

Экспериментальный план I: однофакторные планы

Обзор задач главы

Главы 5 и 6 образуют сцену для этой и следующих глав. В главе 5 мы познакомились с экспериментальным методом, получили представление о независимых, внешних, и зависимых переменных, изучили проблему осложнений и обсудили факторы, связанные с валидностью психологических экспериментов. В главе 6 мы сравнили межсубъектные и внутрисубъектные планы, изучили основные способы контроля за ними (например, случайное распределение и позиционное уравнивание), а также рассмотрели проблему искажений, возникающих в психологических исследованиях. Сцена готова, и очень скоро на ней будет разыграно представление этой и следующих глав. Они расскажут вам о различных экспериментальных планах, на основании которых в психологии строятся экспериментальные исследования. В данной главе рассматриваются планы с одной независимой переменной, принимающей одно или несколько значений. Если используется более одной независимой переменной, то план называется факторным — такой план рассматривается в главе 8. Изучив данную главу, вы:

- сможете определить четыре вида однофакторного плана: планы с независимыми, уравненными и неэквивалентными группами, а также план с повторяемыми измерениями;
- научитесь понимать, когда нужно применять проверку по F -критерию Стьюдента для независимых и зависимых групп;
- научитесь понимать, когда нужно использовать различные виды контрольных групп: группы плацебо и листа ожидания, а также сцепленную контрольную группу;
- поймете этические аспекты использования контрольных групп;
- сможете назвать две главные причины использования более чем двух значений для независимой переменной;
- сумеете решать, когда нужно использовать гистограмму, а когда линейный график;
- поймете, почему при анализе данных однофакторного многоуровневого эксперимента не стоит применять многократную проверку по критерию Стьюдента, а следует использовать метод *ANOVA*.

В главе 3 при обсуждении научной креативности в качестве примера была рассказана история происхождения исследования с лабиринтом. Исследование Смолла, в котором использовалась модель лабиринта Хэмптон Корт, было первым из целой лавины экспериментов с лабиринтами, проведенных в течение первых двух десятилетий XX в. В основном в ранних исследованиях ученые стремились выяснить, какое из чувств крысы наиболее важно для процесса обучения. Вы, вероятно, помните, как в главе ? (вставка 2.3) рассказывалось о проблемах, возникших у Джона Уотсона с борцами против вивисекции из-за серии исследований, в ходе которых он хирургическим путем последовательно лишал крыс одного чувства за другим. Уотсон обнаружил, что способность к изучению лабиринта не снижается, даже если крыса лишается большинства чувств, и сделал вывод, что для запоминания лабиринта крысы пользуются мускульным, или кинестетическим, чувством: они запоминают, что нужно сделать определенное количество шагов, а затем повернуть направо и т. д. Для проверки этой идеи он вместе со своим коллегой из Чикагского университета Харви Карром провел простое, но изящное исследование (Carr & Watson, 1908). После того как группа крыс изучила сложный лабиринт, Карр и Уотсон удалили центральную часть лабиринта, сделав один из его участков короче. Они предположили, что крысы, натренированные проходить более длинный лабиринт, будут буквально наткаться на стены укороченного лабиринта. И действительно, в описании поведения одной из крыс Карр и Уотсон отметили, что она «на полном ходу наткалась на стены. Сильно шаталась и не вернулась к нормальному поведению, пока не прошла (еще) 9 футов» (р. 39). Вторая группа крыс обучалась на более коротком лабиринте, а затем переходила к более длинному. Эти крысы вели себя сходным образом, нередко слишком рано поворачивали и наткались на стены, по всей видимости, ожидая найти поворот. Спустя многие годы после ухода из академической науки, Джон Уотсон назвал это исследование одним из важнейших в своей жизни. Последующие эксперименты с использованием лабиринтов поставили под сомнение кинестетическую гипотезу, но для нас в этой истории важно то, что хорошее исследование не всегда нуждается в чрезвычайно сложном экспериментальном плане. В некоторых случаях двух групп вполне достаточно.

Один фактор — два уровня

Как показано на рис. 7.1, изображающем дерево принятия решений, есть четыре вида плана с одной независимой переменной, которая принимает два значения (т. е. для которой имеется два уровня). К каждому из этих видов приводит последовательность решений, принятых в отношении независимой переменной. Во-первых, эта переменная может быть внутрисубъектной или межсубъектной. Если ее изучают как межсубъектную, она может быть управляемой или субъектной. При управляемой независимой переменной план будет называться **планом с независимыми группами**, если для создания эквивалентных групп используется случайное распределение, и **планом с уравненными группами**, если для этого используется уравнивание, а затем случайное распределение. Как вы узнали в главе 6, чтобы при-

нять решение об использовании уравнивания, необходимо учитывать размер выборки и особенно внимательно относиться к внешним переменным, коррелирующим с зависимой переменной. Если изучается субъектная переменная, группы формируются из разных категорий людей (например, женщины/мужчины, интроверты/экстраверты, либералы/консерваторы). Экспериментальный план при этом иногда называют «*ex post facto*», так как группы формируются после установления наличия у испытуемых определенных характеристик. Такой план также называют планом со «стихийно возникшими группами», или **планом с неэквивалентными группами** (я буду использовать именно этот термин). Поскольку при таком плане группы строятся из разных типов людей, исследователи нередко пытаются по возможности снизить неэквивалентность, уравнивая группы по различным факторам. Например, в исследовании с неэквивалентными группами, в котором сравниваются женщины и мужчины, можно сделать так, чтобы члены обеих групп были одного возраста и принадлежали к одному социоэкономическому классу.

Последний вид однофакторного плана — это **план с повторяемыми измерениями**. Он используется, если независимая переменная является внутрисубъектной и каждый участник исследуется при каждом значении независимой переменной (т. е. измерения повторяются несколько раз). Важнейшие особенности каждого из четырех основных видов экспериментального плана представлены в табл. 7.1. Далее мы приступим к рассмотрению конкретных примеров.

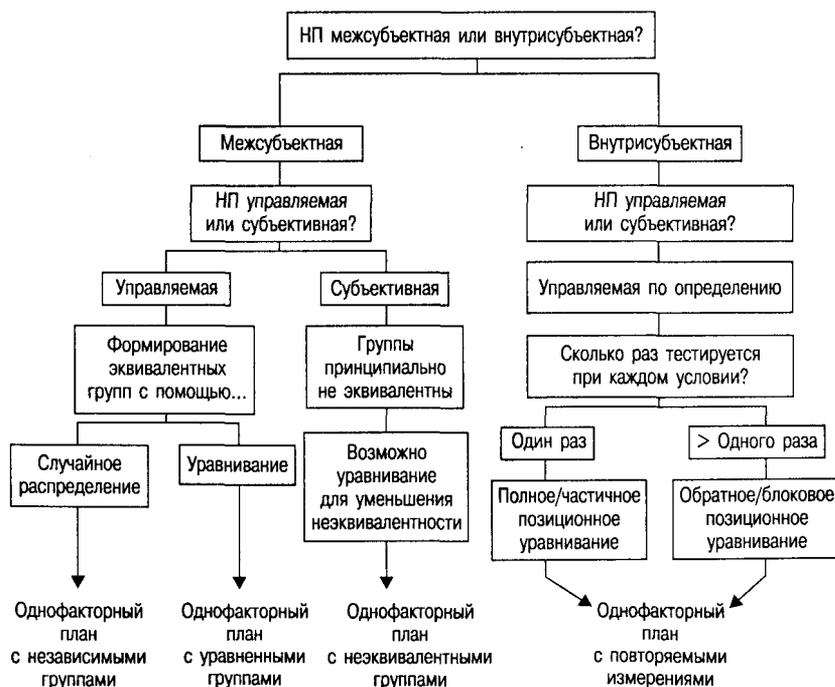


Рис. 7.1. Древо решений - однофакторный план

Таблица 7, 1

Особенности четырех однофакторных планов

Вид плана	Минимальное количество значений, принимаемых независимой переменной	Межсубъектная или внутри-субъектная независимая переменная	Вид независимой переменной	Способ создания эквивалентных групп
С независимыми группами	2	Межсубъектная	Управляемая	Случайное распределение
С уравненными группами	2	Межсубъектная	Управляемая	Уравнивание
С неэквивалентными группами	2	Межсубъектная	Субъектная	Уравнивание может снизить неэквивалентность
С повторяемыми измерениями	2	Внутри-субъектная	Управляемая	Отсутствует

Межсубъектные однофакторные планы

Однофакторные двухуровневые исследования проводятся не так часто, как вы можете решить. Большинство исследователей предпочитают использовать более сложные планы, позволяющие получить более подробные и интересные результаты. Кроме того, мало кто из редакторов журналов достаточно высоко оценит такое простое исследование. Но несмотря на это, в простоте есть особая прелесть, а что может быть проще сравнения двух условий? Ниже приводятся три примера таких исследований.

Пример 6. Независимые группы

Примером эксперимента, проведенного по однофакторному плану с независимыми группами, является хорошо известное исследование Блэкмора и Купера (Blackmore & Cooreg, 1970). Ученые заинтересовались вопросом влияния опыта на развитие зрительной системы. Двухнедельных котят случайным образом распределили по двум значениям независимой переменной, которую можно обозначить как «зрительное окружение». Котят растили в условиях с преобладанием либо вертикальных либо горизонтальных полос. Я думаю, вы понимаете, почему исследование должно было быть межсубъектным, а не внутрисубъектным — было бы бессмысленно растить котят в «вертикальном» окружении, а *затем* в «горизонтальном». В подобных исследованиях испытуемые, принявшие участие в эксперименте с одним значением независимой переменной, уже по сути «использованы» и полученный опыт не дает им возможности «начать сначала» с другим условием.

На рис. 7.2 показан выполненный Блэкмором и Купером набросок устройства для создания «вертикального» окружения. Кот стоит на площадке из оргстекла, с которой видно, как полосы уходят вверх и вниз. Широкий воротник на шее жи-

вотного заставляет его зрительно фокусироваться на стенах камеры. В течение нескольких месяцев котов помещали в «вертикальный» или «горизонтальный» миры на 5 часов в день; остальное время они находились в темноте.

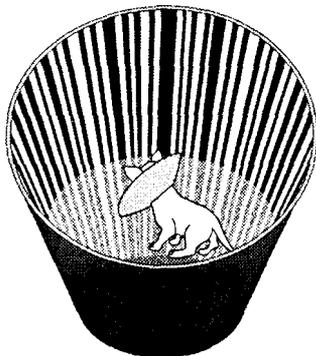


Рис. 7.2. Устройство для эксперимента Блэкмора и Купера (Blakemore & Cooper, 1970)

В конце эксперимента Блэкмор и Купер исследовали поведение животных и измерили у них активность нейронов зрительной зоны коры головного мозга. В целом коты быстро оправились от депривации. «Через 10 часов пребывания в нормальных зрительных условиях они... смогли с легкостью прыгать со стула на пол» (Blackemore & Cooper, 1970, p. 477). Однако коты, выросшие в «вертикальном» окружении, не очень хорошо воспринимали явления, происходящие в горизонтальной плоскости, а вертикальные стимулы вызывали проблемы у животных, привыкших к «горизонтальным» условиям:

Разница стала особенно заметна, когда двух котят, из которых один имел «горизонтальный», а второй — «вертикальный» опыт, одновременно исследовали с помощью длинной белой или черной рейки. Если рейку держали вертикально и трясли из стороны в сторону, один из котят подбегал и начинал играть с ней. Если рейку держали горизонтально, это привлекало другого котенка, а первый оставался безразличным.

