут отсутствовать, если значения IQ сильно отличаются друг от друга. Однако мы должны ожидать, что взаимодействия будут более вероятны тогда, когда школы сильно отличаются друг от друга по различным признакам, а не тогда, когда они сходны.

Часто значительные смещения выборки возникают из-за инерции экспериментаторов, которые не предусмотрели того, что администрация школ, которые представляют репрезентативную выборку, может отказаться от участия в эксперименте. Вследствие этого большинство педагогических исследований осуществляется в школах, где обучается непропорционально большой процент детей университетских преподавателей. Хотя достижение выборочной репрезентативности в полной мере невозможно и такой подход почти полностью отрицается во многих науках, репрезентативность может и должна считаться желательной в педагогических исследованиях. Одним из путей повышения репрезентативности является уменьшение числа испытуемых или классов из одной школы и увеличение числа школ и классов, в которых проводятся эксперименты. Конечно, мы никогда не достигнем того, чтобы эксперименты проводились на выборке, репрезентативной для всех школьных классов. Мы можем лишь шаг за шагом, путем проб и ошибок узнавать, сколь далеко мы можем экстраполировать, как широко обобщать внутренне валидные результаты.

Другие взаимодействия с X. Аналогичным образом взаимодействия X с другими факторами могут рассматриваться как угрозы внешней валидности. Дифференцированное выбывание является, видимо, следствием X, а не фактором, взаимодействующим с ним. Взаимодействие фактора инструментальной погрешности с X уже рассматривалось при обсуждении внутренней валидности, так как специфическое влияние этого фактора при предъявлении X бывает имитацией истинного влияния X (например, когда наблюдатель, который дает оценку, знаком с гипотезой и знает, кто из учащихся подвергался воздействию X). Угрозой внешней валидности является вероятность эффектов, характерных для конкретных средств (тестов, наблюдателей, счетчиков), используемых в этом исследовании. Регрессия не выступает как фактор, взаимодействующий с X.

Действие фактора **естественного развития** связано с проблемой отбора испытуемых: результаты могут оказаться действительными только для данного возраста, данной степени утомления и т.д. Взаимодействие X с фоном означало бы, что эффект наблюдается только в данных условиях эксперимента, и, хотя для этого фона результаты валидны, они не будут получены в других условиях. Успешное воспроизведение результатов в разное время и в различных условиях повышает нашу уверенность в проводимом обобщении, уменьшая вероятность взаимодействия фона и X.

Реакция испытуемых на эксперимент. Весьма серьезным источником нерепрезентативности обычного психологического эксперимента и даже педагогического исследования является пресловутая искусственность экспериментальной обстановки и знание испытуемого о том, что он участвует в эксперименте. Для испытуемого создается проблемная ситуация повышенной сложности, а процедуры и экспериментальное воздействие, на которые он реагирует, имеют не только значение стимулов, но играют роль ключевых признаков, используемых испытуемым с целью разгадать намерения экспериментатора. В такой обстановке для испытуемого характерно актерство, осторожность, установка на угадывание, он ощущает себя подопытным кроликом и т.д. Все это нетипично для нормальных школьных условий и, примешиваясь к эффекту X, сильно затрудняет обобщение результатов. Хотя подобные реакции неизбежны, внутренне валидные эксперименты этого типа все равно следует проводить. Но нужно предупреждать по возможности реакцию испытуемых на эксперимент.

Реакцию на эксперимент может вызвать любой аспект экспериментальной процедуры. Это относится и к процедуре предварительного тестирования (независимо от его содержания), и к его взаимодействию с X. Возможно, оказывает влияние и содержательная сторона процесса тестирования. Процедура рандомизации и распределения испытуемых по режимам эксперимента также может вызывать такую реакцию. Представьте себе, как подействует на класс то, что случайно отобранную половину учеников вдруг отправляют в отдельное помещение. Аналогичную реакцию может вызвать введение какого-либо экстраординарного X.

Эксперимент можно замаскировать следующими приемами. Если X-ы — это варианты обычных мероприятий в классе, проводимых в соответствующее время и предусмотренных школьной программой, то уже многого можно достигнуть, не оповещая о введении экспериментальных воздействий. Далее, О можно проводить под видом регулярных контрольных работ, и тогда второе требование будет соблюдено. Если X предусматривает индивидуальное общение с испытуемым, то для рандомизированного разбиения класса на эквивалентные подгруппы вовсе не обязательно разводить их по разным комнатам и т.д.

Решением проблемы во всех отношениях является использование в качестве единиц рандомизации целых классов, так что экспериментальная и

контрольная группы в результате состоят из случайным образом отобранных классов. Обычно вначале подбираются классы на основе таких признаков, как школа, учитель (когда один учитель ведет несколько классов), предмет, время дня, средний уровень умственного развития и т.д., а затем эти классы случайным образом распределяются по различным режимам эксперимента. Подобных исследований пока немного, но скоро они должны стать нормой.

2.6.3. Проверка статистической значимости эффекта

Проверка статистических гипотез — условие вывода об экспериментальной гипотезе, если исследователь претендует на оценку достоверности полученных данных.

а) **Распространенные ошибки статистической обработки.** Хотя план с предварительным и итоговым тестированием и контрольной группой является стандартным и широко распространенным, даже в этом случае часто применяются неправильные, неполные или не соответствующие задаче статистические процедуры. Применяя обычный критерий Стьюдента (t -критерий), многие подсчитывают два значения t : одно — для различия в показателях предварительного и итогового тестирования по экспериментальной группе, а другое — для аналогичных результатов по контрольной группе. Если первое значение статистически значимо, а второе нет, то делается вывод о наличии эффекта X без прямого статистического сравнения экспериментальной и контрольной групп. Во многих таких случаях применение более адекватного критерия не обнаружило бы значимого различия между группами (например, когда значение t -критерия чуть выше критического в экспериментальной и почти достигает его в контрольной группе).

б) **Использование приращений и ковариационный анализ.** Наиболее широко используемая и приемлемая процедура состоит в определении для каждой группы приращения показателей от предварительного тестирования к конечному и вычислении t -критерия для приращений, наблюдавшихся в экспериментальной и контрольной группах. В большинстве случаев, однако, следует предпочесть ковариационный анализ, в котором показатели предварительного тестирования берутся в качестве сопутствующих переменных и которому предшествует процедура формирования рандомизированных блоков или “распределение по уровням” показателей предварительного тестирования. Использование этого более точного анализа представляется весьма желательным.

