

# Введение в экспериментальные исследования

### Обзор задач главы

Четыре центральные главы этой книги, с 5 по 8, посвящены проблемам, возникающим при разработке экспериментов. В первой части главы 5 рассматриваются основные характеристики эксперимента — варьирование изучаемого фактора (независимой переменной), контроль за остальными факторами (внешние переменные) и измерение результата (зависимая переменная). Из второй части главы вы узнаете, как более или менее качественно осуществленная разработка исследования может влиять на его валидность. Изучив данную главу, вы:

- сможете дать определение управляемой независимой переменной и найти примеры ситуативных, рабочих и инструктивных переменных;
- будете понимать разницу между экспериментальной и контрольной группами;
- научитесь обнаруживать наличие в эксперименте осложняющих переменных и понимать, почему это осложнение затрудняет интерпретацию результатов эксперимента;
- будете различать независимые и зависимые переменные в описании эксперимента;
- будете различать управляемые и субъективные переменные и понимать смысл проблемы интерпретации, которая сопутствует использованию субъективных переменных;
- сможете распознавать факторы, которые могут уменьшить валидность статистических выводов и конструктивную валидность эксперимента;
- сумеете описать различные варианты эффекта снижения внешней валидности эксперимента;
- сможете распознавать и описывать случаи, когда возникает угроза внутренней валидности эксперимента;
- будете понимать этические стороны работы в «фонде испытуемых».

Когда Роберт Вудвортс в 1938 г. опубликовал свою «Экспериментальную психологию», содержание этой книги было уже хорошо известно психологам. Еще в 1909 г. Вудвортс раздавал своим студентам в Колумбийском университете напечатанные на mimeографе копии брошюры «Проблемы и методы в психологии», а за-

тем, вместе с ней, «Руководство к лабораторным занятиям: эксперименты с памятью и т. д.», появившееся в 1912 г. К 1920 г. рукопись составила 285 страниц и была названа «Учебник по экспериментальной психологии». В 1932 г. напечатанный на mimeографе текст был переработан и наконец опубликован в 1938 г. К этому времени бывшие студенты Вудворта уже обучали своих студентов по этому учебнику — он был так широко известен, что издатели объявили о нем просто: «Библия напечатана» (Winston, 1990).

Так называемая колумбийская библия была настоящей энциклопедией, содержала более 823 страниц основного текста и список литературы на 36 страницах. В главах, следующих за вводной, вся информация была разделена на 29 отдельных тем, таких как «Память», «Работа с лабиринтами», «Время реакции», «Ассоциации», «Слуховое восприятие», «Восприятие цвета» и «Мышление». Работая с книгой, студенты изучали методы, применяемые в каждой конкретной области, а также узнавали практически все, что было известно по каждой теме на 1938 г.

Влияние колумбийской библии на обучение экспериментальной психологии невозможно переоценить. Сегодняшнее обучение экспериментальной психологии и в некоторой степени даже структура книги, которую вы сейчас читаете, следуют образцу Вудворта. В частности, он дал определение термину «эксперимент», до этого понимавшемуся широко, как практически любой вид эмпирического исследования (его определение мы используем и сегодня), и противопоставил эксперимент корреляционному исследованию (в наше время мы считаем такое разделение доказанным).

Отличительной особенностью экспериментального метода было управление явлением, названным Вудвортом «независимой переменной», которое влияет на нечто, обозначенное им как «зависимая переменная». Как писал сам Вудворт, экспериментатор «сохраняет постоянными все условия, за исключением одного, являющегося "экспериментальным фактором", или "независимой переменной"». Наблюдаемое следствие представляет собой "зависимую переменную", которой в психологическом эксперименте является определенная характеристика поведения или эмоция, о которой сообщает испытуемый» (Woodworth, 1938, p. 2). Хотя эти термины не были изобретены Вудвортом, он первый вложил в них тот смысл, который мы находим в них сегодня.

Тогда как в экспериментальном методе управляют независимыми переменными, при корреляционном исследовании, согласно Вудворту, «измеряется одна или несколько характеристик одного и того же человека [и] вычисляется корреляция между этими характеристиками. При использовании этого метода "независимые переменные" не используются и со всеми измеряемыми переменными обращаются одинаково» (Woodworth, 1938, p. 3). Более подробно корреляционные исследования рассматриваются в главе 9. Основное внимание этой и трех следующих глав сфокусировано на экспериментальном методе — наиболее мощном исследовательском инструменте выявления причинно-следственных связей.

## **Основные черты экспериментальных исследований**

Со времен Вудворта психологи рассматривают эксперимент как упорядоченное исследование, в ходе которого исследователь непосредственно изменяет некий фактор (или факторы), поддерживает остальные неизменными и наблюдает ре-

зультаты систематических изменений. Факторы, контролируемые в ходе эксперимента, называются независимыми переменными, факторы, поддерживаемые постоянными, — внешними переменными, а измеряемые поведенческие реакции — зависимыми переменными. Рассмотрим эти виды переменных более подробно.

### Способы задания независимых переменных

Любой эксперимент можно охарактеризовать как исследование, в котором изучается влияние  $X$  на  $Y$ .  $X$  — это, по Вудвортсу, **независимая переменная**. Она представляет собой интересующий экспериментатора фактор, изучаемый с целью обнаружить его влияние на поведение. Иногда его называют «управляемый фактор», так как экспериментатор осуществляет над ним полный контроль и сам разрабатывает ситуации, с которыми встретятся участники в ходе исследования. Как вы увидите далее, понятие независимой переменной может быть расширено и распространено также на неуправляемые или субъектные переменные, но сейчас мы будем рассматривать только такие независимые переменные, которые полностью контролируются экспериментатором.

Независимые переменные должны принимать минимум два значения, т. е. в эксперименте должны сравниваться по крайней мере две ситуации (или два условия). Предположим, к примеру, что исследователь хочет узнать влияние различных доз марихуаны на время реакции. В таком исследовании должны сравниваться по крайней мере две различные дозировки. Будут определены «количество марихуаны» как независимая переменная, а также «дозировка № 1» и «дозировка № 2» как два значения независимой переменной. Разумеется, независимая переменная может принимать более двух значений; по сути, использование большего количества значений имеет заметные преимущества — об этом вы узнаете в главе 7, посвященной экспериментальным планам.

Экспериментальные исследования в зависимости от своих целей могут быть либо фундаментальными, либо прикладными, а проводиться они могут как в лаборатории, так и в полевых условиях (о различиях между этими видами опыта подробно рассказывается в главе 3). Эксперименты, проводящиеся в полевых условиях, иногда называются **полевыми экспериментами**. Термин **полевое исследование** — более широкий и касается любого эмпирического исследования за пределами лаборатории, включая экспериментальные исследования и исследования, проводящиеся неэкспериментальными методами.

### Виды независимых переменных

Многообразие факторов, которые можно использовать в качестве независимых переменных, ограничено только пределами фантазии исследователя, но все независимые переменные, которыми управляют в ходе исследования, можно разделить на три пересекающиеся категории: ситуативные, рабочие и инструктивные переменные.

**Ситуативные переменные** — это различные особенности окружения, создаваемого для участников эксперимента. Например, если ученый, проводя исследование готовности людей оказывать помощь, хочет выявить влияние количества наблюдателей на вероятность оказания помощи, то он может создать ситуацию, в которой участники встречаются с человеком, которому нужна помощь. В одних случаях

участник один встречает человека, нуждающегося в помощи, а в других — участник и пострадавший могут быть окружены группой из трех или шести наблюдателей. В этом случае в качестве ситуативной независимой переменной может выступать количество людей, способных оказать помощь, не считая участника эксперимента, а ее значениями будут ноль, три и шесть (число наблюдателей).

Иногда экспериментаторы изменяют вид работы, выполняемой участниками. Один из способов управления **рабочими переменными** — давать для выполнения группам участников разные виды заданий. Например, в исследованиях по когнитивной психологии, чтобы определить наиболее распространенные виды ошибок, участников могут попросить решить несколько разных логических задач. Аналогично лабиринты могут различаться по уровню сложности, при исследовании восприятия могут использоваться разные виды иллюзий и т. д.

**Инструктивными переменными** можно управлять, попросив разные группы участников по-разному выполнить одно задание. Например, в ходе исследования памяти участникам, которым показывают один и тот же список слов, можно дать разные инструкции по способу запоминания. Одних можно попросить создавать визуальные образы для слов, других — устанавливать ассоциации между их смежными парами, а третьих — просто повторять каждое слово по три раза.

В одном исследовании можно также комбинировать различные виды независимых переменных. При исследовании влияния числа людей, их мотивации и сложности задания на способность к его решению одних участников можно поместить в большие, а других — в маленькие комнаты, тем самым воздействуя на фактор скопления людей с помощью ситуативной переменной размера комнаты. Далее, одним участникам в каждой комнате можно выдать сложные кроссворды, а другим — более легкие. Кроме того, с помощью инструктивной переменной можно воздействовать на мотивацию, сказав одним участникам, что они получают \$1 за кроссворд, а другим — \$5.

### Контрольные группы

В некоторых экспериментах независимая переменная — это то, осуществляется или нет экспериментальное воздействие. В таком случае она принимает значения 1 или 0 — это означает, что одни участники испытывают воздействие, а другие — нет. Например, в исследовании влияния сцен насилия, показанных по телевизору, на детскую агрессивность одним детям могут показать передачу с жестокими сценами, а другим — не показывать. Термин **экспериментальная группа** используется для обозначения участников, испытывающих экспериментальное воздействие. Участники, не испытывающие воздействия, принадлежат к **контрольной группе**. В идеале, участники из контрольной группы абсолютно идентичны участникам из экспериментальной группы и отличаются от них только тем, что испытывают экспериментальное воздействие. Таким образом, контрольная группа дает базовые показатели, с которыми сравниваются оценки экспериментальной группы. Можно представить себе это следующим образом: контрольная группа = группа сравнения. В главе 7, первой из двух глав, посвященных экспериментальным планам, вы узнаете о различных видах контрольных групп.

## Контроль за внешними переменными

Вторая особенность экспериментального метода состоит в том, что исследователь пытается контролировать так называемые **внешние переменные**. Они представляют собой неуправляемые факторы, сами по себе не интересующие исследователя, но способные повлиять на изучаемое поведение. Если их поддерживают постоянными, они не представляют опасности для исследования, но при отсутствии адекватного контроля над ними могут определенным образом воздействовать на исследуемое поведение. Такое явление называется осложнением. **Осложнитель** — это любая неуправляемая внешняя переменная, которая «коварирует» с независимой переменной и может повлиять на интерпретацию результатов. Осложняющие переменные сами по себе изменяют то, что изменяют независимые переменные (т. е. эти переменные «коварируют»), и поэтому их воздействие невозможно различить. Таким образом, если в исследовании имеется осложнитель, результаты такого исследования могут быть вызваны как осложняющей переменной, так и независимой, а также их комбинацией. Узнать при этом, чем в действительности вызываются результаты, не представляется возможным.