Blackemore & Cooper, 1970, p. 478

Очевидно, что опыт первых месяцев жизни сильно влияет на развитие мозга.

Пример 7. Уравненные группы

В старых фильмах иногда показывают, как героя держат взаперти, пытаются и лишают сна на 2 или 3 дня, чтобы получить от него определенные сведения. Может ли депривация сна повлиять на ответы, полученные от человека в ходе допроса? Такой эмпирический вопрос поставил перед собой Блэгров (Blagrove, 1996) при проведении интересного исследования с уравненными группами. Выражаясь более точно, он хотел узнать, как подействуют на людей, лишенных сна, вводящие в заблуждение вопросы. Исследователь набрал студентов колледжа для проведения трех различных экспериментов, в каждом из которых должны были участвовать по две группы — одни участники подвергались депривации сна, а другие — нет. Лишенные сна сту-

денты оставались в лаборатории и бодрствовали по 21 часу подряд в ходе первых двух экспериментов и 43 часа в третьем. Постоянное наблюдение «20 смен ассистентов» (р. 50) гарантировало, что испытуемые не спят. Студенты, которых не подвергали депривации, могли спать дома. Переменной уравнивания была «обычная продолжительность сна, сообщаемая испытуемыми» (р. 50). Блэгров хотел уравнивать обычную продолжительность сна, чтобы иметь возможность «контролировать у групп с одинаковой продолжительностью сна различия личностных и связанных с продолжительностью сна характеристик» (р. 50). Средняя продолжительность сна составила 8,4 и 8,5 в первом исследовании, 8,3 и 8,1 во втором и 8,4 и 8,1 в третьем.

Все участники исследования выполнили стандартизованный тест на внушаемость: они прослушали рассказ, а затем ответили на наводящие вопросы (т. е. такие вопросы, на которые нельзя ответить прямо, используя информацию из рассказа). Ответив на вопросы, они получили негативную оценку выполнения задания, после чего их попросили еще раз ответить на те же вопросы, чтобы увидеть, изменят ли они какие-либо из ответов. В целом, на лишенных сна испытуемых вопросы действовали сильнее и они чаще меняли свои ответы. Особенно ярко это проявилось в третьем исследовании, в котором депривация сна продолжалась 43 часа. Процедура уравнивания помогла создать группы, сходные по продолжительности сна, что позволило объяснить различия между группами именно депривацией.

Пример 8. Неэквивалентные группы

Вероятно, под влиянием мегалонгитюдного исследования Термана по изучению одаренных детей (см. вставку 6.1), Кнеппер с соавторами (Knepper, Obrzut & Copeland, 1983) провели заслуживающее внимания исследование, в котором попытались пролить свет на личностные особенности одаренных детей. Исследователи поставили перед собой вопрос, успешнее ли одаренные дети помимо познавательных задач решают социальные и эмоциональные проблемы в сравнении с обычными детьми. Их эксперимент отлично иллюстрирует план с неэквивалентными группами. Независимой субъективной переменной была степень одаренности. Сравнивались два ее значения: одаренный (операционально определенное как $IQ = 130$ и выше) и средний (IQ между 90 и 110). Среднее арифметическое значение IQ составило 136,9 и 102,9 соответственно. Никакой особой процедуры уравнивания использовано не было, но возраст контролировался — в исследовании принимали участие только шестиклассники. Использовался тест *Means-Ends Problem Solving Test*, оценивающий качество решений межличностных (социальных) и внутриличностных (эмоциональных) проблем. Одаренные дети действительно показали более высокие результаты, чем обычные, — это открытие совпадает с выводом Термана о том, что одаренные дети не просто «умные», а имеют также и социальные навыки.

Необходимо сделать одно важное замечание. Вспомните, как в главе 5 рассказывалось о том, что вывод о наличии причинно-следственной связи нельзя сделать, если используются субъектные переменные. Поэтому неверно говорить, что одаренность каким-либо образом вызывает увеличение способности к решению социальных и эмоциональных проблем. Можно лишь сказать, что у одаренных и неодаренных детей способности к решению подобных проблем различаются.

Внутрисубъектные однофакторные планы

Как вы уже знаете, внутрисубъектные планы: а) требуют меньшего количества участников, б) более чувствительны к небольшим различиям в значениях среднего арифметического и в) обычно предполагают использование позиционного уравнивания для устранения проблем последовательности. В случае внутрисубъектного плана с одной независимой переменной, принимающей два значения, можно использовать два вида позиционного уравнивания. Если испытуемые участвуют в экспериментах один раз при каждом условии, применяется полное позиционное уравнивание. Тогда половина участников выполняет задания при условии А, а затем В, а вторая половина — при условии В, а затем А. Если участники исследуются более одного раза при каждом условии, можно применить обратное позиционное уравнивание (АВВА). Именно этот подход использовал Дж. Ридли Струп в первом из трех проведенных им исследований, результаты которых были опубликованы в 1935 г. В списке «10 лучших классических исследований» его исследование заняло бы далеко не последнее место. Чтобы получить о нем более подробную информацию (и побольше узнать о том, что такое свастика), прочитайте вставку 7.1.

ВСТАВКА 7.1

Классические исследования — наиболее часто воспроизводимое психологическое открытие

Именно обратное позиционное уравнивание Дж. Ридли Струп использовал в своей работе, впервые опубликованной им в 1935 г. Это исследование настолько известно, что феномен, который был в нем продемонстрирован, сегодня называется «эффектом Струпа». В сопроводительной статье к повторной публикации данной работы Колин Мак-Леод назвал эффект Струпа «золотым стандартом» измерения внимания и начал свое эссе такими словами: «практически невозможно найти человека, занимающегося когнитивной психологией и при этом не получившего хотя бы краткого знакомства с эффектом Струпа. Эти слова, вероятно, относятся ко всем, кто прошел стандартный вводный курс по психологии - в нем тест Струпа демонстрируется практически обязательно». (MacLeod, 1992, p. 12).

Далее Мак-Леод указал, что эффект Струпа в психологии - это одно из наиболее часто воспроизводимых и упоминаемых открытий. В чем же состоит это открытие?

В исследовании обобщались три эксперимента, проведенные Струпом в ходе работы над докторской диссертацией. Мы остановимся на первых двух, поскольку они иллюстрируют внутрисубъектный план с одной независимой переменной, принимающей два значения, и поскольку в них использовалось обратное позиционное уравнивание. В первом эксперименте 14 женщин и 56 мужчин выполняли два задания, каждое из которых включало чтение названий цветов. Первое условие Струп назвал ЧНЦ («чтение названий цветов, напечатанных черным») (Stroop, 1992, p. 12). Участники должны были прочесть 100 названий цветов (например, ЗЕЛЕНЬИЙ), напечатанных черными чернилами, так быстро и точно, как только возможно. Второе условие Струп обозначил как ЧНЦр («чтение названий цветов, когда цвет чернил и название различны») (p. 16). В этом случае 100 названий цветов были напечатаны цветными чернилами, но при этом цвет чернил не соответствовал названию цвета (например, слово ЗЕЛЕНЬИЙ было напечатано красным). Задание состояло в чтении слова (т. е. правильный ответ - «зеленый»).

Как хороший исследователь, Струп знал о проблеме последовательности и поэтому для ее устранения использовал обратное позиционное уравнивание (АВВА). Разделив каждый из списков стимулов на отрезки, состоящие из 50 слов, Струп выдал половине участников последовательность ЧНЦч-ЧНЦр-ЧНЦр-ЧНЦч, а второй половине - последовательность ЧНЦр-ЧНЦч-ЧНЦч-ЧНЦр. Таким образом, каждый из участников должен был в целом прочитать 200 названий цветов. В первом эксперименте Струп не обнаружил *никаких различий* между выполнением заданий с условиями ЧНЦч и ЧНЦр. Среднее время, требующееся для прочтения 100 слов каждого вида, составило 41,0 и 43,3 секунды соответственно. Таким образом, на чтении цветных названий в условии ЧНЦр не сказалось то, что слова были напечатаны контрастным цветом. Но во втором эксперименте Струп обнаружил огромные различия, и это впоследствии сделало его имя знаменитым. Используя тот же общий план эксперимента, он изменил задание - на этот раз необходимо было называть цвета, а не читать их названия. При одном условии - НЦ («тест на название цвета»), - участники называли цвет образцов квадратной формы. При втором, ключевом условии - НЦСр («тест на название цвета слова, когда цвет чернил и слово различны») - участникам показывали тот же материал, что и при условии ЧНЦр из эксперимента 1, но в этот раз, вместо того чтобы читать название цвета, они должны были называть цвет, которым напечатано слово. Если слово ЗЕЛЕНый было напечатано красным, правильный ответ был «красный», а не «зеленый». В 1935 г. участники встретились с теми же трудностями, что переживают испытуемые сегодня. Поскольку чтение представляет собой в высшей степени глубоко усвоенный, автоматический процесс, он мешает называнию цвета, что приводит к ошибкам и замедлению выполнения задания. Струп обнаружил, что среднее время называния цветов составило 63,3 секунды для условия НЦ и (поразительно) 110,3 секунды для условия НЦСр. Из четырех результатов, представленных Струпом в виде таблиц, я построил гистограмму, показанную на рис. 7.3. Хорошо видно, что эффект Струпа весьма силен.

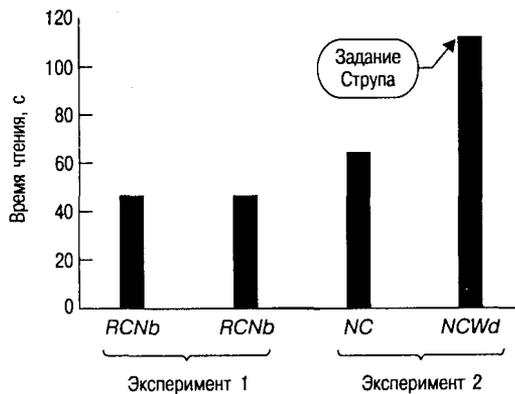


Рис. 7.3. Данные первых двух экспериментов из исследования Струпа

Как я уже отмечал, в действительности в ходе работы над диссертацией Струп провел три эксперимента. Третий из них показывает, что при наличии тренировки испытуемые могут улучшить выполнение задания НЦСр (классическое задание Струпа). Интересная особенность последнего эксперимента заключалась в том, что цветные квадратные образцы из НЦ-теста Струп заменил цветными образцами в форме свастики, что «позволило печатать НЦ-тест, используя оценки, более близкие к использованным в НЦСр-тесте» (Stroop, 1992, р. 18). Свастика - это древний религиозный символ, который получается, если загнуть перекладины традиционного греческого креста (+). Так вышло, что исследование Струпа было опубликовано в том же году (1935), когда свастика стала официальным символом нацистской Германии.