в) **Статистики для рандомизированного распределения целых групп по режимам эксперимента.** Обычные статистики применимы только тогда, когда назначение определенного режима производится в случайном порядке для каждого ученика в отдельности. Когда же режим определяется сразу для целого класса, ошибка, вычисленная согласно этим моделям, будет занижена, так как при рандомизации использовались более крупные единицы и выборка наблюдений содержала меньшее число случайных событий. В качестве исходных данных здесь используются средние значения класса, и эффект X должен быть выделен из колебаний именно этих величин. Они же должны фигурировать в роли сопутствующих переменных при проведении ковариационного анализа.

3. СПОСОБЫ КОНТРОЛЯ ВЛИЯНИЯ АРТЕФАКТОВ

3.1. Формы экспериментального контроля

Отличие экспериментальных гипотез от других научных предположений, в частности, функционирующих как теоретические положения, заключается в том, что они имеют эмпирическое содержание, т.е. являются высказываниями об

опытных закономерностях (“переменная X воздействует на переменную Y так, что...”). **Экспериментальный контроль** обеспечивает достоверность вывода об эмпирически выявляемой зависимости с точки зрения отвержения других правдоподобных объяснений, так или иначе исходящих из “технических условий” ее получения, но недостоверность ее с точки зрения конкретизации теоретической интерпретации причинно-следственного характера. Спор психологических теорий между собой включает и ряд других уровней обсуждения, кроме определения весомости тех или иных интерпретаций. Это и отношение к исходным теоретическим понятиям — конструктам — с точки зрения возможностей их конкретизации, и вопросы о сути психологической причинности, или о понимании принципа действия психологических законов (как детерминистских, вероятностных и т.д.) и многое другое.

С точки зрения контроля переменных следует выделить три его вида, связанные с возможностями: 1) планирования как построения экспериментальных схем, 2) задания разных типов переменных и соответственно разных типов экспериментов, 3) снятия угроз валидному выводу со стороны других факторов, необходимо присутствующих при реализации эксперимента и подлежащих так называемому первичному контролю, т.е. контролю до выбора конкретных схем или безотносительно к экспериментальным планам.

Контроль уровней переменных выводит исследователя на решение вопросов об операционализации переменных теми или иными методическими условиями и о выборе схем, или планов как интраиндивидуальных или межгрупповых сравнений.

1. Контроль посредством экспериментальных схем. План эксперимента задает форму контроля независимой переменной — предъявляются ее разные условия одному и тому же испытуемому или группам испытуемых — и одновременно условия фиксации первичных показателей зависимой переменной. При последующей обработке данных на основе первичных показателей могут быть предложены вторичные показатели, разрабатываемые с целью укрупнения единиц анализа эмпирических данных. Так, переход от матриц переменных к матрицам их интеркорреляций позволяет осуществлять те или иные процедуры сокращения размерности данных. Иногда сложность процедур обработки данных рассматривается как критерий того, что имело место экспериментальное исследование. Так, процедуры факторного анализа, использованные в контексте снижения размерности данных, могут быть ошибочно названы экспериментальными.

Управление независимой переменной означает обычно и контроль **побочных, смешивающихся или сопутствующих переменных**. Смешения независимой переменной с ними представляют угрозы валидности эксперимента. Рассмотрение в учебниках факторов задачи, последовательности задания уровней независимой переменной, времени, состава групп и других в качестве управляемых переменных не обязательно переводит их в статус независимых. То есть экспериментальный контроль включает не только управление независимой переменной, но и формы контроля других переменных. Такие формы контроля, как **контроль несистематической изменчивости, контроль систематических смещений** и разных видов **сопутствующего смешения**, подробно представлены в учебнике Р. Готтсданкера, и мы на них специально останавливаться не будем. Важно, однако, отметить, что не следует смешивать решение вопросов: 1) имел ли место результат действия независимой переменной, т.е. вывод об экспериментальном эффекте, 2) имела ли место именно подразумеваемая в гипотезе каузальная зависимость, т.е. вывод об исследуемой зависимости. Экспериментальные схемы лишь частично служат основанием ответов на эти вопросы.

Обобщение относительно вида установленной зависимости, т.е. решение

вопроса о том, что установлена именно представленная в гипотезе связь, а не какая-то другая, зависит от оценки внутренней валидности эксперимента. Контроль угроз внутренней валидности со стороны тех или иных источников смещений или ненадежности данных рассматривается в отношении к конкретным экспериментальным схемам. Введение уровней вторичной независимой переменной для уточнения эффекта действия первичной независимой переменной означает переход к факторным схемам и служит как улучшению внутренней, так и повышению внешней валидности экспериментов. То есть контроль посредством экспериментальных схем охватывает и управление независимой переменной, и учет всех “третьих” переменных как оснований предпочтений тех или иных планов.

Гипотезы с одним отношением, т.е. формулирующие каузальную зависимость между одной независимой переменной и показателем зависимой переменной, проверяются при использовании так называемых истинных экспериментальных планов (по Кэмпбеллу). Но в любом учебнике по планированию обнаруживается пробел, который можно охарактеризовать как нормативы содержательного планирования. Нет рецептов для формулирования “правильных” гипотез, но есть возможность оценить степень учета исследователем тех азов, без которых гипотеза не может считаться достаточно обоснованной и имеющей вид собственно экспериментальной.

2. Контроль типа переменных. Определение типа переменных с точки зрения соответствия их реальным условиям, рассматриваемым теоретическим конструктам или мысленно моделируемым схемам взаимосвязей между переменными, позволяет исследователю прогнозировать возможности обобщений как выводов об экспериментальной зависимости за пределами экспериментальной ситуации. Цель любого психологического эксперимента — обобщение за его пределами. Перенос рассматриваемой зависимости на другие условия, другие виды деятельности человека, другие ситуации и другие виды переменных означает определенные требования к внешней валидности эксперимента. Обобщение на другие группы испытуемых связано с контролем того обстоятельства, насколько хорошо представленные в эксперименте выборки испытуемых репрезентируют, или представляют, те более широкие группы людей или те популяции (группы людей, отвечающие определенному критерию их сходства), на которые осуществляется перенос зависимости. Способы отбора испытуемых из популяции или подбора испытуемых в группы определяют допустимые границы обобщений. Как показывает Р.Готтсданкер, применительно к межгрупповым экспериментальным схемам следует учитывать связь внешней и внутренней валидности исследования именно благодаря возможностям использованных схем отбора или подбора в группы.