Чтобы проиллюстрировать действие простого осложнителя, рассмотрим один эксперимент. Исследователь поставил перед собой задачу доказать, что студенты, пытающиеся сразу изучить большое количество материала, не имеют таких хороших результатов, как те, кто разделяет учебу на несколько частей. Другими словами, быстрое зазубривание предполагается менее эффективным, чем равномерно распределенное изучение материала. Выбираются три группы студентов и каждой из них дается задание: выучить пять глав учебника по общей психологии. Первой группе дают на изучение материала три часа в понедельник, второй — три часа в понедельник и три часа во вторник, а третьей — по три часа в понедельник, вторник и среду. В пятницу проверяется знание материала участниками всех трех групп (план эксперимента см. в табл. 5.1). Результаты показывают, что участники группы 3 получили самые высокие оценки, более низкие оценки получили участники группы 2, а группа 1 практически не справилась с заданием. Исследователь заключает, что распределенные занятия более эффективны, чем массированные. Согласны ли вы с таким выводом?

**Таблица 5.1**

**Осложнения в гипотетическом эксперименте по распределению нагрузки**

	Понедельник	Вторник	Среда	Четверг	Пятница
Группа 1	3	—	—	-	Экзамен
Группа 2	3	3	-	-	Экзамен
Группа 3	3	3	3	-	Экзамен

*Примечание.* Число 3 в каждой клетке равняется количеству часов, использованных для изучения пяти глав учебника по общей психологии.

Скорее всего, вы не согласитесь с выводом экспериментатора, поскольку в этом исследовании можно легко обнаружить по крайней мере два осложнителя. Распределение занятий у участников действительно различается (1, 2 и 3 дня), но **также**

различается и общее время занятий (3, 6 и 9 часов). Это прекрасный пример осложнителя, так как невозможно определить, вызваны ли результаты одним фактором (распределение нагрузки) или другим (общее время занятий), ведь два этих фактора отлично коварируют. Для описания этой ситуации можно сказать, что «распределение нагрузки осложнено общим временем обучения». Второй осложнитель не так очевиден, но вызывает не меньше проблем. Речь идет о времени, проходящем до проверки результата. Проверка всех групп проводится в пятницу, но время, прошедшее с момента изучения материала, у каждой группы разное. Возможно, группа 3 лучше всего справилась с заданием потому, что ее участники изучали материал позднее и меньше успели забыть. Таким образом, в этом эксперименте фактор распределения нагрузки осложнен одновременно общим временем обучения и временем, прошедшим до проверки. Каждый осложнитель сам по себе может повлиять на результаты, но также они могут воздействовать совместно и изменить общую картину каким-то особым образом.

Взгляните на табл. 5.2. С ее помощью можно легко выявить осложнители. В первом столбце приведены значения независимой переменной, а в последнем — результаты проверки. В двух средних столбцах даны внешние переменные — их необходимо поддерживать постоянными, используя соответствующие стратегии (см. следующую главу). Если они не остаются постоянными (как в данном случае), то возникает осложнение. Очевидно, что результаты можно объяснить различиями в любом из первых трех столбцов, как взятых по отдельности, так и в различных комбинациях. Чтобы исправить ситуацию с осложнителями в данном примере, необходимо добиться, чтобы значения в двух средних столбцах не изменялись.

Таблица 5.2

#### Обнаружение осложнителей

Значения НП — Распределение нагрузки, дни	ВШ Время занятий, часы	ВП 2 Время до проверки, дни	ЗП Результат проверочного теста
1	3	3	Низкий
2	6	2	Средний
3	9	1	Высокий

НП — Независимая переменная.

ВП — Внешняя переменная.

ЗП — Зависимая переменная.

При работе над упражнениями в конце главы вам придется находить осложнители. Представление условия в виде таблицы, подобной табл. 5.2, облегчит вам выполнение задания. Попробуйте усовершенствовать эксперимент с распределением нагрузки. Как отделить осложнители от внешних переменных?

Научитесь помнить о возможности осложняющих факторов и разрабатывать механизмы контроля за ними — значит сформировать один из наиболее сложных навыков, необходимых для научного мышления. Далеко не все осложнители так очевидны, как в примере с массивированной/распределенной нагрузкой. В остальных главах мы будем часто встречаться с этой проблемой и кратко рассмотрим ее при обсуждении так называемой внутренней валидности исследования.

## Измерение зависимых переменных

Третий этап эксперимента — измерение характеристики поведения, предположительно испытавшей воздействие независимой переменной. Для описания характеристик поведения, которые измеряются в ходе эксперимента и являются его результатами, используется термин **зависимая переменная**. Если, как отмечалось ранее, эксперимент можно описать как воздействие  $X$  на  $Y$  и  $X$  — независимая переменная, то  $Y$  — это зависимая переменная. В исследовании влияния телевизионных сцен насилия на детскую агрессивность зависимой переменной является агрессивность. В исследовании фактора распределения нагрузки — результаты экзамена.

Правдоподобность любого эксперимента и возможность сделать существенное открытие отчасти определяются правильностью выбора зависимой переменной, измеряемой в ходе исследования. Как мы уже знаем, на эмпирический вопрос можно получить ответ, только если вопрос сформулирован с определенной степенью точности. Вы можете потратить немного времени и еще раз просмотреть главу о постановке эмпирических вопросов — она напомнит вам о понятии операционального определения. Один из ключевых моментов разработки эксперимента касается операциональных определений характеристик поведения, выбранных и измеряемых как независимые переменные. Воспроизведение эксперимента возможно, только если отдельные характеристики поведения точно определены.

В заключение отметим: очень важно понимать, что в зависимости от задач исследования конструктор может представлять собой независимую, внешнюю или зависимую переменную. В ходе эксперимента конструктором можно управлять как независимой переменной, пытаться контролировать его как внешний фактор или измерять как зависимую переменную. Рассмотрим, к примеру, конструктор «тревожность». Его можно представить в виде управляемой независимой переменной, если сказать участникам, что они будут подвергаться слабому, или напротив, болезненному, электрическому шоку, а затем спросить их, как они предпочитают ждать, в одиночестве или вместе с другими испытуемыми. Тревожность также может выступать как фактор, который в ходе эксперимента необходимо поддерживать постоянным. Например, если требуется оценить влияние семинара, посвященного публичным выступлениям, на умение студентов произносить речи, то не следует производить видеозапись студентов в одной группе, не записав также и другую группу. Если произведена видеозапись всех студентов, то уровень тревожности, вызванной фактором съемки, будет постоянным для всех групп. Кроме того, тревожность может быть зависимой переменной, как, например, в исследовании влияния различных видов экзаменов (например, заданий на выбор ответа или на написание эссе) на тревожность студентов во время сессии. В этом случае можно произвести особое психологическое измерение тревожности. Тревожность можно рассматривать и как характеристику личности, подразумевая, что одни люди обладают ею в более высокой степени, чем другие. Последнюю возможность мы рассмотрим в следующем разделе.

## Управляемые и субъективные переменные

До сих пор термин «независимая переменная» означал некий фактор, непосредственно изменяемый исследователем. В таком случае в ходе эксперимента сравниваются условия, созданные экспериментатором и находящиеся под его контролем, с другими условиями. Но во многих исследованиях сравниваются группы людей, которые отличаются друг от друга иначе, чем это было запланировано при разработке исследования. Тогда сравнение производится между факторами, обозначаемыми как переменные *ex post facto*<sup>1</sup>, переменные стихийно возникших групп, неуправляемые, или субъективные (дополнительные), переменные. Я предпочитаю использовать последний термин, но все они относятся к характеристикам участников исследования, таким как пол, возраст, интеллект, психические или физические нарушения или любые другие свойства личности, которые вы только сможете назвать. При использовании в исследовании субъектных переменных экспериментатор не имеет возможности управлять ими непосредственно, поэтому он *отбирает* людей для эксперимента на основании присущих им свойств.

Для иллюстрации различий между управляемыми и субъективными переменными рассмотрим гипотетическое исследование влияния тревожности на способность людей проходить лабиринт. Тревожностью можно *управлять* непосредственно, создав ситуацию, в которой одна группа будет испытывать тревожность (например, можно сказать, что участники будут выполнять задание перед лицом большого количества зрителей), а вторая — не будет (зрители будут отсутствовать). Человек, согласившийся принять участие в таком исследовании, является потенциальным участником любой из этих групп. С другой стороны, чтобы провести исследование с использованием *субъективной* переменной, можно подобрать две группы, различающиеся по уровню присущей им тревожности, и предложить участникам пройти лабиринт. В первую группу войдут люди с высокой тревожностью (предварительно выявленной с помощью специального теста). Вторая группа будет включать более спокойных людей. Обратите внимание, что существует заметная разница между таким исследованием и экспериментом с управляемой переменной. Если тревожность выступает как субъективная переменная, люди, желающие принять участие в исследовании, не могут произвольно помещаться в любую из двух групп (всегда чего-то боящегося Фреда нельзя помещать в группу с низкой тревожностью), но в зависимости от свойств, которые сформировались у них *еще до* участия в исследовании, должны быть помещены только в одну из них.

Одни исследователи в соответствии с первоначальным значением термина «независимая переменная», примененного Вудвортом, предпочитают использовать его для переменных, непосредственно управляемых экспериментатором. Другие рассматривают субъектные переменные как частный случай независимых переменных на том основании, что экспериментатор обладает некоторой степенью контроля над ними за счет предварительного отбора. Я придерживаюсь второй точки зрения и буду использовать термин «независимая переменная» в более широком смысле. Но если вы поймете разницу между управляемыми и неуправляемыми, или

Имеющие обратную силу {лат.}.



субъективными, переменными, то для вас будет не так важно, используется ли термин «независимая переменная» в широком смысле (управляемая + субъективная) или в более узком (только управляемая).

#### **Пример 4. Использование субъективных переменных**

Эксперимент, проведенный в 1992 г. Шнейдером и Бьерклендом, является хорошим примером одновременного исследования нескольких субъективных переменных. Также он иллюстрирует процесс воспроизведения и дополнения исследований (см. главу 3). Ученые заинтересовались вопросом, как воздействуют на память: а) специальные знания в определенной области и б) общие способности. Предыдущие исследования концентрировались только на знаниях, поэтому важным дополнением, вносимым их исследованием, было «рассмотрение уровня общих способностей детей» (р. 462). Таким образом, в исследовании использовались две субъективные переменные<sup>1</sup>. Первой был объем знаний о футболе. Эта переменная принимала два значения: знаток и новичок. Второй переменной были общие способности, определяемые по «тесту когнитивных способностей». Как это часто бывает при использовании субъективных переменных, исследование состояло из двух частей: первая заключалась в оценке субъективных переменных для формирования отдельных групп, а вторая — в проведении самого эксперимента. В первой части исследования большая группа учеников второго и четвертого классов прошла тест на когнитивные способности и тест, по которому определялись познания в области футбола. По полученным при тестировании результатам дети были отобраны для участия во второй части исследования и разделены на четыре группы: знатоки футбола с высокими способностями, знатоки футбола с низкими способностями, новички в футболе с высокими способностями и новички в футболе с низкими способностями.