Другой способ позиционного уравнивания в случае исследования с двумя условиями, при каждом из которых участники исследуются несколько раз, — простое чередование условий (АВАВ...). Этот подход рассматривается в следующем примере.

Пример 9. Повторяемые измерения

Для проверки нескольких предположений, сделанных на основании выдвинутой Джеймсом Гибсоном теории восприятия, Ли и Аронсон (Lee & Aronson, 1974) провели исследование, посвященное восприятию движения и чувству равновесия (о Джеймсе Гибсоне, муже Элеанор Гибсон, упоминалось в главе 1). В частности, они интересовались тем, как мы сохраняем равновесие в движущейся среде. Исследователи помещали детей в возрасте от 13 до 16 месяцев в аппарат, изображенный на рис. 7.4. Когда ребенок смотрел на дальнюю стену, экспериментатор двигал стены и потолок вперед или назад.

Была выдвинута гипотеза, что движение комнаты вперед (рис. 7.5, а) создает «оптический эффект движения», аналогичный возникающему при движении головы ребенка назад (рис. 7.5, б). Это вызовет у ребенка компенсаторный наклон вперед. Если предположение верно, то движение комнаты вперед должно привести к тому, что ребенок наклонится или даже упадет вперед (рис. 7.5, б). Подобное предположение было сделано для движения комнаты в противоположную сторону.

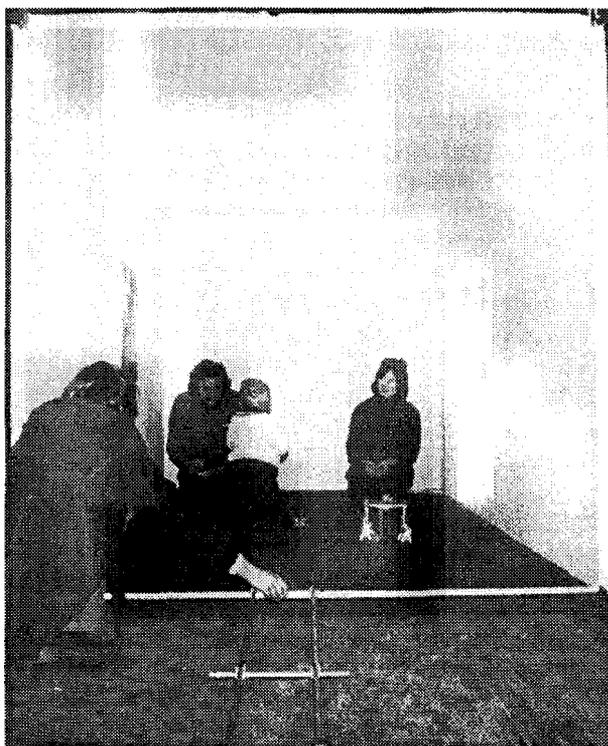


Рис. 7.4. Аппарат, использованный в эксперименте с движущейся комнатой, проведенном Ли и Аронсоном в 1974 г.

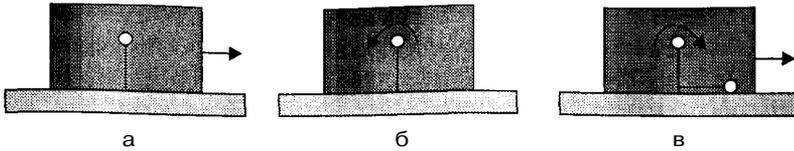


Рис. 7.5. Предсказанное влияние движения комнаты вперед в эксперименте Ли и Андерсона 1974 г.

В противоположность исследованию с кошками, выращенными в «вертикальном» или «горизонтальном» мире, в данном случае не было препятствий для исследования одних и тех же детей при обоих экспериментальных условиях — движении комнаты назад и вперед. Поэтому использовался внутрисубъектный однофакторный план с повторяемыми измерениями. Независимой переменной было направление движения комнаты (вперед или назад), а наклон тела или падение ребенка оценивалось в качестве зависимой переменной. Каждый испытуемый выполнил по двадцать попыток, при этом направление движения комнаты чередовалось от попытки к попытке. У одних детей последовательность начиналась с движения комнаты вперед, а у других — с движения назад. Было набрано семь участников в возрасте от 13 до 16 месяцев, но трое из них испытали сильный стресс и для них эксперимент был прекращен. Реакции оставшихся четырех участников записывались тремя наблюдателями (почему понадобилось более одного наблюдателя?). Потеря равновесия в предсказанном направлении наблюдалась в 82% всех попыток. Наблюдатели выделяли три вида потери равновесия: качание (26% попыток), шатание (23%) и падение (33%).

Один недостаток процедуры позиционного уравнивания, состоящей в простом чередовании условий, (а) и (в), заключается в том, что испытуемые могут легко предсказать, какое условие будет использоваться следующим. Однако Ли и Аронсон справедливо решили, что при подобном возрасте участников эта проблема не повлияет на результаты исследования. Другая причина использования такой разновидности позиционного уравнивания была практической: поскольку ребенок оставался в комнате в течение всего сеанса, если комнату двигали в одну сторону, то при следующей попытке ее необходимо было двигать обратно.

Анализ однофакторных двухуровневых планов

Чтобы выяснить, являются ли различия, обнаруженные между результатами двух условий при двухуровневом плане, значимыми или они вызваны случайностью, требуется провести статистический анализ вывода. Если в обработке результатов эксперимента используется интервальная шкала измерений или шкала отношений, для проведения анализа чаще всего применяется одна из двух разновидностей проверки по критерию Стьюдента. Данная процедура упоминалась ближе к концу главы 4, а более подробно описывается в приложении С. При использовании номинальной или порядковой шкал необходимо использовать другие подходы.

Есть две разновидности проверки по критерию Стьюдента. Первая из них называется **проверкой для независимых групп** и, как ясно из названия, она используется, если две группы участников совершенно независимы друг от друга. Это возможно: а) когда участники исследования случайным образом распределены по

группам или б) если изучаемая переменная является субъективной (например, если участники делятся на женщин и мужчин). Если независимая переменная представляет собой внутрисубъектный фактор или две различные группы людей формируются таким образом, что между ними устанавливаются определенные отношения (например, испытуемые в группе Л уравниваются по интеллекту с участниками группы В), то используется проверка по критерию Стьюдента для зависимых групп (иногда ее называют проверкой по критерию Стьюдента для коррелированных групп). Четырем рассмотренным однофакторным планам соответствуют следующие варианты проверки по критерию Стьюдента:

- проверка по критерию Стьюдента для независимых групп:
 - план с независимыми группами;
 - план с неэквивалентными группами;
- проверка по критерию Стьюдента для зависимых групп:
 - план с уравненными группами;
 - план с повторяемыми изменениями.

По сути, проверка по критерию Стьюдента заключается в нахождении разницы между двумя значениями среднего арифметического и выяснении (с определенной вероятностью), не превышает ли эта разница значения, получаемого при воздействии только случайных факторов. Если она больше этого значения и можно исключить потенциальные осложнители, то исследователь может с высокой вероятностью заключить, что различия действительно существуют, опубликовать исследование и, возможно, получить повышение по службе. В приложении С приведена пошаговая инструкция для обоих видов проверки по критерию Стьюдента и для анализа силы эффекта.

Планы с контрольными группами

Понятие об экспериментальных и контрольных группах, а также о различиях между ними было дано в главе 5. Экспериментальные группы подвергаются определенному воздействию, а контрольные — не подвергаются. В случае плана с повторяемыми измерениями, не подразумевающего наличия разных групп, данное рассуждение переносится на условия исследования, среди которых выделяются экспериментальные и контрольные. Кроме обычной контрольной группы, не подвергающейся воздействию, следует отметить еще три особых вида контрольных групп: контрольные группы плацебо, листа ожидания и сцепленные группы.

Контрольные группы плацебо

«Плацебо» — это вещество, имеющее определенный эффект, но при этом фармакологически неактивное. Иногда пациентам становится лучше, когда они принимают плацебо, но думают, что получают лекарство X. Это происходит потому, что люди верят в то, что лекарство им поможет. В исследованиях члены **контрольной группы плацебо** думают, что подвергаются определенному воздействию, но в действительности этого не происходит. Далее вы поймете, почему может потребоваться использовать такой подход. Предположим, вы хотите выяснить, замедляет ли

алкоголь реакцию человека. Если вы используете обычную экспериментальную группу, которую попросите принять алкоголь, и вторую, сохраняющую трезвость, а затем определите у испытуемых скорость реакции, то может оказаться, что реакция у первой группы будет замедленной. Можно ли будет заключить, что алкоголь снижает скорость реакции? Нет, ведь участники могут придерживаться распространенного мнения, что алкоголь замедляет реакцию, и это знание может оказать на них некоторое влияние. Чтобы избавиться от такого искажения, вам потребуется использовать еще одну группу и дать ей напиток, кажущийся алкогольным (который на вкус невозможно отличить от алкогольного), но таковым не являющийся. Такая группа является контрольной группой плацебо. Следует ли устранять обычную контрольную группу (вообще не принимающую алкоголь)? Не обязательно, ведь эти испытуемые могут дать образец скорости реакции. Если вы используете все три группы и получите следующие значения средней скорости реакции:

экспериментальная группа — 0,32 с;
 контрольная группа плацебо — 0,22 с;
 обычная контрольная группа — 0,16 с,

то сможете сделать вывод, что ожидание людей по поводу действия алкоголя несколько замедляет скорость реакции (с 0,16 до 0,22 с), но и сам алкоголь также оказывает влияние (снижение скорости реакции с 0,22 до 0,32 с).

Примером эксперимента с двумя контрольными группами — обычной и плацебо — является исследование влияния заражения паразитами на различные познавательные и моторные навыки у детей с Ямайки (Sternberg, Powell, McGrane & Grantham-McGregor, 1997). Ученики четвертых и пятых классов с легкой степенью заражения были случайным образом распределены (после уравнивания по полу) на две группы: получающую лекарство и плацебо. Дети из третьей (неэквивалентной) группы не были заражены. Лекарство быстро устранило инфекцию, и по окончании эксперимента его дали также и детям из группы плацебо. Но последствия заражения остались. В сравнении с незараженной контрольной группой, не получающей лечения, дети из двух других групп плохо справились с заданиями на проверку когнитивных способностей. Следовательно, лекарство имело важный медицинский эффект, но мало способствовало усилению когнитивных способностей, снизившихся из-за болезни. Поскольку дети жили в условиях, где паразитические инфекции весьма распространены, Стернберг и др. порекомендовали включить в лечение программу коррекции когнитивных способностей.