Обобщение на те уровни независимой переменной, которые реально не использовались в эксперименте, также может иметь отношение к оценке его внешней валидности. Так, обобщение на другие популяции можно представить как обобщение на другие уровни независимой переменной, если критерий отличия групп лежал в основе межгрупповых сравнений. Однако более тесно проблемы переноса закономерностей на другие ситуации, популяции или виды деятельности связаны с учетом уровней дополнительных переменных.

В зависимости от статуса тех или иных переменных по отношению к экспериментальной гипотезе и возможностей управления ими экспериментатором могут быть использованы такие формы контроля, как стабилизация переменных или их максимальное варьирование. Дополнительные переменные, которые содержательно входят в формулировку психологической гипотезы как условия проявления каузальной связи между независимой и зависимой переменными, обычно поддерживаются на заданном идентичном уровне или начинают выступать в качестве второго фактора (кроме независимой переменной), учет изменений

которого необходим для заключения о каузальном выводе. Так возникают факторные планы для проверки гипотез с одним отношением, т.е. между одной независимой переменной и зависимой переменной.

3. **“Эффект экспериментатора” и первичный контроль.** Психологический эксперимент предполагает взаимодействие исследователя с испытуемым. Общий контекст и мотивация участия в психологическом эксперименте могут быть разными. Обычной для научных исследований является ситуация призыва “добровольцев” (по объявлению, личному приглашению и другим образом). Несколько иными являются ситуация “клиента”, когда сам испытуемый ищет возможность участия в психологическом эксперименте (и с целью самообследования, и при познавательном интересе), или ситуация комплексной экспертизы в условиях психиатрической клиники. В последнем случае очевидна значимость психологических данных как данных обследования в общей жизненной ситуации пациента. При недостаточном уровне интеллектуальной, личностной или поведенческой критичности человека его включенность в выполнение экспериментальных заданий будет иной, чем в обычных условиях актуализации “мотивации экспертизы”. Вынужденное участие характеризует, например, целые группы испытуемых, если проводится сравнение разных методов обучения и используются реальные учебные группы.

Сам экспериментатор в разной степени включен в проведение эксперимента и общение с испытуемым. Так, использование проективных методик, как и других методик толкования, требует от экспериментатора большей личностной включенности во взаимодействие с испытуемым, чем, например, заполнение опросника MMPI или интеллектуальных тестов. Ситуация психофизического эксперимента вообще ограничивает присутствие экспериментатора заданием инструкции и управлением стимуляцией, в то время как в исследованиях мышления взаимодействие с испытуемым необходимо для организации испытуемому обратной связи, коррекции его ошибок в предположениях или рассуждениях по ходу решения задачи. То есть деятельность экспериментатора включает обеспечение сбора данных. Экспериментатор объясняет цель действий и во многом обуславливает мотивационный уровень принятия задачи испытуемым.

Как отмечал Г. Тэшфелл применительно к социально-психологическому исследованию, эксперимент проходит не в социальном вакууме: испытуемый активно строит предвосхищения относительно ожидаемых от него стратегий поведения или желательных ответов. Испытуемый обязательно выдвигает гипотезы об ожидаемых от него результатах. И чем более неопределенной является для испытуемого ситуация, тем более активно он строит ее образ и отмечает в вербальных и невербальных проявлениях экспериментатора все те признаки, которые сигнализируют ему об успешности или неуспешности его деятельности. Эффекты взаимодействий экспериментатора и испытуемого включают, таким образом, как “эффект экспериментатора”, так и “эффект испытуемого”.

Множество исследований показало влияние таких характеристик экспериментатора на результаты эксперимента, как его пол, возраст, престиж в глазах испытуемого, привлекательность, доброжелательность, враждебность, тревожность, предварительное знакомство с испытуемым и т.д. Какие-то закономерности были выявлены в качестве довольно поверхностных (например, возрастание числа рисуемых усатых лиц в присутствии усатого экспериментатора), какие-то могли быть проинтерпретированы более глубокими взаимосвязями переменных. Так, в случае трудных задач мужчины проявляют большую включенность и старательность в решении, если экспериментатор — женщина; женщины добиваются более высокой результативности при тестировании детей, чем мужчины. Было продемонстрировано, что влияние характерологических черт

экспериментатора или взаимодействия черт испытуемого и экспериментатора проявляется в основном там, где может быть показана содержательная зависимость между обсуждаемым свойством и базисным процессом.

Розенталь назвал “эффектом Пигмалиона” непроизвольное способствование экспериментатора в проявлении именно тех закономерностей, в истинности которых он убежден. Этот эффект наведения испытуемого на нужную стратегию был подтвержден исследованиями невербальной коммуникации (например, улыбка экспериментатора предстала в качестве средства позитивного подкрепления ожидаемых форм поведения) и в так называемых экспериментах на “вербальное обусловливание” (в них неопределенные “гм-м” экспериментатора значимо изменяли частоту подкрепленного таким образом ответа испытуемого).

3.2. Контроль за выводом в экспериментальном исследовании

Оценка обоснованности или достоверности выводов, осуществляемых на основе проведенных экспериментов, включает ряд линий рассуждения исследователя, где этапы нормативного логического умозаключения сопутствуют необходимым “прорывам” в обобщениях. **Достоверный вывод** — вывод об экспериментальной гипотезе, который основан на обобщении именно представленной в гипотезе связи.

Возможности ошибиться в уровне обобщения относительно проверяемой психологической гипотезы, в принятии решения о полученной зависимости или проявить логическую некомпетентность остаются всегда, в том числе и в случае получения достоверных экспериментальных результатов. Можно различать два основных пути к неверным, или артефактным, выводам: 1) проведение невалидного эксперимента и принятие ошибочных решений о результатах действия переменных (т.е. о полученных экспериментальных эффектах); 2) ошибки в контроле за выводом, т.е. неверные умозаключения или неверные обобщения. Данный раздел посвящен второй проблеме, и в ней намечены те основные пункты контроля за выводом, учет которых позволяет избежать ошибок на этом последнем этапе исследования: ведь любой эксперимент проводится с целью обобщений за его пределами.