Во второй части исследования дети выполняли задание на классификацию и запоминание: им раздали наборы терминов и попросили разделить их на группы, а сразу после этого — запомнить их. Задание выполнялось дважды, один раз на примере слов, связанных с футболом, а второй раз на примере других слов. (Одна половина детей выполняла задания именно в таком порядке, а другая — в обратном. Этот метод называется позиционным уравниванием и будет обсуждаться в следующей главе.) Результат получался сложный, но главным было то, что дети с высокими способностями в целом более успешно выполнили оба вида заданий, чем дети с низкими способностями вне зависимости от их уровня футбольных знаний. Однако при учете исключительно познаний в области футбола результаты показывали, что знатоки футбола лучше справились с заданием, посвященным футболу. Так, при выполнении задания с футбольными терминами знатоки запомнили 62% информации, а новички — только 44%. Но при выполнении задания со словами, не связанными с футболом, обе группы показали примерно одинаковые оценки: 52% для знатоков футбола и 48% для новичков.

Только исследование, в котором используются *управляемые* независимые переменные, можно назвать экспериментом в самом строгом смысле этого слова. Иногда

<sup>1</sup> В действительности кроме этих было еще три: пол, интерес к футболу и год обучения, но мы сконцентрируемся на двух важнейших.

такой эксперимент называют «истинным» (что звучит несколько претенциозно и приводит к мысли, что все остальные исследования «ложные»). Исследования, использующие независимые субъективные переменные иногда называют исследованиями *ex post facto*, или квазиэкспериментами («квази» в данном случае означает «в определенной степени»)<sup>1</sup>. Иногда (по сути, довольно часто) исследования имеют дело и с управляемыми, и с субъективными независимыми переменными. Очень важно осознавать наличие субъективных переменных, так как они влияют на выводы, которые делаются на основании результатов исследования.

### **Как делать выводы, если используются субъективные переменные?**

Поставьте галочку напротив этого раздела — он чрезвычайно важен. Вспомните: в главе 1 отмечалось, что одной из целей психологических исследований является объяснение поведения. Это означает, что мы хотим узнать, чем вызвано определенное поведение. Проще говоря, если используются управляемые переменные, то выводы о причинах поведения могут быть сделаны, а если субъективные — не могут. Однако суть дела состоит в степени контроля, который осуществляет экспериментатор в каждом из этих случаев.

В случае управляемых переменных эксперимент соответствует описанным в главе 1 условиям выявления причинно-следственной связи. Независимые переменные предшествуют зависимым, коварируют с ними и в случае, если нет осложнений, принимаются как наиболее разумное объяснение результатов. Другими словами, если вы изменяете один фактор и успешно поддерживаете остальные постоянными, то результаты могут объясняться *только* изменяемым фактором. В экспериментальном исследовании с двумя группами участников, не содержащем осложнений, группы будут равноценными (т. е. любые различия будут случайными) по всем показателям, кроме управляемого фактора.

При использовании субъективных переменных экспериментатор также может варьировать определенный фактор (т. е. отбирать участников, имеющих конкретные характеристики), но не может поддерживать все остальные факторы постоянными. Отбор участников по степени предрасположенности к тревожности (высокой или низкой) не гарантирует, что две группы будут эквивалентны по другим показателям. Фактически они могут различаться по многим показателям (например, по степени уверенности в себе), которые способны повлиять на результаты исследования. Если в ходе такого исследования возникает различие между группами, то мы не можем сказать, что его *причиной* является субъективная переменная. В аспекте условий причинно-следственной связи это значит, что если независимая переменная предшествует зависимой и коварирует с ней, то нельзя исключить возможность альтернативного объяснения их взаимосвязи, поскольку отсутствует контроль за некоторыми внешними факторами. При наличии субъективных переменных мы можем говорить лишь о том, что результаты групп различаются по зависимому показателю.

<sup>1</sup> Термин «квази-экспериментальный план» по сути является более широким обозначением любого плана, в котором участники не могут быть случайным образом разделены на группы, которые в дальнейшем будут изучаться (Cook & Campbell, 1979). Такие планы нередко встречаются в прикладных исследованиях и рассматриваются в главе 10.

Пример из социальной психологии поможет понять разницу. Предположим, вы интересуетесь альтруистическим поведением и хотите выяснить его зависимость от самооценки. Исследование можно провести двумя способами. Во-первых, можно непосредственно управлять самооценкой. Для этого участникам предварительно раздают тест личности, а затем сообщают им ложную информацию о результатах тестирования. В зависимости от того, какая информация сообщается, позитивная или негативная, самооценка может быть временно поднята или снижена. Далее можно попросить участников выполнить по желанию какую-либо работу и посмотреть, будут ли люди с высокой самооценкой помогать охотнее. Второй способ проведения исследования заключается в том, чтобы раздать участникам надежный и валидный тест личности для определения уровня самооценки, а затем разделить их на две группы: в одну войдут люди, чьи показатели выше 25%, а в другую — ниже этого уровня. Самооценка в этом случае является субъективной переменной, одна половина участников будет иметь низкую самооценку, а другая — высокую. Как и в первом исследовании, участников из этих двух групп можно попросить о добровольном участии в некоторой деятельности.

В первом исследовании различия в желании выполнять работу могут быть *непосредственно* привязаны к изменениям самооценки. Если все остальные факторы должным образом контролируются, различия в желании оказывать помощь могут быть вызваны *только* временным снижением или повышением самооценки. Во втором исследовании, однако, нельзя сказать, что высокая самооценка является непосредственной причиной желания помогать. Можно сказать лишь, что вероятность оказания помощи людьми с высокой самооценкой выше, чем людьми с низкой самооценкой. Все, что вы можете, — это строить предположения о причинах этого явления, так как участники могут отличаться друг от друга по каким-то другим, неизвестным вам показателям. Например, люди с высокой самооценкой могли уже иметь опыт добровольной помощи, который мог поднять или укрепить их самооценку и увеличить вероятность оказания ими помощи в будущем. Или они могут обладать специальными навыками, способствующими их добровольному участию в данного рода работе (например, навык публичных выступлений). Как вы узнаете из главы 9, в корреляционных исследованиях при попытках сделать выводы проблемы возникают именно из-за трудности интерпретации результатов.

Вернемся ненадолго к исследованию Шнейдера и Бьеркленда (Schneider and Bjorkland, 1992), в котором в качестве субъективных переменных использовались знания о футболе и общие способности. Авторы исследования тщательно избегали выводов о причинно-следственной связи. Слово «причина» никогда не встречается в их статьях, а описание результатов всегда приводится в следующем виде: «оценки у одной группы выше, чем у другой».

Прежде чем переходить к дальнейшему обсуждению валидности экспериментальных исследований, прочитайте вставку 5.1. В ней характеризуются переменные, которые были использованы в одном из экспериментов с куклами бобо (классическом исследовании, знакомом вам из курса по общей психологии), первом исследовании подражательной агрессии. Изучение данного примера позволит вам применить знания о независимых, внешних, и зависимых переменных и увидеть, как управляемые и субъективные переменные могут встретиться в одном исследовании.

## ВСТАВКА 5.1

**Классические исследования— куклы бобо и агрессия**

Спросите любого студента, только что прослушавшего курс по детской психологии, социальной психологии или психологии личности (а возможно, даже и по общей психологии) об экспериментах с куклами бобо, и он сразу поймет, о чем речь, и расскажет вам что-то вроде «Да, я знаю, это эксперименты, которые показали, что дети бьют надувных кукол, если видят, как это делают взрослые». Описание одного из экспериментов поможет нам понять разницу между независимыми, внешними и зависимыми переменными. Результаты исследования были опубликованы Альбертом Бандурой и его коллегами в 1963 г. под названием «Имитация моделей агрессивности, представленных посредством видеозаписи» [*Imitation of Film-Mediated Aggressive Models*, Bandura, Ross, & Ross, 1963].

**Задание независимых переменных**

Эксперимент включал как управляемые так и субъективные переменные. Основной управляемой переменной был вид опыта, предшествующего возможной агрессии. Переменная принимала четыре значения, которым соответствовали три экспериментальные группы и одна контрольная.

*Экспериментальная группа 1:* агрессия в реальной жизни (дети непосредственно наблюдали модель агрессивного поведения взрослых с куклой бобо).

*Экспериментальная группа 2:* агрессия, воспринимаемая в видеозаписи (дети просматривали модели агрессивного поведения взрослых с куклой бобо в видеозаписи).

*Экспериментальная группа 3:* агрессия, воспринимаемая в мультфильмах (дети просматривали мультфильм «Кот Герман» с агрессией, направленной против мультипликационной куклы бобо).

*Контрольная группа:* модель агрессивного поведения не демонстрировалась.

Неуправляемой независимой переменной (субъективной переменной) являлся пол. В исследовании участвовали дети мужского и женского пола из детского сада при Стэнфордском университете (средний возраст - 52 месяца). (В действительности была еще одна управляемая переменная: участникам из групп 1 и 2 показывали модель с участием людей либо того же, либо противоположного пола). Базовая экспериментальная процедура заключалась в том, что детям показывали определенный вид агрессивности (или не показывали, в случае контрольной группы), а затем помещали их в комнату, заполненную игрушками (в числе прочих - кукла бобо), тем самым предоставляя возможность самим проявить агрессивность.

**Контроль за внешними переменными**

Были тщательно устранены некоторые возможные осложнители. Во-первых, для групп 1 и 2 взрослые демонстрировали агрессию, направленную против 5-футовой куклы бобо, а когда детям дали возможность самим поколотить куклу, их поместили в комнату с 3-футовой бобо. Тем самым отношение размеров куклы и человека поддерживалось приблизительно постоянным. Во-вторых, участников во всех четырех группах слегка расстроили перед тем как дали возможность проявить агрессию. Им разрешили немного поиграть с очень привлекательными игрушками, а затем экспериментатор сказал, что эти игрушки не простые и предназначены для других детей. Таким образом все дети испытали приблизительно одинаковое нарастание уровня эмоционального возбуждения до того, как им дали возможность проявить агрессию. Поэтому любые различия в агрессивности должны были быть связаны с подражанием, а не с эмоциональными различиями между группами.

### Измерение зависимых переменных

В этом эксперименте использовались несколько различных способов измерения агрессивности. Агрессивные реакции классифицировались как подражательные, частично подражательные или неподражательные в зависимости от того, насколько близки они были к соответствующей модели поведения. Например, в операциональное определение подражательного агрессивного поведения входили удары куклы деревянным молотком, удары кулаком в нос и всевозможные пинки, частично подражательного - удары молотком по другим предметам и сидение на кукле без нанесения ей ударов, а неподражательная агрессия включала стрельбу дробинками из ружья в другие мишени, кроме куклы бобо, а также агрессивные действия, направленные против других объектов в комнате.