Контрольные группы листа ожидания

Контрольные группы листа ожидания нередко используются в исследованиях по оценке эффективности программ (глава 10) или выявлению действенности психотерапии. Такой план используется, когда члены экспериментальной группы участвуют в программе из-за того, что испытывают определенные проблемы, решению которых эта программа должна помочь. Например, в исследовании Миллера и Дипилато (Miller & DiPilato, 1983) оценивалась эффективность двух видов терапии (релаксации и десенсилизации) для людей, страдающих ночными кошмарами. Исследователи решили использовать контрольную группу, не подвергающуюся воз-

действию, но для того, чтобы испытуемые во всех трех группах были в целом эквивалентны друг другу, члены контрольной группы также должны были страдать от ночных кошмаров. Участникам, попавшим в группу листа ожидания, сказали, что им обязательно помогут, и по окончании исследования им обеспечили лечение, эквивалентное тому, которому были подвергнуты члены экспериментальных групп.

То, что членам группы листа ожидания обеспечивалось лечение, было весьма полезно для здоровья испытуемых, но также это оказало определенное давление на исследователей, которые были вынуждены использовать данную контрольную процедуру только для программ, имеющих относительно небольшую продолжительность. Могут возразить, что использование контрольной группы листа ожидания неэтично, так как люди сначала не получают никакой пользы от программы. Этот вопрос может стать особенно сложным, если в ходе исследования оценивается программа, влияющая на жизнь человека. Более подробно этот вопрос рассматривается во вставке 7.2, защищающей использование контрольных групп в исследованиях.

ВСТАВКА 7.2

Этика — кто войдет в контрольную группу?

При проведении исследования памяти, в котором экспериментальную группу просят формировать визуальные образы, тогда как контрольная группа просто заучивает список слов, вопрос о том, кто должен быть в контрольной группе, не создает этической дилеммы. Однако если эксперимент разрабатывается для оценки некоторой программы или терапевтического воздействия, которое может принести заметную пользу людям или даже продлить их жизнь, то этот вопрос решить нелегко. Например, в известном исследовании влияния личного контроля на здоровье (Langer & Rodin, 1976) нескольким людям, живущим в доме для престарелых, позволили самим планировать свой день, тогда как члены контрольной группы пользовались (в основном) распорядком, установленным для них персоналом. Когда авторы через 18 месяцев вернулись и провели повторное исследование, они обнаружили, что в среднем в сравнении с контрольной группой члены первой группы были более психически и физически здоровы и большее количество из них были живы (Rodin & Langer, 1977). Если бы вы узнали, что один из ваших родственников попал в контрольную группу, вы бы обеспокоились?

Точно так же возникли споры с распределением участников по группам в исследовании с больными раком пациентами (Adler, 1992). В ходе исследования изучалось влияние групп поддержки на психологическое и физическое состояние женщин, страдающих раком груди. Было обнаружено, что женщины в группах поддержки быстрее выздоравливали и даже жили дольше, чем другие (т. е. распределенные в контрольную группу). Некоторые исследователи утверждали, что результаты отражают не пользу, приносимую группами поддержки, а вред, нанесенный членам контрольной группы, которые могли почувствовать себя обделенными или отвергнутыми. Такие чувства могли вызвать стресс, а как известно, стресс неблагоприятно влияет на иммунную систему, что приводит к разнообразным проблемам со здоровьем. Так есть ли хоть доля правды в рис. 7.6? Может ли участие в контрольной группе убить кого-либо? Защитники использования контрольных групп при оценке программ имеют три сильных аргумента. Во-первых, они отмечают, что легко оценивать прошедшие события и говорить уже после свершившегося факта, что «программа, настолько эффективная, как эта, должна быть доступна для всех». Проблема в том, что до «свершившегося факта» далеко не так очевидно то, что программа будет эффективной. Единственный способ узнать это - провести исследование. Например, до начала исследования Лэнгера и Родена в доме для престарелых мож-

но было легко предположить, что участники эксперимента испытают ненужный стресс от того, что должны сами о себе заботиться, и быстрее умрут. Точно так же защитники исследования рака указывают, что в начале исследования довольно мало женщин имели предпочтения по поводу распределения по группам, а несколько из них просто не захотели присоединиться к группе поддержки (Adler, 1992). Следовательно, неверно, что члены контрольной группы чувствовали себя отвергнутыми или обделенными.



Рис. 7.6. Возможные последствия попадания в контрольную группу?
Надпись на рисунке: Хайрем С. Дадсон, 1930-1993, член группы плацебо

Во-вторых, экспериментаторы указывают, что в исследованиях по оценке новых программ или способов терапии редко сравниваются новый способ воздействия и отсутствие воздействия - обычно проводится сравнение нового метода воздействия со старым. Таким образом, членам контрольной группы доступны методы лечения, которые нормально применяются, и, соответственно они получают обычную помощь. Более того, если исследование показывает наличие положительного эффекта от экспериментального воздействия, члены контрольной группы обычно получают возможность испытать его на себе.

В-третьих, лечение обходится дорого и стоит потратить деньги на лучшее из возможного. Однако нельзя определить без подробно разработанного исследования эффективности программы, какой метод лучше. Программы, эмпирически подтвердившие свою эффективность, служат на благо всем людям и нередко могут сохранить или продлить им жизнь.

Помните, как в главе 1 обсуждался вопрос псевдонауки и приводился в связи с этим пример «бессознательно воспринимаемых аудиозаписей»? Как мы узнали, было проведено обширное исследование, показавшее, что положительное влияние этих записей возникает в результате ожидания их эффективности. Ожидания нередко выявляются путем использования контрольных групп плацебо, но мы рассмотрим исследование, в котором использовались одновременно и группы плацебо, и группы листа ожидания. Это исследование позволило по-новому интерпретировать возможности подобных записей.

Пример 10. Использование контрольных групп плацебо и листа ожидания в одном исследовании

Один из самых распространенных вариантов использования «бессознательно воспринимаемых аудиозаписей» — это использование их с целью снижения веса. Эта область предоставляет очень хорошие возможности для продажи подобной продукции. Американцы, к примеру, чтобы сбросить вес, испытывают на себе всевозможные приемы, от диеты с ограниченным потреблением жиров до хирургического вмешательства. Особенно привлекательными для людей оказываются методы, требующие минимальных усилий, а именно это и является отличительной чертой «бессознательно воспринимаемых аудиозаписей» — просто вставьте кассету и очень скоро ваше бессознательное начнет руководить вашим поведением и неизбежно приведет к потере веса. Для изучения эффективности таких записей в деле снижения веса Мерикл и Скейнс (Merikle & Skanes, 1992) провели исследование, в котором использовали и контрольные группы плацебо, и группы листа ожидания. Дав объявление в газеты, они набрали сорок семь женщин и случайным образом распределили их в три группы. Члены экспериментальной группы ($N=15$) получили коммерческую «аудиозапись бессознательного действия», которая, по мнению производителей, должна была помочь слушателям сбросить вес. Члены контрольной группы плацебо ($N=15$) думали, что получили аудиозаписи, специально разработанные для снижения веса, но в действительности им выдали записи для снижения страха лечения зубов (исследователи обладали чувством юмора). Для обычных слушателей эти две записи были неразличимы. Членам контрольной группы листа ожидания ($N=17$) сказали, «что в исследовании уже принимает участие максимальное количество испытуемых и что... их поставят на лист ожидания» (р. 774). Участников из экспериментальной группы и группы плацебо попросили прослушивать записи 1-3 часа в день; при этом членов всех трех групп на протяжении пяти недель еженедельно взвешивали. Каковы были результаты? Потеря веса у членов экспериментальной группы была весьма скромной, но *столько же* сбросили члены контрольной группы плацебо. Это обычный результат для подобного типа исследований, показывающий, что «бессознательно воспринимаемые аудиозаписи» сами по себе не имеют никакого эффекта. Однако интересно, что контрольная группа листа ожидания также потеряла в весе, и примерно столько же, сколько остальные две группы. На основании этого Мерикл и Скейнс заключили, что «бессознательно воспринимаемые записи» дают предполагаемый результат не просто из-за наличия эффекта плацебо. Если бы все дело было в эффекте плацебо, то члены группы плацебо, думая, что на их разум оказано воздействие, сбросили бы вес, а вес испытуемых из группы листа ожидания, которые еще не слушали записи, не изменился бы. Полученные результаты привели авторов к выводу, что воздействие записей заключалось в концентрации внимания людей на существующей проблеме, в данном случае на снижении веса. Испытуемые во всех трех группах «могли потерять вес просто потому, что участие в исследовании увеличило вероятность того, что в ходе эксперимента они станут внимательно относиться к вопросам, связанным со снижением веса» (р. 776). В данном исследовании группа листа ожидания позволила оценить силу эффекта плацебо и получить альтерна-

тивное объяснение популярности «бессознательно воспринимаемых записей». Также, хотя авторы и не отметили этого, результаты исследования выглядят подозрительно похожими на хоторнский эффект, о котором вы узнали в главе 6.

Следует отметить еще один момент, касающийся этого исследования: второй автор, Хизер Скейнс, проводил эксперименты, а первый, Филипп Мерикл, наклеивал этикетки на кассеты. Таким образом, он был единственным, кто знал, какие участники получают записи для потери веса (экспериментальная группа), а какие — записи для снижения страха лечения зубов (группа плацебо). Тем самым исследователи применили двойной слепой метод.

Эквивалентные контрольные группы

Третий вид контрольных групп — **эквивалентные контрольные группы**. Их используют, если члены экспериментальной группы по какой-либо причине участвуют в эксперименте разное количество времени или включены в разные события, происходящие в ходе исследования. Каждый участник из контрольной группы уравнен, или «сцеплен», с одним из членов экспериментальной группы, таким образом для групп в целом затраченное участниками время, а также разновидности событий, с которыми они встречаются, поддерживаются постоянными. Приведенный ниже пример поможет вам лучше понять идею эквивалентных контрольных групп.

Пример 11. Эквивалентные контрольные группы и стресс

Хороший пример эквивалентной контрольной группы можно найти в исследовании Вейсса (Weiss, 1968), посвященном изучению взаимосвязей между управлением стрессорами и здоровьем. К хвостам крыс в ходе эксперимента в случайном порядке посылали слабые электрические разряды. Крысы могли выключить ток (т. е. управлять им) или избежать его воздействия, вращая лапами небольшое колесо (рис. 7.7). Контрольная группа крыс не подвергалась шоку. Крысы, находящиеся в центральной части аппарата, принадлежали к сцепленной контрольной группе. Каждая из них была сцеплена (имела парой) с крысой из экспериментальной группы и получала такое же количество электрических разрядов, но не могла управлять ими. Каждая попытка начиналась с сигнала, который слышали все три животных. У крысы из экспериментальной группы было 10 секунд, чтобы повернуть колесо. Если она не делала этого, то к ее хвосту посылался электрический разряд. При этом сцепленная крыса также получала разряд (вне зависимости от того, что она делала в это время). Таким образом, крысы в экспериментальной и контрольной группах получали абсолютно одинаковое количество раздражителей, вызывающих отрицательную реакцию (т. е. одинаковое количество электрических разрядов), но отличались по возможности контроля за ситуацией. Вейсс заключил, что возможность управлять поступлением тока помогла крысам избежать некоторых разрушительных для здоровья последствий стресса. У крыс из сцепленной группы вероятность развития язвы и потери веса была выше, чем у крыс из экспериментальной и контрольной групп, на которых подобный опыт не оказал почти никакого влияния.