Иногда оценку достоверности выводов связывают с возможностью сделать на основе проведенного эксперимента такие выводы, которые являлись бы обоснованными в случае приближения этого эксперимента к безупречному. Более широкая трактовка достоверности выводов включает оценку их правильности с точки зрения проявленной исследователем логической компетентности и правомерности сделанных обобщений. Однако первым пунктом оценки достоверности выводов является все же оценка правильности принятия решений об экспериментальном эффекте. Не затрагивая проблем статистических выводов, связанных с принятием решения о значимости полученных в эксперименте различий или связей выборочных значений переменных, укажем общие нормативы возможных рассуждений об экспериментальной гипотезе. Это рассуждения, реализуемые исследователем, во-первых, при переходе между различными уровнями проверяемых гипотез: статистических и содержательных (т.е. психологических), экспериментальных и теоретических. Во-вторых, при сопоставлении сформулированного в проверяемой научной гипотезе и конкурирующих психологических объяснений. В-третьих, это “ненормативные”, т.е. не облачаемые в форму логических требований, рассуждения о допустимых уровнях обобщений об исследуемой психологической реальности. Обобщения эти касаются утверждений о переносе выявленных экспериментально закономерностей на другие виды деятельности, ситуации, популяции и т.д.

Признание ограниченности в возможностях обобщения той или иной

закономерности связано не только с оценкой уровней дополнительных переменных. Часто это также решение вопросов о содержательной связи типа психологической регуляции в эксперименте и в ряде жизненных или профессиональных ситуаций, отличаемых по предметной направленности деятельности, о репрезентативности полученных “усредненных” зависимостей для конкретных людей или выборок, о соотношении случайного и закономерного применительно к анализируемому индивидуальному случаю или к группам людей.

Наконец, правильность выводов связана с оправданным переходом от логики экспериментального доказательства “изнутри”, т.е. по отношению к проведенному исследованию, к логике доказательства “извне”, т.е. путем соотнесения обоснованности сделанных обобщений с другими (возможными или имеющимися) теоретическими позициями или доводами “здравого смысла”. Рассмотрение доводов только в пользу представленного в экспериментальной гипотезе психологического объяснения и исключение из обсуждения результатов сопоставления его с другими допустимыми интерпретациями противоречит нормативам экспериментального рассуждения.

3.3. Вывод о гипотезе на основе оценки валидности эксперимента

Проведение любого эксперимента служит цели проверки истинности сформулированного высказывания о каузальной зависимости только с точки зрения соответствия или несоответствия ему эмпирически установленной связи между независимой переменной и зависимой переменной. Интерпретационная часть гипотезы (т.е. собственно психологическое объяснение) не оценивается в самом эксперименте с точки зрения ее истинности. Привлекаемые для описания отношения между независимой переменной и зависимой переменной гипотетические конструкции (понятия, которые соответствуют теоретическим представлениям и функционируют в гипотезе) живут в мире теорий, и оценка адекватности тех или иных теорий включает не только отношение теории к миру эмпирии, но и многие другие компоненты.

Контроль за выводом предполагает более формальный аспект оценки приемлемости проверяемого утверждения с точки зрения анализа логически возможных соотнесений эмпирического результата и оценки валидности как реализованных форм экспериментального контроля. Если в исследовании был получен ожидаемый в соответствии со сформулированной психологической закономерностью экспериментальный эффект, это еще не позволяет автоматически принять или отвергнуть экспериментальную гипотезу или контргипотезу. Вторым условием, влияющим на достоверность вывода, является оценка валидности психологического эксперимента, включая реальные “технические” условия его проведения.

В случае высокой внутренней и операциональной валидности эксперимента при получении данных, ожидаемых в соответствии с экспериментальной гипотезой, считается, что она выдержала проверку и принимается. Однако в силу принципа асимметрии выводов исследователь не может на основании получения ожидаемого экспериментального эффекта считать доказанной ту теорию, исходя из которой эксплицируется экспериментальная гипотеза. То есть вывод звучит примерно следующим образом: полученные экспериментальные данные не противоречат сформулированной экспериментальной гипотезе и, соответственно, выдвинутой теории. То, что психологическая гипотеза выдержала проверку опытным путем, не свидетельствует о “правильности” предполагаемого ею типа психологического объяснения. Понятно, что такое заключение не затрагивает также оценки экспериментального эффекта “извне”, т.е. с точки зрения так называемых конкурирующих объяснений.

Получение эффекта в пользу **нуль-гипотезы** (т.е. статистической гипотезы, которая в экспериментальном исследовании противоположна по направленности утверждения об ожидаемых изменениях зависимой переменной в экспериментальной гипотезе) в условиях высоковалидного эксперимента обозначает возможность строгого вывода в пользу контргипотезы.

Это путь отвержения теоретических положений на основе их экспериментальной проверки, который и стоит за логикой “асимметрии вывода”. То есть это и является случаем строгого отвержения теорий на основе получения “негативных” результатов. Собственно, экспериментальный метод и считается самым строгим путем эмпирического опробования теорий, поскольку он дает возможность отвержения “неверных” теорий, как не соответствующих реальности. Но в логике сопоставления разных психологических гипотез, имеющих разную эмпирическую подкрепленность, прилагаются разные критерии к достоверности такого рода отрицательных выводов. Так, отдельный отрицательный результат сам по себе не влечет отказа от содержательной гипотезы. Часто должно произойти некоторое накопление таких “отрицаний”, чтобы данные в пользу контргипотезы действительно были приняты в качестве достаточных доводов в пользу того, чтобы отвергнуть проверяемую теоретическую гипотезу. Ведь в любом эмпирическом исследовании можно пытаться искать отклонения его от безупречного образца, на основе чего связывать отрицательный эффект с артефактами проведения.

3.4. Вывод о подкреплённости теории экспериментальными фактами

Для **эксперимента с научными целями** обобщение вида полученной каузальной зависимости, если имел место **лабораторный эксперимент**, проходит в три этапа. На первом делается вывод о виде зависимости, на втором — об обоснованности используемой теоретической модели. И уже затем реализуется путь от теории к реальности, т.е. рассматривается соответствие психологическая реальность — ее объяснение с помощью проверенной в лабораторном эксперименте научной гипотезы. Подкрепленность теории экспериментальными данными позволяет распространять предполагаемые в ней обобщения самого высокого уровня на все те виды психологической реальности, которые могут рассматриваться в соответствующем содержательном контексте.

Для **экспериментов с практическими целями** логика вывода иногда упрощается. Обобщения из экспериментов, “дублирующих” или “улучшающих” реальный мир, прямо соотносят установленные психологические закономерности с возможностью их проявления в тех или иных ситуациях или видах деятельности, применительно только к участвовавшему в эксперименте человеку или к другим людям. Так, если проверялась гипотеза о большей эффективности нового метода обучения по сравнению с каким-то традиционным, то получение данных “за” экспериментальную гипотезу прямо интерпретируется как доказательство того, что новый метод “лучше” традиционного. Нередко делается и следующий необоснованный шаг: автоматически считается “доказанным” и то теоретическое звено обобщений, которое легло в основу разработки этого нового метода обучения.