Если кратко сформулировать результаты, то дети в группах 1, 2 и 3 показали значительно более высокую степень агрессии, чем дети в контрольной группе, но при этом степень агрессии была одинаковой при различных моделях поведения. Кроме того, мальчики при любых условиях вели себя более агрессивно, чем девочки. Также некоторые связанные с полом различия наблюдались в особенностях проявления агрессии: девочки «охотнее сидели на кукле бобо, но (в противоположность мальчикам) воздерживались от нанесения ударов» (Bandura et al., 1963, p. 9). Рисунок 5.1 обобщает результаты исследования.

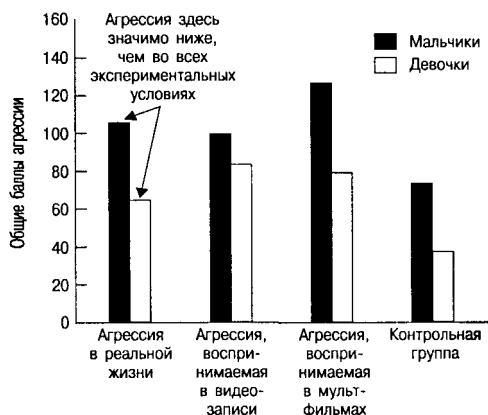


Рис. 5.1. Данные исследования с куклами бобо, проведенного Бандурой, Россом и Россом в 1963 г. для изучения влияния подражания на агрессию

## Валидность экспериментальных исследований

В главе 4 было введено понятие валидности в применении к измерениям. Этот термин также применяется к эксперименту в целом. Так же как измерение считается валидным, если измеряется именно то, что предполагается, психологическое исследование считается валидным, если оно дает такое объяснение поведения, которое, как предполагается, оно должно давать. Эта часть главы познакомит вас с четырьмя различными видами валидности по схеме, предложенной Куком и Кэмпбеллом (Cook and Campbell, 1979) для полевых исследований, но применимой также и для любых других психологических исследований. Эти четыре вида — валидность статистических выводов, конструктивная валидность, внешняя и внутренняя валидность.

## Валидность статистических выводов

В предыдущей главе вы познакомились с применением статистики в психологии. В частности, вы узнали о видах шкалы измерений, проверке гипотез и получили представление о различиях между описательной статистикой и статистикой вывода. **Валидность статистических выводов** связана с границами корректного использования статистики и получением на основании статистического анализа правильных выводов.

Статистическая валидность исследования может снижаться по нескольким причинам. Во-первых, исследователь может неверно проводить анализ данных или нарушать условия, которые необходимо соблюдать при проведении конкретного вида анализа. Например, данные исследования могут быть расположены на порядковой шкале, а следовательно, потребуется использование определенной статистической процедуры. Исследователь, однако, по ошибке может использовать вид анализа, подходящий только для данных, представленных с помощью интервальной шкалы или шкалы отношений. Во-вторых, исследователь может избирательно сообщить об одних результатах анализа, полученных в соответствии со сделанными предположениями, и скрыть остальные (догадайтесь, какие) — такая практика граничит с мошенничеством (см. главу 2). Третий фактор, который способен снизить статистическую валидность исследования, — это надежность проводимых измерений. Если зависимые показатели ненадежны, возникает множество ошибок, что снижает вероятность обнаружения значимого влияния одних показателей на другие. Если такое влияние действительно существует (т. е. нулевая гипотеза должна быть отвергнута), а низкая надежность не позволяет его обнаружить, то результатом будет ошибка 2-го рода по вышеприведенной классификации.

Грамотные исследователи проводят планирование статистического анализа одновременно с разработкой экспериментального плана. По сути, эксперимент не может быть разработан без выяснения способов анализа данных.

## Конструктивная валидность

В предыдущей главе конструктивная валидность описывалась в связи с измерением психологических конструктов: она свидетельствует одновременно о том, действительно ли в ходе тестирования измеряется некий гипотетический конструкт (например, задержка удовольствия) и соответствует ли этот конструкт экспериментальной реальности. В случае экспериментальных исследований термин **конструктивная валидность** имеет сходное, но несколько отличающееся значение: он обозначает, насколько адекватно определяются независимые и зависимые переменные, используемые в исследовании. В случае исследования влияния показа по телевидению жестоких сцен на детскую агрессивность для определения конструктивной валидности можно задать следующие вопросы: а) отражают ли наилучшим образом отобранные экспериментатором программы контраст между программами, содержащими сцены жестокости, и без них, и б) является ли они наилучшим из возможных способ измерения агрессии? Если в исследовании сравниваются жестокие мультипликационные персонажи (например, Элмер Фадд, стреляющий в Багса Банни) с обычными персонажами, (например, Винни-Пух), то на это могут возразить, сказав, что фантастические образы никак не могут влиять на агрессив-

ное поведение детей, поэтому более валидный способ управления независимой переменной, называемой «уровень жестокости, представленной в видеозаписи», будет заключаться в показе детям реальных видеозаписей людей, демонстрирующих разную степень жестокости.

Аналогично критике может быть подвергнут применяемый в конкретном исследовании способ измерения агрессии — и это действительно вызвало проблемы при исследовании данного явления. По вполне очевидным этическим причинам вы не должны разрабатывать исследование, в котором участники будут колотить друг друга. Агрессия определялась операционально самыми различными способами, и некоторые из них могут показаться вам более валидными (например, предварительно разозленные испытуемые считали, что наказывают ударом тока других испытуемых), чем другие (например, гудение автомобилей раздраженных водителей). Так же как в случае конструктивной валидности, которая обсуждалась в предыдущей главе, посвященной валидности измерений, валидность определения независимых и зависимых переменных вырабатывается по мере того, как в ходе исследования накапливаются данные и появляется возможность их описания при помощи последовательной (и сводимой к единому основанию) системы.

### **Внешняя валидность**

Психологов-экспериментаторов часто критиковали за то, что о студентах-второкурсниках и белых крысах они знают очень много, а обо всем остальном — очень мало. По сути, это критика **внешней валидности** — возможности генерализации данных, полученных в ходе исследования, за пределами содержания конкретных экспериментов. Чтобы исследование обладало наивысшей степенью внешней валидности, должна существовать возможность генерализации его результатов по трем направлениям: на другие популяции, другие условия и другое время.

### **Другие популяции**

Именно к этому разделу относится замечание насчет крыс и второкурсников. Как мы узнали из главы 2, часть споров о правомерности исследований с животными касается того, насколько важны открытия, сделанные в ходе таких исследований, для объяснения человеческого поведения. Что касается второкурсников, вспомните, что Милграм сознательно не хотел использовать студентов и в качестве объектов для исследования подчинения и отобрал взрослых людей из нейтральной в этом отношении популяции. Однако, о большинстве психологов нельзя сказать того же. Обзор, проведенный Сирсом (Sears, 1986) в ходе исследования по социальной психологии, показал, что 75% исследований, опубликованных в этой области, в 1980 г. проводились с использованием студентов. Повторный обзор для исследований, опубликованных в 1985 г., дал результат 74%. Сирс утверждал, что особенности студентов колледжей как популяции легко могут исказить общие выводы о социальных явлениях. В сравнении с основной популяцией, к примеру, студенты колледжей обладают более высокими когнитивными способностями, они более эгоистичны и восприимчивы к социальным воздействиям, а также легче меняют свое отношение к различным вопросам. По Сирсу, если в исследовании изучаются вопросы, связанные с этими особенностями, результаты, полученные при участии

студентов, могут не быть пригодными для других групп людей. Ученый считал, что исследователи должны расширить свои базы данных и воспроизводить важные открытия на различных популяциях. Но он также указал, что во многих областях исследований (например, при изучении восприятия) данные особенности студентов практически не влияют на получаемые результаты и, кроме того, студентов очень много, и поэтому они являются классическим примером удобной выборки (глава 4). Как показано во вставке 5.2, к этой группе применяется особый этический подход.

#### ВСТАВКА 5.2

### **Этика — набор участников - все становятся членами фонда**

Большинство психологов-исследователей работают в колледжах и университетах и поэтому окружены потенциальными «морскими свинками» в виде студентов. Поскольку студенты не всегда охотно принимают участие в исследованиях, многие психологические факультеты университетов организуют так называемые «фонды испытуемых», или, как это называется в последнее время, «фонды участников». Этот термин обозначает группу студентов, обычно посещающих вводный курс по психологии, которых просят участвовать в исследованиях как в обязательных для данного курса занятиях. Если вы студент большого университета, то, возможно, уже имели опыт «добровольного» участия в двух или трех экспериментах, чтобы избежать снижения оценки за весь курс. Если в большом университете курс по общей психологии каждый семестр посещают 800 студентов и каждый студент записывается для участия в трех экспериментах, то исследователи получают возможность работать с 2400 участниками.

Фонды испытуемых облегчают исследователям поиск участников, а их существование оправдывают тем, что участие в исследованиях входит в состав образовательного процесса (Kimmel, 1996). В идеале, участвуя в экспериментах и изучая исследуемые психологические явления, студенты получают более глубокое понимание исследовательского процесса. Чтобы сделать участие студентов действительно добровольным, кроме непосредственного участия в исследовании им предлагают на выбор несколько вариантов. Однако проблемы все равно существуют. Критики утверждают, что при формировании фондов участников в действительности не соблюдается принцип добровольности, что альтернативные занятия (например, оформление документов) зачастую настолько тягостны, что студентов по сути вынуждают участвовать в исследованиях, а также что приобретаемый опыт скорее скучный, чем поучительный (Kern, 1988). В поддержку такого мнения были проведены исследования. В исследовании Зибера и Сакса (Sieber and Saks, 1989) было обнаружено, что 89% из 366 рассмотренных факультетов содержат фонды, не соответствующие одной или нескольким рекомендациям ассоциации АРА. Тем не менее многие факультеты стараются сделать так, чтобы студенты могли извлечь пользу из подобных занятий. Например, во время дебрифинга после эксперимента по изучению памяти участникам/студентам могут рассказать о том, как данное исследование связано с главой X из учебника, используемого во вводном курсе. Также на многих факультетах студентам в качестве альтернативы предлагают творческие виды деятельности. Студенты, не участвующие в исследованиях могут: а) наблюдать проводимые эксперименты и записывать свои наблюдения, б) участвовать «за кулисами» в проектах преподавателей факультета (например, кодировать данные), в) участвовать в других мероприятиях научного коллектива или г) слушать лекции ученых, приехавших в данное учебное заведение (Kimmel, 1996). Некоторые исследования показали, что студенты в основном находят участие в экспе-



риментах весьма полезным, особенно если исследователи действительно стараются «привязать» их участие к процессу обучения, протекающему на занятиях по общей психологии (например, Landrum & Chastain, 1999; Leak, 1981).