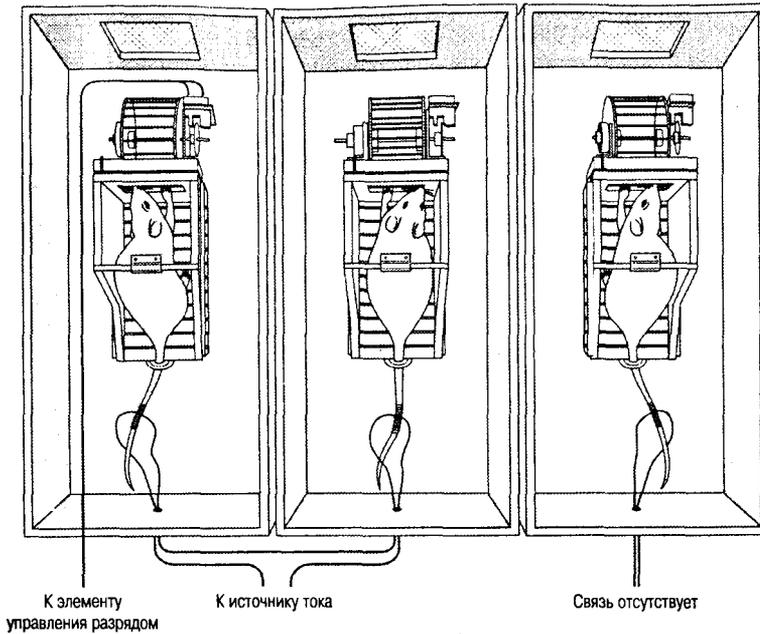


Рис. 7.7. Данные первых двух экспериментов из исследования Вейсса

Вы уже знакомы с экспериментальным планом, использованным Вейссом, — его использовала исследовательская группа Селигмана при изучении выученной беспомощности. Если вы еще раз обратитесь к обсуждению теорий, приведенному в главе 3, то обнаружите описание процедуры с эквивалентными контрольными группами, использованной в триадном плане Селигмана. Другой пример эквивалентной контрольной группы дает исследование Брэди, посвященное изучению язвы у «исполнительных» обезьян, описанное во вставке 5.3. По сути, при разработке своего исследования Вейсс скорректировал методологические недостатки (т. е. проблему отбора испытуемых), обнаруженные в исследовании Брэди (Weiss, 1977).

Один фактор — более двух уровней

Если в экспериментах используется одна независимая переменная, ситуация, когда изучаются только два ее значения, является скорее исключением, чем правилом. В большинстве однофакторных исследований используется три или более значений независимой переменной, поэтому такие планы часто называют **однофакторными многоуровневыми планами**. Так же как и двухуровневые, многоуровневые планы могут быть меж- и внутрисубъектными, а также планами с независимыми, уравненными или неэквивалентными группами, или планами с повторяемыми измерениями.

Межсубъектные многоуровневые планы

Явное преимущество многоуровневых планов заключается в том, что они позволяют исследовать нелинейные эффекты. Рассмотрим простой пример межсубъектного плана. Предположим, вы интересуетесь влиянием различных доз кофеина на скорость реакции. Вы разрабатываете эксперимент, в котором сравниваете два уровня дозировки (1 и 3 мг), получаете результаты, представленные на рис. 7.8, и делаете вывод, что кофеин как стимулятор ускоряет реакцию. При увеличении дозировки скорость реакции нарастает по прямой (т. е. линейно).

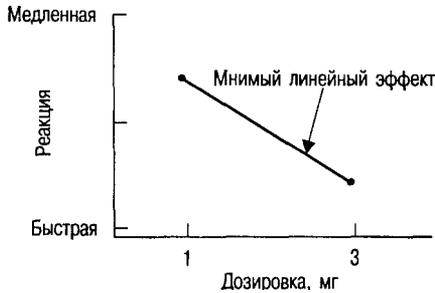


Рис. 7.8. Гипотетическое влияние кофеина на скорость реакции - два уровня

Далее предположим, что другой исследователь провел такое же исследование, но использовал многоуровневый план, включающий четыре дозы кофеина (1, 2, 3 и 4 мг) — пример повторения (1 и 3 мг) и дополнения (2 и 4 мг) вашего эксперимента. Такое исследование может дать результаты, представленные на рис. 7.9.

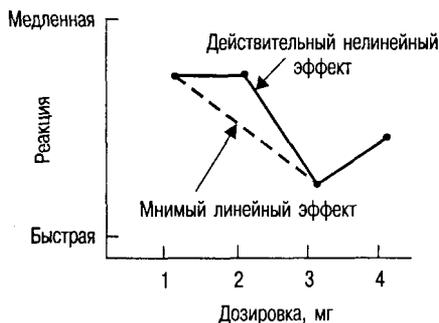


Рис. 7.9. Гипотетическое влияние кофеина на скорость реакции - четыре уровня

Таким образом, в точности воспроизведены ваши результаты для условий с дозировкой в 1 и 3 мг, но в целом результаты для четырех условий ставят ваши выводы под сомнение. Вместо вывода о том, что кофеин увеличивает скорость реакции, в данном случае можно заключить, что а) увеличение дозы кофеина повышает скорость реакции, но лишь по достижении уровня в 2 мг и б) кофеин увеличивает скорость реакции только до определенного момента — после 3 мг кофеин снижает ее. Это значит, что результаты больше не показывают линейную зависимость и являются нелинейными. В целом, преимущество многоуровневых планов состоит

в том, что они более информативны и зачастую дают более интересные результаты, чем двухуровневые.

Один из самых известных в психологии графиков иллюстрирует нелинейный эффект. Он показывает, как время, прошедшее с момента запоминания, влияет на объем забытого материала, и объясняет, почему через два дня после экзамена очень часто говорят «Я ничего не помню!» Эта кривая впервые была получена в новаторском исследовании памяти, проведенном Германом Эббингхаузом (см. вставку 7.3).

ВСТАВКА 7.3

История — нелинейные результаты - кривая забывания Эббингхауза

Немецкий психолог XIX в. Герман Эббингхауз (1850-1909) известен своим новаторским исследованием памяти и забывания. В те времена, когда психология переживала свое детство и руководств о том, «как проводить психологические исследования», еще не существовало, Эббингхауз провел серию исследований, ставших стандартом точности и методологической строгости. Цель исследования заключалась в изучении образования ассоциаций в человеческом сознании, и первой задачей было найти материал, свободный от ассоциаций. Способ решения этой задачи является одним из лучших примеров креативности в психологии: учебный образовал последовательности, состоящие из согласных, гласных и снова согласных. Такие последовательности, обозначаемые как CVC (consonants, vowels, consonants), более известны как бессмысленные слоги. Эббингхауз создал около 2300 таких слогов. В течение нескольких лет, демонстрируя потрясающую настойчивость, а возможно, и полное отсутствие личной жизни, Эббингхауз запоминал, а затем старался вспомнить последовательности CVC. Да, действительно, он был единственным испытуемым. Он систематически изменял такие факторы, как количество слогов в списке, количество экспериментальных попыток на один список и время, разделяющее попытки друг от друга. Эббингхауз опубликовал результаты своего исследования в небольшой монографии, названной «Память: исследование по экспериментальной психологии» (Memory: A Contribution to Experimental Psychology, 1885/1964).

Одно из самых известных своих исследований, в ходе которого были получены нелинейные результаты, Эббингхауз посвятил изучению процесса забывания с течением времени. Был поставлен следующий эмпирический вопрос: какой объем из запомненного материала сохраняется спустя различные отрезки времени? Процедура состояла в том, чтобы запомнить восемь списков по 13 бессмысленных слогов в каждом, подождать некоторое время, а затем попытаться выучить их еще раз. Временные интервалы составили 20 минут, 1 час, 9 часов, 1 день, 2 дня, 6 дней и 31 день. Эббингхауз зафиксировал общее время первоначального и повторного заучивания восьми списков. Их разность составила величину «экономии», которую ученый перевел в проценты, разделив на время первичного запоминания. Таким образом, если начальное запоминание заняло 20 минут, а повторное - 5 минут, то было сэкономлено 15 минут или 75% ($15 / 20 \times 100$), от времени первичного запоминания.

Результаты, полученные Эббингхаузом, показаны на рис. 7.10. Вероятно, вы также сможете найти их в главе, посвященной памяти, учебника по общей психологии. Несложно заметить, что процент запомненных слогов снижался со временем, но это снижение не носило равномерного или линейного характера. Очевидно, что эффект был нелинейным. Сначала забывание было очень сильным, но затем оно замедлялось. Таким образом, через 20 минут было сэкономлено только около 60% (58,2) от времени первичного запоминания. Другая часть кривой показывает, что между интервалами в неделю (экономлено 25,4%) и месяц (экономлено 21,1%) не было больших отличий.

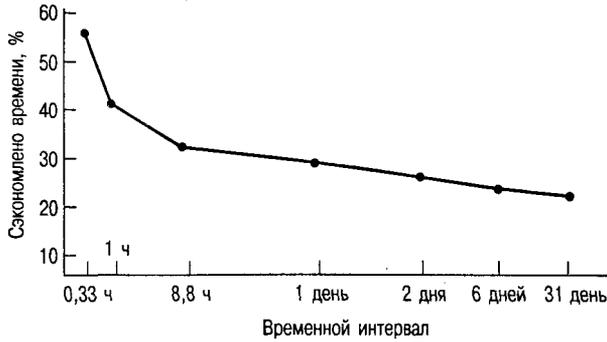


Рис. 7.10. Кривая забывания Эббингхауза - нелинейные результаты

Есть несколько интересных методологических особенностей исследования Эббингхауза. Например, чтобы сделать постоянным время показа слогов, ученый установил метроном на 150 ударов в минуту и читал каждый слог CVC в течение ровно одного удара. Также он старался изучать списки слогов в одних и тех же условиях, в одно время суток и не использовать никаких особых техник для запоминания, кроме простого повторения. Кроме того, он работал, только когда чувствовал достаточную мотивацию и мог «концентрировать внимание на утомительном задании и его цели» (Ebbinghaus, 1885/1964, p. 25).

Кроме выявления нелинейного эффекта однофакторные многоуровневые планы также используются для проверки альтернативных гипотез. Именно этот подход приводился в пример в главе 3 при обсуждении положительных сторон фальсификации. Превосходный пример этого являет собой исследование Брэнсфорда и Джонсона (Bransford & Jonson, 1972).