При этом исследователь выходит за рамки принципа фальсификации гипотез. Этот принцип гласит, что при неотвержении экспериментальной гипотезы еще не может считаться доказанной истинность теории, так как всегда можно ожидать новых опровергающих фактов и вновь сформулированных объяснений (исходя из других теорий). В этом смысле любая гипотеза всегда открыта для дальнейшей проверки.

Применительно к “полевым” экспериментальным исследованиям,

характеризующимся высокой **конструктивной валидностью**, вопрос о широте допустимых обобщений в случае “позитивных” результатов может решаться положительно, даже если соответствующие теоретические гипотезы были разработаны для других условий или популяций. Однако вариация условий и популяций не должна превышать каких-то пределов, за которыми использование тех же самых конструктов уже обесмысливается, поскольку главенствующую роль начинают играть другие факторы. Доказанность, например, многих социально-психологических гипотез может пониматься именно как соответствие их определенным условиям социума; и невозможность их перенесения для интерпретации в другие условия не означает снижения их эмпирической подкрепленности.

Проблема доказанности теоретических гипотез, т.е. обобщений более высокого уровня, чем постулируемая экспериментальная гипотеза, связывается со следующими моментами. Во-первых, это методологические споры на уровне собственно рационального знания, т.е. сопоставления систем теоретических построений. Во-вторых, это оценка теории с точки зрения развития целостной исследовательской программы, включающей анализ системы экспериментов в рамках той или иной школы. В-третьих, это анализ количества значимых результатов при проверке данной гипотезы в программе исследований.

3.5. Проблема возникновения новых гипотез

Описав общие черты экспериментального метода, мы не затронули вопроса “откуда берутся экспериментальные гипотезы”. Их выдвигает исследователь, решая научные задачи или желая достигнуть практических целей средствами научного познания. Поскольку не существует правил, в соответствии с которыми исследователь должен был бы выводить новые закономерности на основании новых опытных данных, то он должен формулировать гипотезы. Новые гипотезы выдвигаются либо для объяснения вновь наблюдаемых, подчас неожиданных явлений, либо для устранения оставшихся незамеченными противоречий в ранее существовавших концепциях. “Новые” гипотезы в этом утверждении как бы противопоставляются “старым”, т.е. выдержавшим проверки опытным путем. В психологической науке новые гипотезы возникают часто в системах новых интерпретаций, нового понимания предмета изучения и сосуществуют как “современники” в рамках разных психологических школ.

Кроме того, существует проблема разноуровневых обобщений: между объяснением той или иной зависимости на уровне обсуждения результатов действия переменных и на уровне каузального описания полученной эмпирической зависимости как закономерности всегда остается некий пробел, требующий от исследователя “прорыва в обобщении”. Этот пробел заполняется исследователем, во-первых, привнесением в психологическое объяснение определенных представлений о том, как действуют психологические законы или психологические механизмы регуляции деятельности, поведения, общения. Во-вторых, этот пробел заполняется недостающими звеньями между описанием зависимости и обобщением. Определить достаточную для понимания закономерности полноту обобщения достаточно сложно, и в психологии существует масса понятий, конкретизируемых в разной степени полно. Например, понятия “материнская депривация”, “интеллектуальная стратегия” или “уровень притязаний” могут использоваться в контексте построения разных гипотетических объяснений наблюдаемых зависимостей и сами раскрываться в разной степени полно со стороны входящих в них переменных. Новые объяснения и новые гипотезы в психологии могут возникать именно на пути все более глубокого и полного охвата тех эмпирических зависимостей, которые сами по себе не являются новыми, но допускают переинтерпретацию в результате развития самих

психологических понятий.

Наконец, новые гипотезы возникают, когда старая проблема вводится в новый контекст обсуждения, начинает анализироваться в комплексе других, подчас и непсихологических, подходов. В рамках такой области исследований, как “инженерия знаний”, сосуществуют разные подходы по отношению к решению ряда вопросов, связанных с конкретизацией компонента интеллектуальных стратегий, т.е. затрагивающих такую старую область, как психология мышления. Психологический анализ использования информационных технологий позволяет формулировать новые гипотезы относительно психологических механизмов опосредствования интеллектуальной деятельности человека, по-новому взглянуть на возможности развития мышления человека, в этой же области знаний разрабатываются модели и в рамках так называемой компьютерной метафоры, реализующей один из вариантов редукционизма в психологических объяснениях. Таким образом, новая гипотеза не всегда означает движение на пути развития психологического знания. Критерий “более новая” не может служить синонимом “более содержательной” гипотезы.

Итак, эксперименты проводятся для опытной проверки гипотезы. Подход к психологической интерпретации в целях обобщения полученного знания включает оценку того, насколько хорошо представленные в гипотезе гипотетические конструкты позволяют исследователю осуществлять обоснованные переходы между разными уровнями обобщения выявленной зависимости и используемых психологических представлений. Эта обоснованность задается на этапах содержательного планирования и вновь анализируется в обсуждении полученных результатов. Эти содержательные проблемы могут, однако, решаться при разной степени рефлексии продланного пути при доказательстве психологической гипотезы. Логика организации вывода должна включать при реализации экспериментального метода следующие компоненты:

1) гипотетико-дедуктивный вывод об эмпирической реальности с указанной асимметрией вывода о научной гипотезе (“отбросить предположение, если факты ему противоречат”); **дедуктивный вывод** — умозаключение, реализующее путь обобщения от общего к частному;

2) построение планов, или экспериментальных схем, в рамках индуктивного вывода о результате действия экспериментального фактора и возможности причинного объяснения изменения зависимой переменной; **индуктивный вывод** — вывод о том, можно ли рассматривать управляемую независимую переменную в качестве причинно действующих условий, вызывающих экспериментальный эффект;

3) вывод об экспериментальной гипотезе на основе анализа полученного эффекта путем соотнесения результата с оценкой валидности эксперимента;

4) обоснование содержательных оснований обобщений зависимости за пределами эксперимента.

3.6. Проблема артефактных выводов

Ошибки на пути к конечным обобщениям приводят к формулировкам неверных, или артефактных, выводов. **Артефактные выводы** — совокупность неверных умозаключений или обобщений, приводящих к недостоверным обобщениям относительно вида экспериментальной или корреляционной зависимости.