Ассоциация APA (1982, р. 47-48) разработала особые требования к отбору студентов для участия в исследованиях. Вот основные из них:

- 1) студенты должны быть осведомлены о требованиях курса, прежде чем на него записываться;
- 2) студенты в первый день занятий по данному предмету должны получить подробное описание требований курса, в том числе ясное описание альтернативных видов деятельности, в случае если они не хотят участвовать в исследованиях;
- 3) альтернативные виды деятельности должны быть равными участию в исследовании по времени и трудоемкости, а также, аналогично участию в экспериментах, должны иметь образовательное значение;
- 4) все предложения о проведении исследований с использованием фондов участников должны получать одобрение соответствующего исследовательского совета;
- 5) особое внимание следует обратить на вежливое обращение со студентами;
- 6) должна существовать простая и понятная процедура, посредством которой студенты могут пожаловаться на некорректное обращение. Оценка по данному предмету при этом не должна быть снижена;
- 7) все остальные требования этического кодекса APA должны неукоснительно выполняться;
- 8) на факультете должен быть разработан механизм для периодической проверки правил использования фонда участников.

«Проблема второкурсников» — это только один пример затруднений, связанных с распространением результатов исследования на другие группы. Другой пример касается пола участников. Некоторые из наиболее известных психологических исследований ограничены изучением только мужчин (или, но гораздо реже, только женщин), но распространяют свои выводы на людей обоих полов. Вероятно, самый известный пример — это исследование Лоренса Колберга, посвященное изучению нравственного развития детей. В ходе исследования, проведенного в 1963 г., Колберг просил мальчиков-подростков прочитать, а затем разрешить различные коротко сформулированные нравственные дилеммы. На основании полученных ответов Колберг разработал теорию о шести стадиях морального развития, закрепившуюся в литературе по психологии развития. Человек, находящийся на самой высокой стадии, поступает в соответствии с универсальными принципами справедливости и соблюдения прав личности.

Теория Колберга была подвергнута критике за низкую внешнюю валидность. К примеру, Кэрол Гиллиган утверждала (Gilligan, 1982), что модель Колберга не учитывает половых различий в особенностях мышления и способах принятия моральных решений. Мужчины могут уделять огромное внимание правам личности, а женщины больше внимания обращают на сохранение взаимоотношений между людьми, а следовательно, ответы женщин на некоторые из дилемм Колберга могут не соответствовать столь «высокой ступени развития», как ответы мужчин. По мнению Гиллиган, это связано с недостатками модели Колберга, составившего выборку из подростков исключительно мужского пола.

Также психологи-исследователи обращают внимание на перенос результатов из одной культуры в другую. Например, считается, что в «индивидуалистических» культурах личность ставится выше группы и высоко ценятся личная ответственность и инициатива. С другой стороны, в «коллективистских» культурах группа оценивается выше личности (Triandis, 1995). Выводы исследований, базирующихся на одной культуре, могут не получить универсального применения. Приведем один пример. В США большинство детей учат высоко ценить личные достижения, а, например, в Японии дети считают, что если они выделяются из коллектива, то тем самым снижают ценность остальных членов группы, и поэтому личные достижения ценятся не столь высоко. В одном исследовании было обнаружено, что у американских студентов личные достижения связаны с положительными эмоциями, а у японских — с *отрицательными* (Kitayama, Markus, Matsumoto & Norasakkunkit, 1997). Решить, что положительное отношение к личным достижениям — это общечеловеческая особенность, значит совершить ошибку. Но означает ли это, что во всех психологических исследованиях должны проводиться межкультурные сравнения? Конечно нет. Это лишь означает, что выводы необходимо делать очень внимательно и всегда указывать группу, изучаемую в ходе исследования.

### Другие условия

Обладающие внешней валидностью результаты можно распространить не только на другие группы людей, но и на другие условия — они действительно также для других наборов стимулов. Этот вопрос лег в основу критики лабораторных исследований, упомянутой в главе 3: иногда говорят, что такие исследования искусственны и слишком далеки от реальной жизни. Вспомните, как при обсуждении фундаментальных и прикладных исследований рассказывалось, что ответом работающих в лаборатории исследователей на критику явился тезис Аронсона об экспериментальном реализме. Наиболее важно, чтобы люди действительно были вовлечены в исследование, а жизненный реализм при этом уходит на второй план.

Тем не менее попытки изучать психологические явления в реальных жизненных условиях привели к важным открытиям во многих областях психологии. Хороший пример тому — история исследований человеческой памяти. Большую часть XX в. исследования памяти проводились преимущественно в лабораториях, где бесчисленные второкурсники запоминали бесконечные списки слов, бессмысленных слогов, последовательности цифр и т. д. Эти исследования позволили получить большое количество информации об общих механизмах памяти, и это очень важно, но не было ясно, можно ли открытые в лаборатории принципы распространить на ситуации реальной жизни. Переломной точкой стала работа ученого Корнельского университета Ульриха Нейссера «Познание и реальность» (Neisser, *Cognition and Reality*, 1976). В своей работе он указал, что лабораторная традиция в когнитивной психологии, приводя к важным открытиям, тем не менее не способна в достаточной мере пролить свет на процессы обработки информации в реальных условиях, и призвал к исследованиям, обладающим **экологической валидностью**, — исследованиям, важным для каждодневной познавательной деятельности людей, старающихся приспособиться к окружающим условиям. Нейссер заявил, что экспериментальная психология «должна приложить серьезные усилия для понима-

ния особенностей познания, протекающего в обычных условиях и в процессе естественной целенаправленной деятельности. Это означает не прекращение лабораторных экспериментов, а переход от исследования переменных, которыми легко управлять, к изучению экологически важных переменных» (р. 7).

На призыв Найссера откликнулись многие (но, конечно, не все) исследователи, занимающиеся вопросами познания. В 1980-1990-х гг. увеличилось число исследований памяти свидетелей (Loftus, 1979) и долговременной памяти — как, например, запоминание содержания школьных курсов, таких как испанский язык (Bahrick, 1984). Сам Найссер провел интересный анализ памяти Джона Дина (Neisser, 1981), главного советника Белого Дома, положившего конец попыткам президента Ричарда Никсона скрыть незаконные действия во время уотергейтского скандала в начале 1970-х гг. Его показания в конгрессе повлияли на развитие скандала и в конечном счете привели к отставке Никсона. Показания Дина заняли 245 страниц текста и были настолько подробными, что некоторые обозреватели назвали его живым диктофоном. Позднее выяснилось, что описанные Динем встречи в Овальном кабинете были записаны также Белым Домом. Сравнение этих записей с показаниями Дина дало Найссеру прекрасную возможность оценить предполагаемую фотографическую память Дина, которая в конце концов оказалась не столь совершенной. Дин смог вспомнить основные вопросы, обсуждавшиеся на таких встречах, но пропустил множество деталей. Для обсуждаемой темы внешней валидности важно, что исследование Найссера отлично показывает, как наши знания о механизмах памяти могут быть расширены с помощью изучения явлений за пределами лаборатории.

### **Изменение временного контекста**

Еще один вопрос, связанный с внешней валидностью, касается долговечности результатов. Одно из самых известных в истории психологических исследований — это исследования конформности, проведенные Соломоном Ашем в 1950-х гг. (например, Ash, 1956). Эксперименты проводились в исторический период, когда в США главенствовали консервативные силы, люди опасались «красной угрозы», а конформность и послушание властям высоко оценивались обществом. Аш обнаружил, что студенты колледжей очень восприимчивы к давлению сверху. Действительно ли это и сегодня, в другом политическом климате? Будет ли фактор воздействия на конформность, найденный Ашем (групповое согласие), главенствовать и в наши дни? Вообще, исследования, посвященные более фундаментальным процессам (например, познанию), выдерживают проверку временем лучше, чем исследования социальных факторов, включенных в определенный исторический контекст.

Обобщая изложенное ранее, можно сказать, что валидность данных увеличивается, если их можно применить к другим людям, другим условиям и другой эпохе. Но должны ли поэтому ученые разрабатывать исследования, включающие множество разных социальных групп, проводить эксперименты в различных условиях и повторять их каждые десять лет? Конечно, нет. Внешняя валидность не определяется по одному исследованию. Она представляет собой еще один пример важности процесса повторения и дополнения. Для ученого, разрабатывающего исследование, важность внешней валидности теряет значение в сравнении с важностью нашей следующей темы.

### **Внутренняя валидность**

Последний вид экспериментальной валидности, описанный Куком и Кэмпбеллом (Cook and Campbell, 1979), называется **внутренняя валидность**. Этот термин обозначает методологическое совершенство эксперимента и отсутствие в нем осложнений. При обладающем внутренней валидностью исследовании экспериментатор уверен, что результаты, полученные измерением зависимой переменной, непосредственно связаны с независимой переменной, а не с каким-нибудь другим неконтролируемым фактором. В исследовании, отягощенном осложняющими факторами, как мы видели на примере эксперимента с массиванной/распределенной нагрузкой, результаты невозможно интерпретировать. Результат может быть связан с влиянием независимой переменной, осложняющей (осложняющих) переменной (переменных) или их комбинации, при этом невозможно выбрать одно из нескольких различных объяснений.

### **Угроза для внутренней валидности**

Внутреннюю валидность исследования может снизить любой неконтролируемый внешний фактор, однако есть несколько вопросов, требующих особого рассмотрения (Cook & Campbell, 1979). Так называемая угроза для внутренней валидности наиболее опасна при отсутствии контрольной группы. Эта проблема иногда возникает в исследованиях по оценке программ (глава 10). Часто такая угроза возникает в длительных исследованиях, в ходе которых проводятся несколько измерений. Например, участники могут пройти предварительное тестирование, подвергнуться особому экспериментальному воздействию, а затем пройти заключительное тестирование. В идеале воздействие должно иметь определенный положительный эффект, который можно определить, обнаружив различия между результатами предварительного и заключительного тестирования. Второй вид угрозы возникает, если сравниваются «неэквивалентные» группы. Эта так называемая проблема отбора испытуемых может переплетаться с другими видами угрозы для внутренней валидности.