Пример 12. Многоуровневый план с независимыми группами

Ученые, посвятившие себя когнитивной психологии и занимающиеся вопросами понимания новой для нас информации, показали, что понять новую идею легче, если она окружена определенным контекстом. Например, главу из книги будет легче понять, если вы сначала прочтаете обзор и задачи этой главы. Я надеюсь, вы уже обнаружили это при работе с данным учебником. Исследование Брэнсфорда и Джонсона показывает влияние контекста на понимание. В ходе исследования участников просили прочитать и понять следующий текст. Попробуйте и вы:

Если шарики лопнут, звук не сможет быть услышан, поскольку это произойдет слишком далеко от нужного этажа. Закрытое окно также не даст возможности распространиться звуку, так как большинство зданий обладают весьма неплохой звукоизоляцией. Поскольку вся процедура зависит от равномерного потока электричества, разрыв провода также может вызвать проблемы. Конечно, человек может кричать, но человеческий голос недостаточно силен, чтобы разнести так далеко. Дополнительная проблема состоит в том, что у инструмента может порваться струна. Тогда сообщение останется без аккомпанемента. Очевидно, что чем меньше будет расстояние, тем лучше. Тем самым будет меньше потенциальных проблем. При личном контакте возможностей потерпеть неудачу будет гораздо меньше (Bransford & Jonson, 1972, p. 392).

Я думаю, вашей реакцией на этот пассаж будет что-то вроде «Хм?» Именно так реагировало большинство участников данного исследования. Однако Брэнсфорд и Джонсон обнаружили, что пониманию можно помочь, использовав определенный контекст.

Психологи разработали однофакторное исследование с независимыми группами, в котором независимая переменная принимала пять значений. Испытуемые, случайным образом распределенные в контрольную группу, выполняли задание, аналогичное только что выполненному вами: они читали текст и старались запомнить как можно больше идей из 14, содержащихся в тексте. В результате в среднем они запомнили 3,6 идей, что совсем не впечатляет. Чтобы проверить, сможет ли простое повторение улучшить запоминание, вторую группу попросили прочитать рассказ дважды. Это не помогло — испытуемые запомнили 3,8 идей. Третьей группе предварительно показали рисунок (рис. 7.11, а), а затем попросили прочитать и пересказать текст. Участники из этой группы запомнили 8,0 идей из 14. Очевидно, что рисунок явился общим контекстом, помогшим участникам понять рассказ. Но обязательно ли *сначала* рассмотреть рисунок, а затем прочитать текст? Да, обязательно. При четвертом условии эксперимента участники читали рассказ, *затем* рассматривали рисунок, а потом пересказывали прочитанный текст. Они запомнили 3,6 идей — ровно столько, сколько запомнили члены контрольной группы. Пятой группе предложили частичный контекст. Прежде чем читать рассказ, они рассмотрели рис. 7.11, б, содержащий все отдельные части рисунка 7.11, а, но иначе расположенные. Испытуемые из этой группы запомнили в среднем 4,0 идей. Полученные результаты можно изобразить графически, как показано на рис. 7.12.

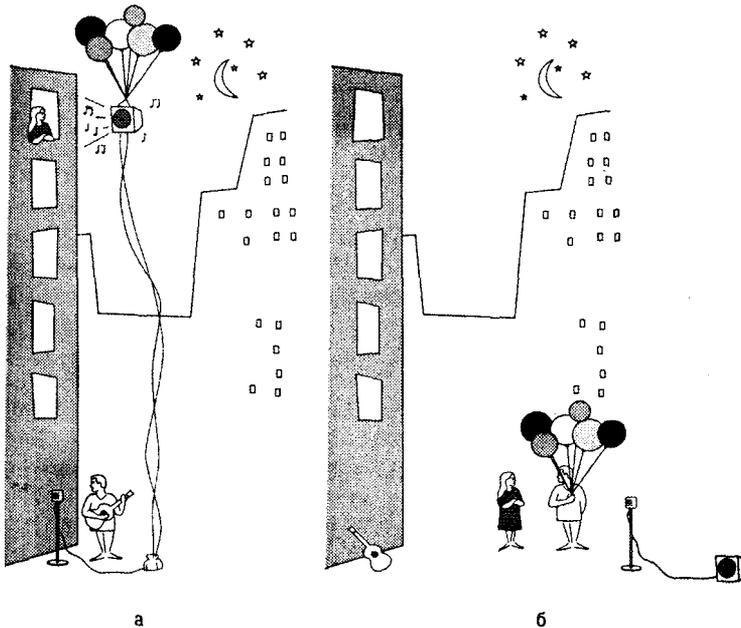


Рис. 7.11. Рисунок, обеспечивающий: а) контекст и б) частичный контекст для исследования Брэнсфорда и Джонсона (Bransford & Jonson, 1972)

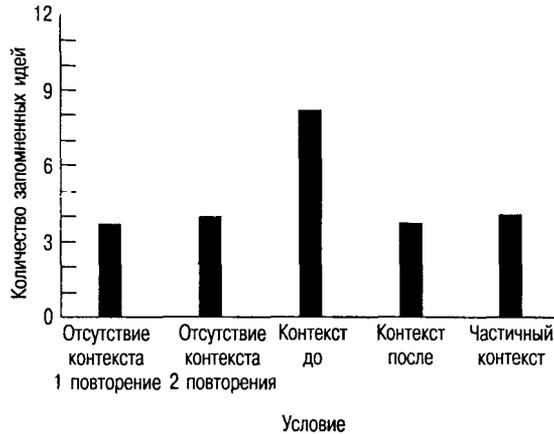


Рис. 7.12. Данные по пяти условиям исследования Брэнсфорда и Джонсона, представленные в виде гистограммы

Аналогичное исследование с двумя группами: «Отсутствие контекста — одно повторение» и «Контекст до», было бы весьма интересным и показало бы улучшение понимания при введении контекста в форме рисунка. Но использование всех четырех условий делает исследование *действительно* интересным и позволяет отбросить (т. е. фальсифицировать) некоторые альтернативные факторы, которые иначе будут расценены как улучшающие запоминание. Таким образом, контекст улучшает понимание, но *только* если он введен предварительно. Так как представление контекста после прочтения текста не улучшает запоминания, можно сделать вывод, что контекст делает это, облегчая первичную обработку информации, а не ее последующее восстановление в памяти. Можно также отбросить предположение, что простое повторение улучшает запоминание, ведь удваивание количества повторений текста не способствовало запоминанию. Кроме того, недостаточно просто показать отдельные части рассказа (как в случае с частичным контекстом) — они должны быть расположены в определенном порядке и отражать суть запоминаемого материала.

Внутрисубъектные многоуровневые планы

Тогда как в случае однофакторного двухуровневого плана с повторяемыми измерениями можно применить лишь отдельные варианты позиционного уравнивания, выход за пределы двух уровней дает возможность использовать все его виды. Если каждое условие изучается один раз для каждого испытуемого, доступны и полное, и частичное позиционное уравнивание. Если каждое условие исследуется несколько раз для каждого испытуемого, можно использовать обратное позиционное уравнивание или блоковую рандомизацию. В нижеследующем примере каждое условие изучалось только один раз и для позиционного уравнивания использовался латинский квадрат.

Пример 13. Многоуровневые планы с повторяемыми измерениями

Может ли прослушивание музыки Моцарта сделать вас умнее? Некоторые считают, что может, — подобный феномен был назван «эффект Моцарта». Несмотря на отсутствие фактов, поддерживающих эту идею, родителей пытаются убедить, что

музыка Моцарта способна дать их детям оружие для IQ-битвы. Существует даже веб-сайт, посвященный эффекту Моцарта (www.mozarteffect.com), где можно заказать различные записи и книги. В истинно псевдонаучной манере сайт доказывает действенность этого явления, приводя отзывы покупателей и единичные свидетельства. В описании одной из записей, продающейся на сайте (концерта для скрипки), утверждается, что прослушивание концерта «увеличит вербальные и эмоциональные способности, улучшит концентрацию внимания, память и навыки интуитивного и пространственного мышления», а также что «высокая частота звуков скрипки тренирует слух, а следовательно, стимуляция мозга уравнивается прекрасной гармонией» (что бы это ни значило). Каковы же основания для таких необычайных утверждений? Поиски приводят нас к небольшому исследованию, опубликованному в 1993 г. в журнале *Nature*. В нем было показано, что прослушивание произведений Моцарта по 10 минут в день производит кратковременное (т. е. эффект длится недолго) усиление способности к пространственному мышлению у студентов колледжа (Rauscher, Shaw, & Key, 1993). Занимающиеся когнитивной психологией ученые скептически отнеслись к информации даже о таком кратковременном эффекте и неоднократно пытались повторить это исследование. Все попытки были неудачными. Одно из таких исследований провели Стил, Болл и Ранк (Steele, Ball & Runk, 1977).

В своем исследовании Стил и его коллеги использовали три условия: прослушивание произведений Моцарта в течение 10 минут, прослушивание успокаивающих звуков естественных природных явлений (например, мягкий шум дождя) в течение 10 минут и отсутствие прослушивания каких-либо звуков — в течение 10 минут участники в сидели в тишине и старались расслабиться. Все 36 участников исследовались при каждом условии, а следовательно, план был межсубъектным многоуровневым. Хотя легко можно было провести полное позиционное уравнивание (шесть различных последовательностей условий, шесть участников, случайным образом распределенные в каждой из шести последовательностей), авторы решили использовать латинский квадрат размером 3x3, подразумевающий участие 12 испытуемых, случайно распределенных в каждом из рядов квадрата. Чтобы предотвратить искажение, которое могло возникнуть, если бы участники знали, что изучается эффект Моцарта, им «сказали, что эксперимент посвящен влиянию релаксации на запоминание» (Steel et al., 1997, p. 1181). Задание на проверку памяти заключалось в воспроизведении в обратном порядке набора чисел. Если стимул был «6-8-3-1-7», то правильный ответ должен был быть «7-1-3-8-6». В ходе каждой попытки испытуемые слушали Моцарта, мягкие звуки дождя или сидели в тишине, а затем выполняли три задания на запоминание чисел. Каждое задание состояло из 9 чисел, представленных в случайном порядке. Таким образом, участники могли набрать от 0 до 27 очков.

Результаты исследования являются статистически значимыми, но они отнюдь не способствуют продажам записей Моцарта. Среднее количество правильно запомненных чисел было практически одинаковым для всех трех условий: 18,53 для записей Моцарта, 18,50 для записи шума дождя и 18,72 для контрольного условия. Однако наблюдался значительный эффект тренировки. Вне зависимости от порядка следования условий участники продемонстрировали улучшение результатов от первого набора чисел для запоминания к третьему (средние оценки для них соста-

вили соответственно 15,64, 19,14 и 20,97). Так нужно ли родителям ставить записи произведений Моцарта своим детям? Конечно, нужно, ведь это прекрасная музыка. Сделает ли она их умнее? Нет, но зато они наслаждаются классической музыкой, что само по себе хорошо.