Артефактные выводы могут быть следующего характера:

1) неверный вывод из-за неверных статистических решений. Например, исследователь может “проглядеть”, что нужно отвергнуть как экспериментальную, так и контргипотезу и необходим поиск так называемой третьей конкурирующей

гипотезы;

2) неверный вывод о действии независимой переменной из-за невнимания к строгой оценке валидности эксперимента, за которым кроется ошибка принять в качестве положительного “подтверждения” экспериментальной гипотезы артефактный результат или возможность “проглядеть” в эксперименте истинную зависимость;

3) неверные обобщения из-за неучета существенных дополнительных переменных или ошибок в понимании соотношения теоретического утверждения и эмпирического базисного высказывания в экспериментальной гипотезе;

4) замена или искажение рассмотренных нормативов экспериментального вывода оценочными суждениями, апелляцией к авторитету и другими неявными “уступками” логики рассуждения содержательно необоснованным критериям. Эти ошибки связаны с ценностными отношениями к содержанию проблемы или выводам и недостаточной критичностью к тому, что в исследовании оказалось упущенным.

На последней группе “ошибок в выводах” остановимся специально, поскольку они грозят содержательным обобщениям и в тех исследованиях, которые были хорошо спланированными. В специальной литературе обсуждается ряд следующих ошибок, наиболее часто встречающихся в психологических исследованиях.

А. “Недостающие звенья” в причинных объяснениях

Каузальные факторы, гипотетически ответственные за возникновение того или иного феномена или обуславливающие связи переменных, могут описываться без достаточного детального рассмотрения, что приводит к возникновению “недостающего звена” между объяснением и обобщением.

Например, делается вывод о том, что “материнская депривация” у детей приводит к “расстройствам адаптации”. Однако фактор “материнской депривации” не является элементарным, а может быть рассмотрен в качестве множества факторов, каждый из которых оказывает свое непосредственное влияние на возникновение расстройств адаптации.

Другим примером мог бы служить анализ взаимосвязей между переменными “агрессивность” и “предпочтение просмотра телепередач с агрессивным содержанием”. Гипотеза о связи агрессивности и просмотра телепередач с агрессивным содержанием проверялась в лонгитюдном корреляционном исследовании, где переменные были измерены для одних и тех же субъектов в возрасте 8-9 лет и затем 18-19 лет. Если бы не был применен план перекрестно-отсроченных корреляций, т.е. из анализа были бы изъяты данные об изменениях во времени (с промежутком в десять лет) связи между измерениями “А1 и В2” и “А2 и В1”, то было бы утеряно наиболее важное основание выводов о взаимосвязи А и В переменных.

Б. Неправомерное выделение основной причины

При формулировании обобщений, включающих объяснения эмпирических зависимостей, какая-либо частная причина может быть принята за основную и исчерпывающую.

Например, при объяснении связи материнской депривации с расстройствами адаптации в качестве основной и исчерпывающей причины последних может быть предложен частный фактор, являющийся одним из многих, входящих в понятие материнской депривации, например отсутствие физического контакта с матерью. При выводах о связи показателей IQ с уровнем полученного образования те факторы, которые стоят за “попаданием” ребенка в более

привилегированное учебное заведение, могут то выдвигаться на первое место (социальный и образовательный ценз семьи, этническая принадлежность или другие переменные), то занимать место третьестепенных по сравнению с рассматриваемыми автором (например, IQ родителей, креативность и т.д.). Одни и те же переменные могут занимать разноуровневые положения — измеряемых или латентных переменных. Понятно, что в каждом конкретном исследовании веер возможных объяснительных гипотез ограничен числом охваченных переменных и предположением авторов о главенствующей роли одних и подчиненной — других переменных.

Претензии на полное или исчерпывающее каузальное объяснение всегда можно постараться “вынести за скобки”, чтобы оценить подкрепленность авторской гипотезы другими конкурирующими объяснениями.

В. Подмена одного высказывания другим

Исследователь может некритично предположить, что одна зафиксированная им ситуация гарантирует содержание другой, и сообщить только о последней.

Например, может быть предложено сообщение: “Никто из испытуемых не читает журнал “Плейбой”, в то время как более корректным сообщением было бы: “Никто из испытуемых не признался, что читает журнал “Плейбой”. Или же автор, используя ответы испытуемых типа “часто”, “не очень часто”, “редко” и т.п., не пытается найти критерий оценки (или размерности) субъективной шкалы испытуемого, что приводит к неразличению высказываний: “Испытуемый поступает таким образом часто” и “Испытуемый считает, что он поступает таким образом часто”.

При качественной интерпретации данных, например при использовании в психологии мышления протоколов “рассуждения вслух”, это оборачивается проблемой неразличения двух уровней их анализа: объяснения с позиций самого испытуемого, как он осуществляет решение задачи, и объяснения с позиции экспериментатора, использующего некоторую схему членения высказываний в вербальном протоколе. Если при оценке личностью своих качеств психолог обычно сохраняет критичную позицию, выражающуюся в различении высказываний: “Он говорит, что он такой” и “Он такой, каким он себя описывает”, то в других областях психологической реальности такой критичной исследовательской позиции может не прочитываться, либо же психолог сознательно стремится ее избежать.

Г. Ошибка ценностных суждений

Содержательные выводы о подкрепленности психологической гипотезы опытными данными вводятся в контекст ценностных отношений так, что желаемая оценка значимости эффекта с точки зрения социального или этического критерия связывается с проблемой “доказанности” психологических посылок автора.

Например, оценка преимуществ нового метода обучения может связываться с тем, что этот метод способствует воспитанию интеллектуальной элиты. Тогда в зависимости от конкретных социально-политических условий в выводы об эффективности метода могут включаться ценностные пожелания типа “метод плох, потому что он направлен на воспитание элиты, а нам нужен такой метод, который бы был адресован всем” либо “метод хорош, потому что он направлен на воспитание элиты, а именно с воспитанием интеллектуальной элиты у нас в обществе проблемы”. И оба таких утверждения могут быть в равной степени необоснованными, поскольку трудно желаемый прогноз о возможностях метода соотносить с внешним критерием социальной оценки интеллектуального развития людей.

Другой вопрос — это вопрос о том, что те или иные психологические результаты имеют разный ценностный резонанс с точки зрения возможности и желательности их использования на практике. Но оценка актуальности тех или иных психологических разработок не может выступать критерием истинности проверяемых в них психологических гипотез. Подмена же выводов типа “что получено” (в результате исследования) на типы “что ценного в том, что получено, для практики” или “что делать” и есть механизм возникновения ошибок ценностных суждений.