### **Предварительное и заключительное тестирование**

Верно ли, что студентам легче изучить общую психологию с помощью компьютеризированного курса, где каждый может продвигаться вперед соразмерно своей собственной скорости? Если в колледже будет проводиться программа по снижению экзаменационной тревожности, как можно показать, что она действует эффективно? Если вы обучите людей различным стратегиям запоминания, улучшит ли это их память? Во всех этих эмпирических вопросах речь идет о том, изменятся ли люди, усвоив некоторый опыт (курс, программа, тренировка памяти). Стандартный способ обнаружить изменения — проверить людей до получения ими данного опыта (провести **предварительное тестирование**), а затем после получения опыта (провести **заключительное тестирование**). Идеальным результатом для только что описанных примеров будет, если люди в ходе заключительного тестирования покажут, что они: а) знают общую психологию лучше, чем вначале, б) меньше волнуются перед экзаменами, чем раньше, и в) демонстрируют улучшение памяти. В ходе

стандартного исследования сравниваются экспериментальная и контрольная группы; последняя при этом не подвергается никакому воздействию:

Экспериментальная группа:	предварительное тестирование	<i>Воздействие</i>	заключительное тестирование
Контрольная группа:	предварительное тестирование		заключительное тестирование

При отсутствии подходящей контрольной группы возникает угроза внутренней валидности эксперимента. Предположим, вы хотите оценить эффективность программы помощи студентам колледжа, страдающим от экзаменационной тревоги (т. е. тем, которые имеют неплохие навыки обучения и знают материал, но так сильно волнуются во время экзаменов, что это мешает их успешной сдаче). В начале лекций первокурсники заполняют различные опросники, в том числе один, выполняющий функцию предварительного тестирования для изучения тревожности. Предположим, что оценки могут варьироваться от 20 до 100, и чем больше значение, тем выше уровень тревожности. Студентов, получивших высокие оценки, просят участвовать в программе по снижению экзаменационной тревоги, включающей тренинги на расслабление и выработку учебных навыков, а также другие методики. Спустя три месяца студенты снова выполняют тест на тревожность и его результаты выглядят следующим образом:

Предварительное тестирование	<i>экспериментальное воздействие</i>	заключительное тестирование
90		70

Таким образом, средняя предварительная оценка у студентов, выбранных для участия в программе, равна 90, а средняя заключительная — 70. Если предположить, что разница статистически значима, какой вывод можно сделать? Была ли эффективна программа? Чем были вызваны изменения, экспериментальным воздействием, или же могли сказаться другие факторы? Я надеюсь, вы понимаете, что есть несколько способов объяснения таких результатов. Читайте дальше.

### **История и взросление**

Иногда между предварительным и заключительным тестированием вклинивается некоторое событие, не связанное с экспериментальным воздействием, но вызывающее большие изменения. В таком случае угрозой для внутренней валидности исследования является история. Предположим, к примеру, что в колледже из предыдущего примера решили, что система оценок мешает обучению и что знания по всем предметам впредь должны оцениваться по принципу сдал/не сдал. Допустим, что такое решение было принято после проведения предварительного тестирования тревожности в середине экспериментальной программы по ее снижению. Заключительное тестирование может показать значительное снижение тревожности, но очень вероятно, что такой результат будет вызван изменением системы оценок, а не самой программой. Разве вы не станете меньше волноваться по поводу данного курса об исследовательских методах, если оценка не будет иметь никакого значения?

Кроме того, в программе по снижению экзаменационной тревоги студенты-первокурсники принимают участие в самом начале своей учебной деятельности, поэто-

му изменение оценок предварительного и заключительного тестирования может быть связано со **взрослением** студентов по мере привыкания к жизни колледжа. Как вы, вероятно, помните, первый семестр в колледже — это время значительных изменений в жизни студентов. Если исследование занимает продолжительный период, то взросление студентов обязательно скажется на его результатах.

Обратите внимание, что если будет использоваться контрольная группа, то экспериментатор сможет учесть влияние и истории, и взросления. Если будут получены приведенные ниже результаты, то это влияние можно будет исключить, а программу снижения экзаменационной тревожности считать эффективной:

Экспериментальная группа:	предварительное тестирование 90	<i>экспериментальное воздействие</i>	заключительное тестирование 70
Контрольная группа:	предварительное тестирование 90		заключительное тестирование 90

С другой стороны, если оценки контрольной группы при заключительном тестировании также снижаются до 70, и история, и взросление, а также оба фактора одновременно могут являться объяснением изменений оценок экспериментальной группы.

### Регрессия

Регрессия означает движение в обратную сторону — в данном случае в направлении средней оценки. Поэтому явление, которое я собираюсь описать, иногда называют **регрессией к среднему**. В большом наборе оценок большинство из них группируются вокруг среднего арифметического и лишь некоторые (крайние) оценки будут заметно отстоять от него. Представьте, что вы случайным образом выбираете несколько оценок из нормального распределения, показанного на рис. 5.2. Большинство оценок группируются вокруг среднего, так что если вы выбираете случайным образом, то наиболее вероятно, что вы выберете оценку, близкую к среднему арифметическому (X слева на рис. 5.2). Предположим, однако, что вы выбрали одну, далеко отстоящую от средней оценки (т. е. крайнюю оценку — Y). Верно ли, что при следующем выборе вы, вероятнее всего, получите:

1. Ту же оценку еще раз?
2. Оценку еще более крайнюю, чем первая?
3. Оценку ближе к среднему арифметическому, чем первая?

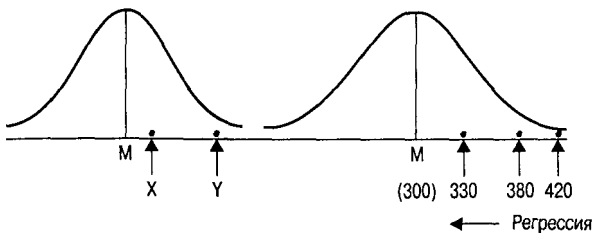


Рис. 5.2. Регрессия к среднему

Я полагаю, вы выберете третий вариант, и это будет означать, что вы понимаете идею регрессии к среднему. Приведем более конкретный пример (относящийся к правой части рис. 5.2). Предположим, вы знаете (на основании нескольких сотен бросков), что в среднем Тэд бросает бейсбольный мяч на 300 футов. Затем он бросает мяч на 380 футов. Если вы будете спорить о его *следующем* броске, на что вы поставите?

1. На 380 футов.
2. На 420 футов.
3. На 330 футов.

И опять я думаю, что вы выберете третий вариант, тем самым еще раз подтвердив, что понимаете идею регрессии к среднему. Но как это связано с нашими предварительными и заключительными исследованиями?

Для многих исследований, включающих предварительное и заключительное тестирование, участники отбираются на основании того, что в ходе предварительной проверки они показали результаты, близкие к крайним. Так, в исследовании экзаменационной тревожности отбирались те участники, чьи показатели тревожности были очень высоки. Заключительное тестирование может показать улучшение показателей, которое возможно будет связано с регрессией, а не с программой улучшения памяти. И снова контрольная группа, состоящая из участников с высокой тревожностью, поможет исследователю заметить наличие регрессии. Например, следующие результаты предполагают, что регрессия могла оказать некоторое влияние<sup>1</sup>, но тем не менее воздействие самой программы было более сильным и перекрыло регрессию. Вы понимаете, почему это так?

Экспериментальная группа:	предварительное тестирование 90	<i>экспериментальное воздействие</i>	заключительное тестирование 70
Контрольная группа:	предварительное тестирование 90		заключительное тестирование 80

Регрессия может привести к различным проблемам и вероятно именно она была виновна в проведении исследований, ставящих под сомнение эффективность хорошо известной программы *Head Start*. Этот случай будет рассматриваться в главе 10 для иллюстрации проблем, возникающих при осуществлении крупномасштабных федеральных программ.

### Тестирование и инструментарий

**Тестирование** рассматривается как угроза внутренней валидности, когда сам факт проведения предварительного тестирования влияет на послеэкспериментальные оценки. Если проводится многократное тестирование, то может появиться эффект

<sup>1</sup> Обратите внимание на использование выражения «может оказать», а не «оказывает», так как возможно, что оценки контрольной группы изменились с 90 до 80 под влиянием одного из факторов, угрожающих внутренней валидности. Если эти факторы будут исключены, регрессия станет определенной.

привыкания, а кроме того, некоторые аспекты доэкспериментального тестирования могут сделать участников более восприимчивыми к определенным особенностям программы. Например, если объектом эксперимента является компьютеризированный курс по общей психологии, который каждый может изучать с необходимой ему скоростью, то в ходе предварительного тестирования могут проверяться знания по данному предмету и участники могут обратить внимание на вопросы, о которых они ничего не знают, а в дальнейшем в ходе обучения уделять им больше внимания и в результате лучше справиться с заключительной проверкой.

**Инструментарий** становится проблемой, если между предварительной и заключительной проверкой измерительные инструменты как-либо изменяются. В упомянутом курсе для самостоятельного изучения по общей психологии предварительные и заключительные тесты состоят из разных заданий с одинаковым уровнем сложности. Если, к примеру, заключительный тест будет легче, то его результаты покажут увеличение знаний, не соответствующее действительности. Иногда инструментарий вызывает проблемы, если измерение включает наблюдение. Наблюдатели лучше справятся с работой, попрактиковавшись в ней, а это означает, что инструмент при заключительном тестировании будет отличаться от используемого при предварительном тестировании (в данном случае он будет более точным).

Так же как и проблемы, вызванные историей, взрослением и регрессией, связанные с инструментарием осложнения можно устранить с помощью контрольной группы. Единственным исключением является случай предварительной сенсбилизации, когда экспериментальная группа в ходе заключительного тестирования имеет некоторое преимущество перед контрольной группой, так как полученная при предварительном тестировании информация поможет им в ходе эксперимента сконцентрироваться на пробелах в своих знаниях, тогда как члены контрольной группы не будут иметь такой возможности.

### Предварительное оценочное тестирование: план Соломона

Иногда влияние предварительного тестирования можно оценить с помощью особого плана, называемого планом Соломона для четырех групп (Solomon, 1949). Он изображается следующим образом:

Экспериментальная группа 1:	предварительное тестирование	<i>экспериментальное воздействие</i>	заключительное тестирование
Экспериментальная группа 2:		<i>экспериментальное воздействие</i>	заключительное тестирование
Контрольная группа 1:	предварительное тестирование		заключительное тестирование
Контрольная группа 2:			заключительное тестирование

Сравнение двух экспериментальных групп позволяет определить влияние предварительного тестирования на экспериментальное воздействие, приводящее к изменениям результатов заключительного тестирования. Аналогично сравнение результатов тестирования членов контрольных групп показывает, воздействует ли само предварительное тестирование на выполнение заключительных тестов. Идеальный



результат использования плана Соломона при оценке программы по снижению экзаменационной тревоги выглядит следующим образом:

Экспериментальная группа 1:	предварительное тестирование 90	<i>экспериментальное воздействие</i>	заключительное тестирование 70
Экспериментальная группа 2:		<i>экспериментальное воздействие</i>	заключительное тестирование 70
Контрольная группа 1:	предварительное тестирование 90		заключительное тестирование 90
Контрольная группа 2:			заключительное тестирование 90

Результаты, свидетельствующие о том, что низкая тревожность при заключительном тестировании вызвана воздействием предварительного тестирования, будут следующими:

Экспериментальная группа 1:	предварительное тестирование 90	<i>экспериментальное воздействие</i>	заключительное тестирование 70
Экспериментальная группа 2:		<i>экспериментальное воздействие</i>	заключительное тестирование 90
Контрольная группа 1:	предварительное тестирование 90		заключительное тестирование 70
Контрольная группа 2:			заключительное тестирование 90

План Соломона — хороший способ оценки влияния предварительного тестирования, но затраты на его реализацию довольно велики. В сравнении с обычным планом, подразумевающим наличие экспериментальной и контрольной групп, он требует больше участников, а также проведения двух отдельных экспериментальных программ.