Представление данных

При составлении отчетов о результатах исследований необходимо решить, каким образом представить данные. Есть три возможности. Во-первых, их можно представить в повествовательной форме — такой подход хорош, если экспериментальное исследование включает два или три значения независимой переменной (как, например, в исследовании эффекта Моцарта), но будет весьма скучным при увеличении количества данных. Возможно, вы уже заметили это, когда читали о результатах исследования Брэнсфорда и Джонсона, содержащего пять условий (Bransford & Jonson, 1972). Второй способ — поместить результаты в таблицу. Для результатов исследования Брэнсфорда и Джонсона можно построить таблицу, подобную табл. 7.2.

Таблица 7.2

Данные исследования Брэнсфорда и Джонсона, представленные в виде таблицы

Таблица 1. Среднее количество запомненных идей как функция различных контекстов для запоминания и воспроизведения материала

Условие	Средняя оценка	Стандартное отклонение
Отсутствие контекста — 1 повторение	3,60	0,64
Отсутствие контекста — 2 повторения	3,80	0,79
Контекст до	8,00	0,65
Контекст после	3,60	0,75
Частичный контекст	4,00	0,60

Примечание. Максимальная возможная оценка равна 14.

Третий способ представления данных — это график. С его помощью можно представить исследование Брэнсфорда и Джонсона, как показано на рис. 7.12. Обратите внимание, что на графике, созданном для экспериментального исследования, зависимая переменная всегда откладывается по вертикальной оси (Y), а независимая — по горизонтальной (X). Как вы узнаете в следующей главе, ситуация несколько усложняется, если используется более одной независимой переменной. Однако вне зависимости от количества независимых переменных зависимая переменная всегда откладывается по вертикальной оси.

Что использовать, график или таблицу, исследователь решает по своему усмотрению. График может выглядеть очень эффектно, если получены большие различия между результатами или обнаружено взаимодействие (глава 8). Таблицы обычно используют, если данных так много, что график невозможно будет понять, или если исследователь хочет показать читателям точные значения среднего арифметического (в случае графика о точных значениях можно будет лишь догадываться). Единственное правило, которым необходимо руководствоваться, — это то, что одни и те же данные нельзя представлять одновременно и в виде таблицы, и в виде

графика. В целом, данные должны быть представлены таким способом, чтобы результаты, на получение которых вы потратили столько сил, были отображены наиболее ясно и понятно.

Виды графиков

Обратите внимание, что я представил данные исследования Брэнсфорда и Джонсона в виде гистограммы. Почему нельзя сделать это с помощью линейного графика, как на рис. 7.13? В данном случае это не слишком хорошая идея. Проблема заключается в сущности конструкта, использованного в качестве независимой переменной, и непрерывности этой переменной. **Непрерывная переменная** — это переменная, у которой существуют промежуточные значения, а это значит, что она существует в определенном континууме. В качестве примера можно привести дозировку лекарства. В исследовании, в котором сравниваются дозы лекарства в 3,5 и 7 мг, дозировка является непрерывной переменной, ведь мы можем, если потребуется, использовать также 4 или 6 мг. Для изображения результатов в случае непрерывной независимой переменной можно использовать линейный график. Чтобы предположить эффективность промежуточных значений, можно провести интерполяцию по имеющимся точкам и по полученной линии оценить их влияние. В исследовании с лекарствами график может иметь вид, показанный на рис. 7.14. При этом исследователь может быть вполне уверен в оценке эффективности промежуточных значений дозировки, одно из которых помечено звездочкой на рис. 7.15.

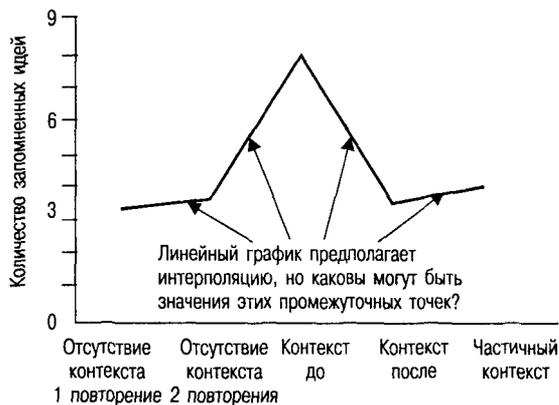


Рис. 7.13. Данные исследования Брэнсфорда и Джонсона, неверно представленные в виде линейного графика

Конечно, если в исследовании используется два значения независимой переменной, довольно сильно отстоящих друг от друга, а зависимость в действительности нелинейная, интерполяция может вызвать проблемы. Так, если в ходе исследования сравниваются дозы лекарства в 2 и 10 мг и получена прямая, изображенная непрерывной линией на рис. 7.16, то если учесть, что истинную зависимость отображает кривая, показанная пунктирной линией, интерполяция эффекта дозы в 5 мг приведет к огромной ошибке. Такое исследование лучше проводить с использованием однофакторного многоуровневого плана.

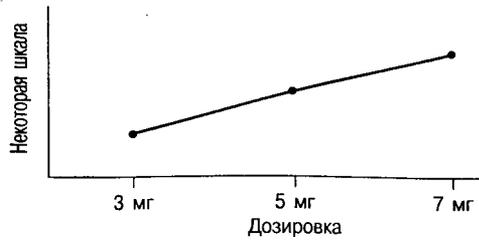


Рис. 7.14. Правильное использование линейного графика в случае непрерывной переменной (дозировка лекарства)

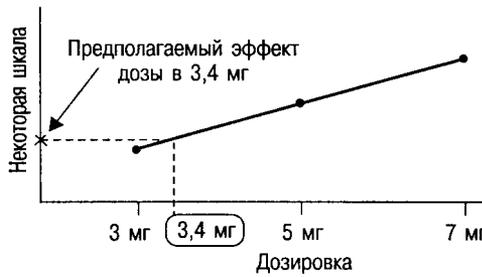


Рис. 7.15. Интерполяция точек линейного графика

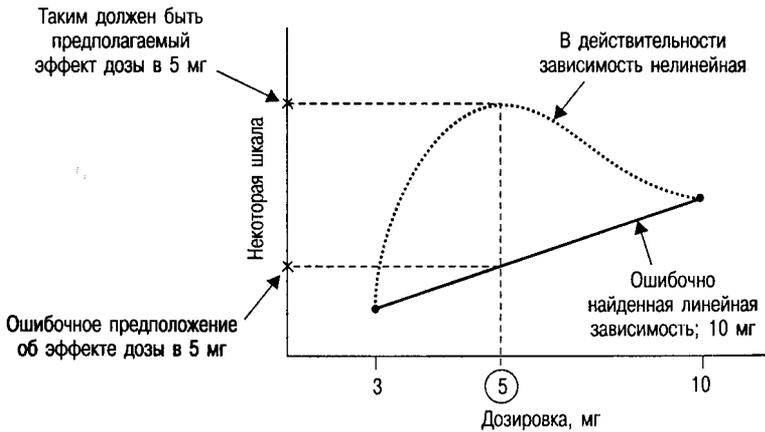


Рис. 7.16. Проблемы интерполяции в случае нелинейной зависимости и большого промежутка

В случае дискретной независимой переменной, каждое значение которой представляет отдельную область, а промежуточных точек просто не существует, ситуация в корне меняется. В таком случае невозможно провести интерполяцию, а следовательно, связать точки с помощью прямой означает предположить существование промежуточных точек, которых в действительности не существует. Поэтому при использовании дискретных переменных, как в исследовании Брэнсфорда и Джонсона (рис. 7.12), обычно строятся гистограммы. Основное правило такое:

Если переменная непрерывна, лучше использовать линейный график, также можно использовать гистограмму.

Если переменная дискретна, лучше использовать гистограмму, линейный график использовать нельзя.

В целом, гистограмму можно использовать как для непрерывных, так и для дискретных данных, а линейный график можно использовать только для непрерывных данных. Обратитесь ко вставке 4.3 — она напомнит вам об этических аспектах представления данных. Легко дезинформировать наивных читателей отчета об исследовании, изменив расстояния на шкале Y . Вы как исследователи ответственны за то, чтобы честно представить результаты эксперимента и использовать для этого способ, позволяющий наилучшим образом проиллюстрировать полученные данные.

Анализ однофакторного многоуровневого плана

Как вы уже знаете, если в случае однофакторного двухуровневого плана независимая переменная измеряется с помощью интервальной шкалы или шкалы отношений, для проверки нулевой гипотезы можно использовать коэффициент Стьюдента. Вы можете подумать, что для многоуровневого плана, как, например, в исследовании Брэнсфорда и Джонсона, достаточно будет найти коэффициенты Стьюдента для всех пар условий (например, для контекста до и контекста после). К сожалению, все не так просто. Трудность заключается в том, что проведение многократных проверок по критерию Стьюдента увеличивает риск возникновения ошибки 1-го рода. Чем больше подобных проверок вы проводите, тем больше вероятность ошибочно обнаружить значимые различия. Чтобы охватить все пары условий в исследовании Брэнсфорда и Джонсона, потребуется найти 10 коэффициентов Стьюдента.

Вероятность сделать по крайней мере одну ошибку 1-го рода при проведении многократных проверок по критерию Стьюдента можно оценить по формуле

$$1 - (1 - \alpha)^c,$$

где c — это количество проведенных сравнений.

Таким образом, если для исследования Брэнсфорда и Джонсона найти все возможные коэффициенты Стьюдента, то возникнет очень высокая вероятность (4 из 10) сделать по крайней мере одну ошибку 1-го типа:

$$1 - (1 - 0,05)^{10} = 1 - (0,95)^{10} = 1 - 0,60 = 0,40.$$

Чтобы избежать проблем, связанных с проведением многократных проверок по критерию Стьюдента в случае однофакторных планов, исследователи используют особую процедуру, носящую название «метод ANOVA» (ANalysis Of VAriance — дисперсионный анализ). «Однофакторный» означает наличие одной независимой переменной. По сути, однофакторный метод ANOVA проверяет наличие некоторого «общего» значения среди различных значений независимой переменной. Так, в исследовании с тремя уровнями нулевая гипотеза будет следующей: «уровень 1 = уровень 2 = уровень 3». Однако отклонить нулевую гипотезу еще не значит понять, какой из знаков равенства использован ошибочно. Чтобы точно определить

общее значение, необходимо провести так называемую «последующую проверку» или «анализ *post hoc*» (после факта). В ходе последующей проверки для исследования с тремя уровнями после того, как общий анализ *ANOVA* обнаружит наличие такого значения, будет проведен анализ всех трех сравниваемых пар. Если *ANOVA* не находит общего значения, последующая проверка обычно не производится — ее используют только в том случае, если в дальнейшем возникнут определенные предположения насчет конкретной пары условий. В приложении *C* показан однофакторный метод *ANOVA* и распространенный вариант последующего тестирования, названный «*HSD-тест*» Хаки.