Д. Редукционизм выводов

В психологии редукционизм рассматривается как подмена психологических объяснений непсихологическими или как поиск объяснения определенного числа различных явлений посредством сведения их к одному принципу объяснения. Он может проявляться как подведение эмпирических выводов под интерпретационные схемы, лежащие зачастую в сфере других наук или других областей знания (культурологии, социологии, физиологии и т.д.), или как организация выводов в рамках принятой объяснительной парадигмы тех или иных психологических школ. Тогда, говоря о редукционизме, другие исследователи подчеркивают неадекватность используемых понятий психологической реальности или схем получения эмпирических данных предмету исследования.

На основе таких вариантов редукционизма при обсуждении авторами заявленных психологических гипотез могут делаться выводы, упрощенно представляющие исследуемые базисные процессы или связи между переменными. Редукционизм социологического, психофизиологического или другого толка — это методологические выводы о структуре психологических объяснений.

Существенно, что принимаемые научным сообществом критерии “правильных” выводов изменяются во времени, со сменой исследовательских парадигм и стилей мышления, а также социальных установок авторов. Относительность понятия достоверных выводов тем более понятна, если учитывать, что одни и те же эмпирические результаты могут отвечать на разные вопросы и вводиться в разные контексты других данных и других объяснительных схем.

Е. Неправомерные апелляции к авторитету

Утверждение о том, что некоторый эксперт (группа экспертов) имеет некоторое мнение, может предлагаться в качестве достаточного или значимого довода для подтверждения этого гипотетического суждения. Ошибка суждения здесь имеет место в том случае, если вместо прошедших проверку на истинность доводов для оценки содержания гипотезы привлекается ссылка на мнение авторитета. В подобном контексте будет более корректно оценивать не авторитетность эксперта, а обоснованность содержательных аргументов в пользу данного мнения.

Для содержательной оценки выводов из исследования неважными являются регалии ученого — важна аргументация в данном конкретном случае. Ведь авторитет, т.е. признанный специалист в этой области знаний, может сделать вывод такого рода: “Если Вы согласны со мной по 9 из 10 отстаиваемых мною утверждений, то Вы должны согласиться и с десятым”. Этот пример приводится, в частности, австралийскими психологами при рассмотрении доводов, выдвигаемых Г. Айзенком для защиты гипотезы о наследственных факторах интеллекта. Было не раз отмечено, что апелляция к мнению экспертов, которые якобы заведомо лучше разбираются в проблеме, чем другие исследователи, не

попавшие в группу экспертов, возникает всякий раз, когда очевидны ценностные или социально-политические установки общества в подходе к тем или иным проблемам. Это происходит также в тех случаях, когда научный вес, научный или личностный авторитет исследователя столь высок, что обсуждение отстаиваемых им гипотез выводится за рамки требований к доказанности суждений на уровне эмпирических или логических доводов.

Ж. Апелляции к факту

Ошибка суждения заключается в следующем. В качестве достаточного или значимого довода предлагается утверждение о том, что доказательство, подтверждающее какую-либо точку зрения, является установленным или общепринятым фактом.

Довод необходимо аргументировать, так как подобное доказательство с большой вероятностью будет являться спорным. Его спорность будет особенно велика, если в качестве “факта” предлагается описание какого-либо феномена в терминах какой-либо конкретной теории. Подобное описание может не считаться установленным или общепринятым для сторонника другой теории. Наконец, апелляция к феноменальной данности психологического события еще не есть довод, поскольку не менее важен вопрос, ответ на который связывается с такой апелляцией. Система рассуждений, в которую включается описание психологической реальности, и рождает “психологический факт”.

Принятие решения о том, что тот или иной психологический факт имел место, может основываться на очень разных системах доказательств. Так, для использования метода наблюдения характерна проблема ограничения интерпретации, которая всегда включается в описание наблюдаемых явлений. Для экспериментального метода принятие решений о виде устанавливаемого факта предполагает гораздо более строгий контроль пути от сбора данных до утверждений о полученных зависимостях (как психологических фактах). Но какой бы метод ни имелся в виду, всегда в выводах представлен тот уровень обобщений, в рамках которого только и имеет смысл говорить об установленных фактах.

ЗАДАНИЯ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

- 1. Составьте логическую схему базы знаний по теме курса.**

2. Раскройте факторы, угрожающие внутренней или внешней валидности (при выполнении этого задания студент имеет право выбора).

3. Выявите ошибки в предложениях:

а) Психологические выводы — предположения, сформулированные относительно психологической реальности в рамках использования тех или иных психологических представлений.

б) Безупречный эксперимент — мысленный образец для оценки реально проводимых экспериментов, в котором число проб или число испытуемых оцениваются с точки зрения охвата всех источников вариабельности переменных.

в) Субъективный процесс — реконструируемый в теории или в гипотетических конструктах процесс, определяющий показатели зависимой переменной в психологическом эксперименте.

4. Дайте соответствующее (правильное) определение понятиям.

Понятия	Определения
Зависимая переменная	управляемое экспериментальное воздействие (фактор), т.е. активно изменяемая и контролируемая исследователем переменная, представленная на двух или более уровнях (качественных или количественных).
Независимая переменная	измеряемая в эксперименте переменная, изменения которой причинно обусловлены действием независимой переменной.
Валидность исследования	план организации многоуровневого эксперимента, в котором вторым экспериментальным фактором выступает порядок предъявления разных условий независимой переменной.
Латинский квадрат	оценка проведенного эмпирического исследования с точки зрения правильности его организации и возможности считать полученные результаты и сделанные выводы достоверными.

5. Перечислите известные Вам виды экспериментов (по форме их проведения).

- а)
- б)
- в)
- г)
- д)
- е)
- ж)
- з)
- и)
- к)

ТРЕНИНГ УМЕНИЙ

1. Пример выполнения упражнений тренинга на умение №1.

Задание

В экспериментальном исследовании принимали участие две группы испытуемых, мужчин и женщин, а также два экспериментатора, мужчина и женщина. Определите наличие внешней переменной и примените контроль внешней переменной.

Решение

Предварительно заполните таблицу, подобрав к каждому алгоритму конкретное соответствие из данной ситуации.

№ п/п	Алгоритмы	Конкретное соответствие данной ситуации предложенному алгоритму
1	Определение влияния внешней переменной на зависимую переменную.	Влияние переменной состава групп и состава экспериментаторов.
2	Выявление возможности контролировать внешнюю переменную.	Влияние внешней переменной можно контролировать с помощью техники балансировки.
3	Применение техники контроля внешней переменной.	Применение техники балансировки: Группа 1 (эксперимент) Мужчины — экспериментатор мужчина Мужчины — экспериментатор женщина Женщины — экспериментатор мужчина Женщины — экспериментатор женщина Группа 2 (контроль) Мужчины — экспериментатор мужчина Мужчины — экспериментатор женщина Женщины — экспериментатор мужчина Женщины — экспериментатор женщина

Решите самостоятельно следующие ситуации:

Ситуация 1

В экспериментальном исследовании выявлялись психофизиологические свойства слухового анализатора испытуемых. Внешние звуки, шумы и вибрации могут повлиять на результаты эксперимента. Примените контроль внешней переменной.