### **Проблемы, связанные с участниками**

Угроза внутренней валидности может также исходить от участников исследования. Кук и Кэмпбелл в 1979 г. выделили две связанные с этим проблемы.

#### **Эффект отбора испытуемых**

Характерной особенностью экспериментального исследования с управляемой независимой переменной является то, что его участники, находящиеся в разных условиях, эквивалентны по всем показателям, кроме независимой переменной. В следу-

ющей главе вы узнаете, как формировать эквивалентные группы при помощи методов случайного распределения и уравнивания. Если группы не эквивалентны друг другу, может возникнуть **эффект отбора испытуемых**. Предположим, что исследователь на примере двух разделов курса по общей психологии хочет сравнить обычный лекционный курс с другим, состоящим из лекций и дискуссионных занятий. Правила учебного заведения: а) не позволяют исследователю случайным образом разделить студентов на две группы, каждая из которых будет посещать один из курсов, а также б) требуют подробного разъяснения студентам сущности обоих курсов. Таким образом, студенты могут записаться на изучение любого из разделов — в этом и заключается трудность. Допустим студенты, посещающие курс с лекциями и дискуссиями, справились с ним лучше посещающих обычный лекционный курс. Что явилось причиной такой разницы? Связано ли это с организацией курса (наличие дискуссий) или с самими студентами, *выбравшими* этот курс? Может быть, эти студенты имели более развитые речевые способности (а следовательно, были более склонны к обсуждениям), чем другие студенты. Если кратко охарактеризовать сложившуюся ситуацию, то можно сказать, что возникло осложнение, связанное с отбором испытуемых для двух сравниваемых групп.

Эффект выбора может также взаимодействовать с другими факторами, угрожающими внутренней валидности. Например, в исследовании с двумя группами некоторое историческое событие может подействовать на одну группу и не подействовать на другую. В этом случае возникнет осложнитель история  $\chi$  выбор. Аналогично у двух групп могут быть разные показатели взросления и регрессии, они могут по-разному реагировать на факт тестирования и на них по-разному может влиять инструментарий.

Хорошим примером, демонстрирующим эффект выбора, является (к сожалению) одно из классических исследований по психологии. Это новаторское исследование в области психологии здоровья сегодня известно под названием «язва у "исполнительных" обезьян». Во вставке 5.3 вы найдете описание этого методологически несовершенного (а также сомнительного с этической точки зрения) эксперимента и его последствий.

#### ВСТАВКА 5.3

### **Классические исследования- проблема отбора и исполнительные обезьянки**

В апрельском номере за 1958 г. журнала «Сайентифик америкэн» была напечатана примечательная статья, озаглавленная «Язва у "исполнительных" обезьян». В ней описывались эксперименты Джозефа В. Брэди из Военного исследовательского института Уолтера Рида. Заинтересованный военными исследованиями, показавшими взаимосвязь между желудочной секрецией солдат во время медицинского осмотра и последующим развитием язвы желудка, Брэди решил исследовать взаимосвязь между эмоциональными и физическими расстройствами путем изучения эмоционального поведения обезьян, принадлежащих к виду макак-резус. К глубокому сожалению ученого, многие из обезьян умерли, а вскрытие показало, что большинство из них страдали язвой желудка. В основном эти обезьяны использовались для

формирования у них условного избегания. В ходе эксперимента их пристегивали к специальным стульям, позволяющим реагировать на идущий к ступням электрический разряд только нажатием рычага.

Брэди предположил, что развитие язвы могло быть связано или с повторяющимся разрядом, или со стрессом, вызванным процедурой выработки условного избегания. Чтобы выявить возможность развития язвы из-за воздействия электрического разряда, применив метод фальсификации, ученый разработал особый эксперимент. Он поместил двух обезьян на соседние стулья; одна из них, «исполнитель» (обратите внимание на аналогию с вечно занятым и испытывающим стресс исполнительным директором), могла избегать разряда, подаваемого каждые 20 секунд, нажав на рычаг в промежутке между двумя разрядами. У контрольной обезьяны рычаг не работал, и она испытывала разряд каждый раз, когда управляющая обезьяна пропускала 20 секунд и получала разряд. Таким образом, обе обезьяны испытывали разряд одинаково часто, но только одна из них могла контролировать ситуацию. Так продолжалось в течение 6 часов, а затем следовал 6-часовой отдых. Далее, через 24 часа эксперимент возобновлялся - проводились две попытки по 6 часов с двумя периодами 6-часового отдыха (при этом обезьяны ни разу не покидали стульев).

Через двадцать три дня (!) обезьяна-исполнитель умерла, после чего обнаружили, что у нее была язва желудка. У контрольной обезьяны язва не была обнаружена. Брэди повторил эксперимент с другой парой обезьян и получил те же результаты. В дальнейшем он привел отчет об исследовании с четырьмя парами животных (Brady, Porter, Conrad & Mason, 1958), в котором сделал вывод, что психологический стресс, вызванный ситуацией, когда приходится управлять не только собственной судьбой, но также и судьбой подчиненного, может привести к проблемам со здоровьем, в данном случае к язве желудка. Догадываетесь, почему армия профинансировала это исследование?

Описание данного исследования было приведено во многих учебниках по общей психологии, а публикация в «Сайентифик америкэн» привлекло к нему еще большее внимание. Однако тщательное изучение использованного Брэди метода показало, что в нем имело место осложнение из-за процедуры выбора субъектов. Брэди разделил обезьян на две группы не случайным образом, а на основании того, как быстро они освоили процедуру избегания. Сначала все восемь обезьян испытывались как исполнители, а затем те, которые реагировали быстрее, были сделаны исполнителями на время эксперимента. Брэди не знал, что животные различаются по присущей им эмоциональности и более эмоциональные быстрее реагируют на болевое воздействие, поэтому он невольно сделал исполнителями более эмоциональных животных (т. е. предрасположенных к язве желудка), а более спокойных поместил в контрольные условия.

Первым на эту проблему указал Вейсс в 1968 г. Его исследования с крысами были лучше организованы (например, Weiss, 1977) и показали результаты, *противоположные* результатам Брэди. Как вы узнаете из рассказа об эквивалентных контрольных группах в главе 7, Вейсс обнаружил, что язва развилась у меньшего количества животных, имеющих возможность контролировать электрический разряд, и у большего количества животных, не имеющих такой возможности.

Еще один аспект исследования Брэди, который вы, вероятно, уже отметили, заключается в том, что сегодня Комиссия по использованию животных потребует существенных изменений в процедуре исследования. Изучение взаимосвязи между стрессом и физическими расстройствами является жизненно важным, но его можно исследовать с помощью менее жестокой процедуры, когда обезьян не будут привязывать к стульям и оставлять практически бездвиженными в течение нескольких недель.

### Истощение

Испытуемые не всегда участвуют в экспериментах вплоть до их завершения. Некоторые исследования могут проводиться довольно долгое время, и люди могут прекратить свое участие, потерять интерес или даже умереть. В ходе некоторых исследований участники могут испытывать дискомфорт и воспользоваться правом отказа от участия в дальнейших экспериментах. Поэтому вследствие различных причин в начале исследования в нем могут участвовать 100 человек, а в конце остаться только 60. Эту проблему иногда называют «смертностью» испытуемых, но я предпочитаю не использовать столь печальное выражение и вместо него применяю термин **истощение**. Истощение является проблемой, так как если одни люди более склонны к тому, чтобы бросить исследование, чем другие, то в конце исследования состав участников качественно будет отличаться от исходного. Данная проблема сходна с проблемой выбора, поскольку в результате группа, начинающая исследование, становится неэквивалентной группе, завершающей его. Чтобы проверить, различаются ли продолжающие исследование участники и покидающие его, можно сравнить оценки предварительного тестирования или другие показатели, полученные до начала экспериментов, у обеих групп. Если покинувшие исследование и продолжающие участвовать в нем испытуемые не различались в начале, то общие выводы в конце исследования не пострадают, даже несмотря на потери в связи с истощением.

На этом мы завершаем введение в экспериментальный метод. В трех следующих главах будут подробно рассматриваться вопросы, связанные с этой темой. Глава 6 откроется описанием различий межсубъектных и внутрисубъектных планов (или планов с повторяющимися измерениями). Далее в ней будут обсуждаться проблемы, связанные с управлением экспериментальными исследованиями, в том числе вопросы создания эквивалентных групп для экспериментов с межсубъектными планами, контроля за эффектом последовательности для внутрисубъектных планов, а также искажения, возникающие из-за того, что и экспериментаторы и участники — это живые люди. Главы 7 и 8 посвящены рассмотрению различных планов исследований, начиная от планов с одной независимой переменной (глава 7) и заканчивая планами с множеством независимых переменных, известными как факторные планы (глава 8).

## Резюме

### Основные черты экспериментальных исследований

Психологический эксперимент включает в себя задание независимых переменных, контроль за внешними переменными и измерение зависимых переменных. Независимые переменные возникают при проведении сравнений и при создании экспериментальных условий, находящихся под непосредственным контролем исследователя. Управляемые независимые переменные подразумевают помещение участников в разные условия, предложение им разных заданий или разных инструкций для выполнения заданий. Внешние переменные — это факторы, которые сами по себе не интересуют исследователя. Невозможность управлять ими приводит к проблеме, называемой осложнением. При наличии осложнителя результаты могут быть

вызваны как независимой переменной, так и осложнителем. Зависимые переменные — это характеристики поведения, измеряемые в ходе исследования. Они должны быть точно (операционально) определены.

### **Управляемые и субъективные переменные**

В некоторых психологических исследованиях сравниваются группы участников, определенным образом отличающихся друг от друга еще до начала эксперимента (например, могут различаться пол, возраст, интровертированность). В таком случае интересующая исследователя независимая переменная выбирается для эксперимента, а не регулируется непосредственно и называется субъективной переменной. Психологические исследования часто проводятся с использованием и управляемых, и субъективных переменных. В исследованиях, достаточно хорошо контролируемых экспериментатором, выводы о причинах и следствиях могут быть сделаны только при использовании управляемых переменных, в исследованиях с субъективными переменными это невозможно.

### **Валидность экспериментальных исследований**

Есть четыре стороны валидности психологических исследований. В валидном исследовании правильно применяется статистический анализ (валидность статистических выводов), адекватно определяются независимые и зависимые переменные (конструктивная валидность), его результаты можно распространить за пределы проведенного эксперимента (внешняя валидность), а кроме того, оно не содержит осложняющих переменных (внутренняя валидность).

### **Угроза для внутренней валидности**

Угрозу для внутренней валидности эксперимента могут представлять различные факторы. В плохо контролируемых исследованиях, включающих сравнение результатов предварительного и заключительного тестирования, часто появляются такие осложняющие факторы, как история, взросление, регрессия, тестирование и инструментарий. Проблема отбора возникает, когда сравниваются две неэквивалентные группы участников, различия между которыми существуют еще до начала исследования (например, как в исследовании Брэди, посвященном изучению образования язвы желудка у «исполнительных» обезьян). Так же проблема отбора может взаимодействовать с другими факторами, угрожающими внутренней валидности. В долгосрочных исследованиях истощение может привести к возникновению особой разновидности проблемы отбора — небольшая группа участников, оставшихся к концу исследования, может заметно отличаться от более многочисленной группы, набранной в начале.