В ходе однофакторного метода *ANOVA* вычисляется «оценка F », или «отношение F ». Как же, как коэффициент Стьюдента, отношение F показывает, насколько вероятно то, что найденные различия в значениях среднего арифметического вызваны случайностью или связаны с влиянием других факторов (возможно, независимой переменной). Метод *ANOVA* очень широко применяется психологами, и если на занятиях по статистике вы еще не успели подробно познакомиться с ним, изучите примеры, приведенные в приложении *C*. Как же необходимо знать, что обычно, если независимая переменная принимает всего два значения, используется проверка по критерию Стьюдента. Но этот способ не единственный — в такой ситуации также можно применить однофакторный метод *ANOVA*. По сути, проверку по критерию Стьюдента можно рассматривать как особый случай метода *ANOVA*, применяемый, когда независимая переменная принимает два значения.

Все планы, рассмотренные в данной главе, имеют одну общую особенность — наличие одной независимой переменной. В главе 8 будет сделан следующий логический шаг и рассмотрены планы с несколькими независимыми переменными — «факторные планы».

Резюме

Один фактор — два уровня

Простейший экспериментальный план содержит одну независимую переменную, принимающую два значения (два уровня). Переменная в таком плане может быть межсубъектной или внутрисубъектной. Межсубъектными переменными можно управлять непосредственно или их можно отбирать как субъективные факторы. Если переменная управляемая, испытуемых можно распределить по группам случайным образом (план с независимыми группами) или уравнивать по переменной, являющейся потенциальным осложнителем, а затем случайным образом распределить (план с уравненными группами). В случае субъективной переменной используется план с неэквивалентными группами. Однофакторные планы с внутрисубъектной переменной иногда называют планами с повторяемыми измерениями (как в случае известного исследования Струпа). Для статистической оценки исследований с двумя уровнями независимой переменной обычно проводят проверку по критерию Стьюдента (при условии данных, полученных с помощью интервальной шкалы или шкалы отношений).

Планы с контрольными группами

Планы с контрольными группами подразумевают наличие хотя бы одного условия, при котором не осуществляется экспериментальное воздействие. Есть следующие разновидности контрольных групп: группы плацебо, часто используемые при исследовании лекарств; группы листа ожидания, которые используют при оценке эффективности определенной программы или терапии; а также эквивалентные контрольные группы, участники которых тщательно уравниваются с испытуемыми из экспериментальной группы по определенному фактору, требующему неотрывного контроля.

Один фактор — более двух уровней

При сравнении только двух уровней экспериментальной переменной результаты обязательно будут линейными, так как график, построенный по этим результатам, будет содержать всего две точки. Однако некоторые зависимости нелинейны (например, кривая забывания Эббингхауза) — их можно обнаружить, используя более двух значений независимой переменной. Увеличение количества уровней можно также использовать для проверки и отклонения (фальсификации) альтернативных гипотез. Как и в случае с двумя уровнями, многоуровневые планы могут быть либо межсубъектными, либо внутрисубъектными. Результаты можно наглядно представить с помощью гистограммы, если независимая переменная является дискретной, или в виде линейного графика, если переменная непрерывна. Обычно статистическая оценка исследований, в которых используется более двух значений независимой переменной, проводится с помощью однофакторного дисперсионного анализа (*ANOVA*) (при условии данных, полученных с помощью интервальной шкалы или шкалы отношений).

Задания для повторения

Выбор ответа

1. Чтобы проверить, влияет ли опыт чтения на выполнение задания Струпа, исследовались дети четырех различных возрастов (5, 7, 9 и 11 лет). (Более взрослые дети предположительно имели больший опыт чтения; использовалось только ключевое условие, названное Струпом НЦСр.) Как вы охарактеризуете экспериментальный план?
 - 1) однофакторный многоуровневый с независимыми группами;
 - 2) однофакторный с повторяемыми измерениями;
 - 3) однофакторный многоуровневый с неэквивалентными группами;
 - 4) однофакторный двухуровневый с уравненными группами.
2. Что общего имеют все однофакторные планы с повторяемыми измерениями?
 - 1) они всегда включают контрольную группу;
 - 2) участники исследуются при каждом из экспериментальных условий более одного раза;

- 3) наиболее предпочтительным методом создания эквивалентных групп является уравнивание;
 - 4) каждый участник исследуется при каждом из экспериментальных условий.
3. Какой план иллюстрируется примером, посвященным изучению влияния депривации сна на особенности ответов на наводящие вопросы?
 - 1) внутрисубъектный многоуровневый;
 - 2) с независимыми группами;
 - 3) с уравненными группами;
 - 4) с неэквивалентными группами.
 4. Важнейшая особенность сцепленных контрольных групп состоит в том, что,
 - 1) члены этих групп получают плацебо;
 - 2) на испытуемых в этих группах оказывается такое же воздействие, как на членов экспериментальных групп;
 - 3) по сравнению с членами экспериментальных групп испытуемые в этих группах исследуются при меньших значениях независимой переменной;
 - 4) они используются, если исследователь хочет обнаружить нелинейную зависимость.
 5. Предположим, вы пытаетесь повторить исследование памяти, проведенное Эббингхаузом. В понедельник у вас уходит 20 минут на запоминание списка СВС, а во вторник — 15 минут на его повторение. Какова оценка экономии?
 - 1) 25%;
 - 2) 75%;
 - 3) 5%;
 - 4) ее невозможно вычислить, не зная количество слогов в списке.

Короткие эссе

1. Что общего имеют и чем различаются планы с независимыми группами, планы с уравненными группами и планы с неэквивалентными группами?
2. Почему в примере, посвященном сравнению людей, подвергавшихся и не подвергавшихся депривации сна, был использован план с уравненными группами, а не план с независимыми группами? Какова была переменная уравнивания?
3. Опишите эффект Струпа и использованный автором экспериментальный план.
4. Опишите две разновидности проверки по критерию Стьюдента и на примере планов, изученных в начале данной главы (однофакторные многоуровневые), расскажите о возможностях их использования.
5. С помощью примера, посвященного влиянию алкоголя на скорость реакции, объясните, для чего используются контрольные группы плацебо.
6. На примере исследования «бессознательно воспринимаемых записей» покажите, для чего используются контрольные группы листа ожидания.

7. На примере исследования Вейсса, посвященного изучению взаимосвязей между управлением и стрессом, объясните, для чего нужны эквивалентные контрольные группы.
8. На примере гипотетического исследования влияния кофеина на скорость реакции покажите, как, используя многоуровневые планы, можно обнаружить нелинейный эффект.
9. На примере эксперимента Брэнсфорда и Джонсона, посвященного изучению влияния контекста на запоминание, покажите преимущества использования более двух значений независимой переменной.
10. Опишите, когда лучше всего использовать: а) линейный график, б) гистограмму. Объясните, почему линейный график нельзя использовать в исследовании, в котором сравнивается скорость реакции у женщин и мужчин.

Упражнения

Упражнение 7.1. Определения вида экспериментального плана

Для каждого из описанных ниже исследований укажите независимую и зависимую переменные, особенности первой (межсубъектная или внутрисубъектная; управляемая или субъективная) и назовите, какой экспериментальный план был использован.

1. В исследовании, посвященном изучению влияния булимии на восприятие размера тела, две группы женщин одного возраста (члены одной из групп страдали булимией, а второй — нет) рассматривают серию рисунков, изображающих женщин разного размера, и указывают, какой размер более всего соответствует их восприятию собственного тела.
2. Студентов колледжа, участвующих в исследовании когнитивных карт, попросили с помощью прибора для определения направления точно указать направления, в которых находятся три объекта, не видимые из лаборатории и располагающиеся на разных расстояниях от нее.
3. Три группы дошкольников (случайным образом распределенные по 50 человек в группу) участвуют в исследовании, посвященном изучению настойчивости при выполнении заданий, в котором варьируется время до получения вознаграждения. Детям во всех трех группах раздали трудные головоломки и попросили собирать их, пока не надоест. Одной группе сказали, что по окончании работы все получат по 5 долларов. Вторая группа получит 5 долларов через два дня после окончания эксперимента, а третья — через 4 дня.
4. Для изучения воздействия тесноты на решение задач участников попросили решить серию словесных головоломок, находясь при этом либо в больших, либо в маленьких комнатах. Чтобы получить одинаковое среднее значение вербального IQ в группах, исследователи измерили вербальный интеллект участников, а затем распределили их по двум условиям.

Упражнение 7.2. Результаты

Для каждого из приведенных ниже исследований определите, каким образом представить результаты: в виде линейного графика или гистограммы, а затем создайте график, точно отображающий результаты.

1. В исследовании, посвященном изучению влияния марихуаны на непосредственное запоминание списка слов, участники случайным образом распределены на три группы: экспериментальную группу, контрольную группу плацебо и обычную контрольную группу.

Результат А. Марихуана снижает запоминание, а ожидание действия марихуаны не влияет на запоминание.

Результат В. Марихуана снижает запоминание, но ожидание действия марихуаны также снижает запоминание.

Результат С. Кажущееся неблагоприятное действие марихуаны на запоминание связано только с эффектом плацебо.
2. С помощью надежного и валидного теста исследователь оценивает уровень независимости у трех групп студенток, которые провели в колледже 2 месяца. Высокий уровень независимости означает возможность справляться с делами самостоятельно. Одна группа (О300) состоит из студенток, живущих в общежитии, чей дом находится за 300 или более миль от студенческого городка; студентки из второй группы (О100) живут в общежитии, а их родители живут на расстоянии менее 100 миль от студенческого городка; студентки из третьей группы живут дома (Д).

Результат А. Живущие дома студентки более независимы, чем проживающие в общежитии.

Результат В. Чем дальше дом от студенческого городка, тем более независимым является человек.

Результат С. Живущие дома студентки и члены группы О300 очень независимы, а студентки из группы О100 — нет.
3. Животные изучают лабиринт, и в процессе этого фиксируются совершаемые ими ошибки (т. е. неправильные повороты). В конце каждой попытки, дойдя до цели, они получают пищевое вознаграждение. Одной группе крыс еду выдают сразу по достижении цели (задержка 0). Второй группе еду выдают через 5 секунд после достижения цели (задержка 5 секунд).

Результат А. Задержка подкрепления затрудняет обучение.

Результат В. Задержка подкрепления не влияет на обучение.
4. Игроки в баскетбол совершают три серии по 20 свободных бросков в кольцо, испытывая при этом возбуждение различной силы: слабое, среднее или сильное. Предполагается, что слабое возбуждение возникает, если неудачный бросок наказывается пробежкой вокруг игрового поля (т. е. наказание минимально и не вызывает особого возбуждения). Среднее возбуждение означает два круга вокруг поля, а сильное — четыре круга (т. е. достаточно большое наказание, чтобы создать сильное возбуждение, скорее всего, в виде тревожности). Используется план с повторяемыми измерениями; проведено позиционное уравнивание.

Результат А. Существует линейная зависимость между возбуждением и выполнением задания: рост возбуждения снижает качество выполнения задания.

Результат В. Существует нелинейная зависимость между возбуждением и выполнением задания: задание выполняется хорошо только при среднем уровне возбуждения.