Ситуация 2

В экспериментальном исследовании принимали участие две группы школьников: “совы” и “жаворонки”. Определите наличие внешней переменной и примените контроль внешней переменной.

Ситуация 3

В экспериментальном исследовании последовательно предъявлялись три цвета: красный, желтый и зеленый. Определите наличие внешней переменной и примените контроль внешней переменной.

ПСИХОЛОГИЧЕСКИЙ ПРАКТИКУМ

ЮНИТА 7

ЭКСПЕРИМЕНТ: ВНУТРЕННЯЯ И ВНЕШНЯЯ ВАЛИДНОСТЬ

Редактор Н.М. Пилипенко
Оператор компьютерной верстки А.Б. Кондратьева

Изд. лиц. ЛР № 071765 от 07.12.1998

Сдано в печать

НОУ “Современный Гуманитарный Институт”

Заказ

Тираж

Современный Гуманитарный Университет

ГЛОССАРИЙ

№ п/п	Новые понятия	Содержание
1	2	3
1	Экспериментальный метод	метод, который предполагает разработку плана управления независимой переменной
2	Психологические гипотезы	предположения, сформулированные относительно психологической реальности в рамках использования тех или иных психологических представлений
3	Безупречный эксперимент	мысленный образец психологического эксперимента, в котором отсутствуют угрозы достоверному выводу
4	Бесконечный эксперимент	мысленный образец для оценки реально проводимых экспериментов, в котором число проб или число испытуемых оцениваются с точки зрения охвата всех источников вариабельности переменных
5	Идеальный эксперимент	эксперимент, в котором изменению подвергается только независимая переменная, а зависимая переменная контролируется
6	Гипотеза	научное предположение в виде высказывания, истинность или ложность которого неизвестны, но могут быть проверены опытным путем (эмпирически)
7	Лабораторный эксперимент	эксперимент в специально созданных условиях, позволяющих выделить независимую переменную путем контроля всех других условий, с которыми может смешиваться ее влияние
8	Базисный процесс	реконструируемый в теории или в гипотетических конструктах процесс, определяющий показатели зависимой переменной в психологическом эксперименте
9	Независимая переменная	управляемое экспериментальное воздействие (фактор), т.е. активно изменяемая и контролируемая исследователем переменная, представленная на двух или более уровнях (качественных или количественных)

10	Функциональный контроль независимой переменной	основное условие для заключения о влиянии управляемых экспериментатором воздействий на измеряемые зависимые переменные
11	Дополнительное варьирование	способ повышения уровня обобщения выводов из экспериментального исследования, который предполагает систематическое варьирование дополнительных условий, важных для распространения каузальной зависимости на множество других ситуаций
12	Зависимая переменная	измеряемая в эксперименте переменная, изменения которой причинно обусловлены действием независимой переменной
13	Латинский квадрат	план организации многоуровневого эксперимента, в котором вторым экспериментальным фактором выступает порядок предъявления разных условий независимой переменной
14	Валидность исследования	оценка проведенного эмпирического исследования с точки зрения правильности его организации и возможности считать полученные результаты и сделанные выводы достоверными
15	Концептуальная репликация	проверка одной и той же теоретически предполагаемой зависимости разными методическими средствами
16	Внутренняя валидность	критерий оценки планирования и проведения исследования, позволяющий считать достоверными выводы относительно представленной в гипотезе зависимости
17	Внешняя валидность	критерий оценки исследования с точки зрения решения проблем соответствия сформулированных в гипотезе переменных и реализованных методически условий тем, на которые осуществляется обобщение полученных результатов
18	Гипотеза каузальная	причинно-следственная гипотеза, которая включает содержательную интерпретацию взаимосвязи между независимой и зависимой переменными, формулируется как высказывание о влиянии причинно действующих условий на изучаемый базисный процесс

1	2	3
19	Фактор фона	конкретные события, которые происходят между первым и вторым измерением наряду с экспериментальным воздействием
20	Фактор естественного развития	изменения испытуемых, являющиеся следствием течения времени, например, взросление, усиление голода, усталости
21	Эффект тестирования	влияние выполнения заданий, применяемых для измерения, на результаты повторного испытания
22	Инструментальная погрешность	нестабильность измерительного инструмента, при которой изменения в калибровке инструмента могут вызвать изменения в результатах измерения
23	Статистическая регрессия	фактор, имеющий место тогда, когда группы отбираются на основе крайних показателей и оценок
24	Фактор отбора испытуемых	неэквивалентность групп по составу, вызывающая появление систематической ошибки в результатах эксперимента
25	Фактор отсева испытуемых	неравномерность выбывания испытуемых из сравниваемых групп в ходе эксперимента
26	Реактивный эффект	возможное уменьшение или увеличение восприимчивости испытуемых к экспериментальному воздействию под влиянием предварительного тестирования
27	Контроль угроз внутренней валидности	способ контроля, который осуществляется благодаря устранению источников несистематической изменчивости при приближении эксперимента к бесконечному или идеальному
28	Контроль угроз внешней валидности	способ контроля, который осуществляется благодаря тому, что это эксперимент полного соответствия
29	Проверка статистических гипотез	условие вывода об экспериментальной гипотезе, если исследователь претендует на оценку достоверности полученных данных

30	Достоверный вывод	вывод об экспериментальной гипотезе, который основан на обобщении именно представленной в гипотезе связи
31	Гипотетические конструкты	понятия, которые соответствуют теоретическим представлениям и функционируют в гипотезе
32	Нуль-гипотеза	статистическая гипотеза, которая в экспериментальном исследовании противоположна по направленности утверждения об ожидаемых изменениях зависимой переменной в экспериментальной гипотезе
33	Дедуктивный вывод	умозаключение, реализующее путь обобщения от общего к частному
34	Индуктивный вывод	вывод о том, можно ли рассматривать управляемую независимую переменную в качестве причинно действующих условий, вызывающих экспериментальный эффект
35	Артефактные выводы	совокупность неверных умозаключений или обобщений, приводящих к недостоверным обобщениям относительно вида экспериментальной или корреляционной зависимости