## **Задания для повторения**

### **Выбор ответа**

1. В главе 4 в качестве примера использования номинальной шкалы было описано исследование Шехтера, показавшее, что ожидающие воздействия сильного электрического шока люди предпочитают ждать вместе с другими. Ис-

следование можно подвергнуть критике за то, что все его участники были женщинами. Это будет критикой:

- 1) конструктивной валидности исследования;
  - 2) внешней валидности;
  - 3) внутренней валидности;
  - 4) экологической валидности.
2. Для участия в программе по снижению повышенного кровяного давления отобрали 50 добровольцев, которым остался один год до выхода на пенсию. Через год после выхода людей на пенсию все 50 человек продолжали участвовать в программе, а их кровяное давление снизилось. Какой угрожающий внутренней валидности фактор затрудняет интерпретацию данных этого исследования?
- 1) истощение;
  - 2) отбор испытуемых;
  - 3) инструментарий;
  - 4) история.
3. В исследовании памяти экспериментатор хочет показать, что использование зрительных образов улучшает запоминание. Одних участников попросили придумывать образы, а других — просто повторять каждое слово по два раза. Поскольку формирование образов занимает больше времени, чем повторение слов, создающей образы группе слова демонстрировали с более низкой скоростью (4 секунды на каждое), чем группе, повторяющей слова (2 секунды на слово). Какое из следующих утверждений верно?
- 1) скорость показа осложняется инструктивной независимой переменной;
  - 2) скорость показа является независимой переменной;
  - 3) исследование обладает внутренней валидностью, но не обладает внешней;
  - 4) исследование хорошо спланировано. Инструктивная переменная является управляемой независимой переменной.
4. Какое из следующих утверждений об известном исследовании с куклами бобо верно?
- 1) все независимые переменные были управляемыми;
  - 2) исследователи управляли фактором «бобо», используя одну и ту же надувную куклу как для взрослых, так и для детей;
  - 3) дети имитировали модели, только если они были представлены в виде мультфильмов;
  - 4) управляемая независимая переменная являлась ситуативной переменной.
5. В исследовании программы снижения тревоги средняя предварительная оценка в экспериментальной группе была 79 (максимальная = 100), а средняя заключительная — значительно более низкой: 68. Наиболее вероятно, что такое изменение было вызвано:
- 1) регрессией;

- 2) взрослением;
- 3) эффективностью программы;
- 4) любой из вышеперечисленных факторов (или их комбинация) мог вызвать подобное изменение.

### Короткие эссе

1. На примере исследования тревожности проиллюстрируйте разницу между а) управляемыми и б) субъективными независимыми переменными.
2. На различных примерах покажите разницу между ситуативными, рабочими и инструктивными независимыми переменными.
3. Что такое осложнитель и почему его наличие затрудняет интерпретацию результатов исследования?
4. Считается, что в исследовании с использованием субъективных переменных выводы о причинах явлений не могут быть сделаны. Почему?
5. Опишите условия, при которых может снизиться валидность статистических выводов эксперимента.
6. Опишите три вида условий, при которых может снизиться внешняя валидность.
7. Охарактеризуйте различия внутренней и внешней валидности исследования.
8. Объясните, каким образом наличие контрольной группы может снизить угрозу внутренней валидности. В качестве примера используйте историю, взросление или регрессию.
9. Объясните, как можно использовать план Соломона дня оценки влияния предварительного тестирования участников.
10. На примере исследования Брэди проиллюстрируйте проблему отбора.

## Упражнения

### Упражнение 5.1. Выявление переменных

В каждом из перечисленных ниже исследований выявите независимую (независимые) переменную (переменные), значения независимой (независимых) переменной (переменных) и зависимую (зависимые) переменную (переменные). Определите, являются ли независимые переменные управляемыми или неуправляемыми. Укажите, какая шкала использовалась для измерения зависимых переменных.

1. В исследовании когнитивных карт сравнивалась способность первокурсников и студентов старших курсов точно указывать местоположение зданий кампуса. Некоторые здания располагаются в центре кампуса поблизости от дорог, по которым часто ходят студенты, а другие — на периферии студенческого городка. Участников исследования попросили оценить (по шкале от 1 до 10) степень уверенности в том, что они правильно указали направление, и записать количество ошибок (в процентах).

2. В исследовании эффективности нового лекарства для лечения депрессии одни участники принимали его, а другие думали, что принимают. Третья группа участников не подвергалась экспериментальному воздействию. По окончании программы участники заполнили Опросник депрессии Бека, а затем опытные наблюдатели оценили (по 10-балльной шкале) уровень их депрессии.
3. В исследовании обусловливания, проведенном И. П. Павловым, у голодных животных вырабатывался условный рефлекс слюноотделения при ЗВУКОВОМ сигнале, связанном с получением пищи. Одним животным, прежде чем выдать еду, предъявляли сигнал, а затем снова выключали его. На других животных воздействовали сигналом до самого кормления. Третья группа животных получала еду до включения сигнала. Экспериментаторы отмечали начало саливации и ее объем для определенного промежутка времени.
4. В исследовании по психолингвистике развития двух-, трех- и четырехлетним детям показывали кукол и просили разыгрывать отдельные сценки, чтобы определить, способны ли они использовать определенные грамматические правила. В одних случаях детей просили разыгрывать сценки, используя активный залог (Эрни побил Берта), а в других — используя пассивный залог (Эрни был побит Бертом). Оценивалось, правильно ли дети разыгрывают сценки (две оценки) и как быстро они приступают к игре.
5. В исследовании способностей к запоминанию лабиринта одни крысы на протяжении 30 попыток получали корм после прохождения лабиринта, другие корма не получали, третьих не кормили на протяжении первых 15 попыток, но кормили в конце следующих 15 раз, а последнюю группу крыс кормили первые 15 раз и не кормили вторые 15 попыток. Исследователь отмечал все сделанные ошибки (неправильные повороты) и время прохождения лабиринта.
6. В исследовании, посвященном готовности к оказанию помощи, к посетителям торгового центра подходит либо хорошо одетый студент, либо неряшливый и спрашивает, где находится туалет или ближайший магазин *Kmart*. Расположившийся неподалеку экспериментатор отмечает, была ли оказана помощь.

### **Упражнение 5.2. Обнаружение осложнителя (осложнителей)**

Для каждого из приведенных ниже исследований определите независимые и зависимые переменные, значения, принимаемые независимыми переменными, а также укажите хотя бы одну внешнюю переменную, контроль за которой не был адекватным (т. е. которая создала осложнитель). Для выполнения задания воспользуйтесь подходом, показанным в табл. 5.2.

1. Фирма, проводящая тестирование, пытается выяснить, можно ли новой клюшкой для гольфа (клюшка 1) посылать мяч дальше, чем это делают с помощью клюшек трех других конкурирующих торговых марок (клюшки 2-4). Для этого набрали 20 добровольцев — мужчин, профессионально играющих в гольф. Каждый гольфер наносит 50 ударов клюшкой 1, затем 50 клюшкой 2, далее 50 клюшкой 3 и 50 клюшкой 4. Для большего реализма эксперимент



проходит на настоящей площадке для гольфа. Используются первые четыре лунки — участник совершает первые 50 ударов от первой метки для мяча, затем 50 от второй и т. д. Первые четыре лунки находятся на расстоянии 380-400 ярдов, попадание в каждую дает 4 очка.

2. Исследователь хочет определить способность страдающих шизофренией пациентов определять длительность различных событий. Выдвинута гипотеза, что сильный шум будет снижать точность. Проверка участников проводится двумя способами: одних участников исследуют в условиях «тишины» в маленькой звуконепропускаемой комнате, которую используют для проверки слуха, а других — в комнате медицинского персонала, где играет музыка с постоянной (и довольно высокой) громкостью. По причине существующего расписания пациенты, которым не разрешено самостоятельно покидать палаты (т. е. немного более опасные), могут участвовать в исследовании *только* в понедельник, а те, которые могут свободно передвигаться (т. е. немного менее опасные), — *только* в четверг. Кроме того, в четверг проводится проверка слуха, поэтому звуконепропускаемая комната может использоваться только в понедельник.
3. Экспериментатор хочет определить, улучшится ли память у людей, использующих визуальные образы. Участники (женщины) разделяются на две группы — в первой используют технику создания зрительных образов, а во второй запоминают путем механического повторения. Первой группе предлагают для изучения список из 20 существительных, обозначающих конкретные понятия (для них легче, чем для абстрактных, сформировать образы), а второй — 20 абстрактных слов (таких, которые особенно легко произносить, а следовательно, повторение не составит труда), имеющих такую же частоту использования, как и конкретные слова. Для согласования метода представления слов со способом их запоминания участникам из первой группы слова демонстрируют визуально (на экране компьютера). Для контроля возможной «компьютеробии» участники из второй группы также сидят у экранов компьютеров, запрограммированных на чтение слов. После прослушивания соответствующих наборов слов участникам дается 60 секунд на их воспроизведение в произвольном порядке.
4. Социальный психолог изучает готовность к оказанию помощи другим людям. Он познакомился с двумя аспирантами, готовыми участвовать в проведении экспериментов. Первый (Нед) хорошо одевается, а второй (Тед) не слишком заботится о своей внешности. В ходе эксперимента к посетителям торгового центра подходит либо хорошо одетый Нед, либо потрепанный Тед и просит двадцатипятицентовик на кофе. Находящийся поблизости экспериментатор записывает, дают ли покупатели деньги. Исследование проводится с 8 до 9 часов вечера, причем Нед работает в понедельник, а Тед — в четверг.

### **Упражнение 5.3. Операциональные определения**

В главе 3 вы познакомились с операциональными определениями и выполнили упражнение на формулировку таких определений для нескольких конструктов, используемых в психологических исследованиях. В данном упражнении вы буде-

те играть роль экспериментатора, разрабатывающего исследование. Для каждой из четырех приведенных ниже гипотез:

- а) определите независимую (независимые) переменную (переменные), решите, сколько значений независимой (независимых) переменной (переменных) вы будете использовать, и задайте эти значения;
- б) определите зависимую переменную для каждого исследования и
- в) дайте операциональные определения независимым и зависимым переменным.
  1. Люди скорее помогут нуждающемуся, если необходимость помощи не будет вызывать сомнений.
  2. Способность концентрироваться на задании снижается, если человек чувствует себя в тесноте из-за окружения других людей.
  3. Хорошие игроки в боулинг лучше играют в присутствии зрителей, а посредственные — хуже.
  4. Животные быстрее запоминают сложный лабиринт, если они испытывают умеренное возбуждение. Они справляются с заданием хуже при сильном или слабом возбуждении. Если лабиринт легкий, способность к его прохождению постепенно повышается при нарастании возбуждения.