

Роберт Готтсданкер: Основы психологического эксперимента

Роберт Готтсданкер. Глава 2. ОСНОВЫ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА

Если вы хотите экспериментально проверить, помогают ли музыкальные радиопередачи заучивать слова французского языка, вы легко сможете сделать это, повторив один из экспериментов, описанных в предыдущей главе. Скорее всего, вы спланируете свой эксперимент по образцу Джека Моцарта. Вы заранее определите оба условия независимой переменной, будете заниматься в одно и то же время дня и каждый этап эксперимента фиксировать документально. Вместо четырех фортепьянных пьес вы могли бы заучивать четыре списка слов следующим образом: слушая радио, без радио, без радио, с радио. Иначе говоря, вы сможете применять ту же экспериментальную схему, что и Джек.

Вполне возможно, что вам будут понятны некоторые основания ваших собственных действий. Но что-то наверняка останется неясным, и прежде всего — последовательность условий независимой переменной, т. е. сама экспериментальная схема. Вашей вины в этом нет, ведь экспериментальных схем вы еще не проходили. В настоящей главе этот недостаток будет устранен. Конечно, можно провести эксперимент и путем простого подражания образцу, но гораздо лучше понимать то, что вы делаете. Двух идентичных экспериментов не бывает, и слепое копирование экспериментальной схемы часто приводит к затруднениям. Например, Йоко могла бы применить в своем эксперименте регулярное чередование двух условий (сортов сока), как это делали в эксперименте с ткачихами (использование или неиспользование наушников). Но тогда она наверняка знала бы название тестируемого сока, а как раз этого она и старалась избежать»-используя случайную последовательность. К тому же если вы не будете знать оснований различных планов и схем, вам трудно будет оценить качество экспериментов, о которых вы будете читать. А, как вы помните, научить вас этому — одна из главных целей нашей книги.

В данной главе мы сравним те планы, по которым строились эксперименты в главе 1, с менее удачными планами проведения тех же экспериментов. Образцом для их сравнения будет «безупречный» эксперимент (.который практически неосуществим). Анализ такого рода позволят рассмотреть те основные идеи, которыми мы руководствуемся при создании и оценке экспериментов. В процессе этого анализа мы введем в наш словарь несколько новых терминов. В итоге мы определим, что безупречно и что нет в тех трех экспериментальных схемах, которые использовались в главе 1. А эти схемы представляют три способа упорядочивания, или три вида последовательностей предъявления различных условий независимой переменной, применяемые в эксперименте с одним испытуемым.

После изучения материала этой главы вы сможете компетентно и не подражая чужому эксперименту спланировать свой собственный. В конце главы нам будут заданы вопросы по следующим темам:

1. Степень приближения реального эксперимента к безупречному.
2. Факторы, нарушающие внутреннюю валидность эксперимента.
3. Систематические и несистематические источники нарушения внутренней валидности.

4. Методы повышения внутренней валидности, способы первичного контроля и экспериментальные схемы. 5. Некоторые новые термины из словаря экспериментатора.

ПРОСТО ПЛАНЫ И ПЛАНЫ БОЛЕЕ УДАЧНЫЕ

Несомненно, первое условие проведения эксперимента — это его организация, наличие плана. Но не всякий план можно считать удачным. Предположим, что эксперименты, описанные в главе 1, проводились иначе, по следующим планам.

1. Пусть в первом эксперименте ткачиха вначале носила наушники 13 недель, а затем 13 недель работала без них.

2. Положим, Йоко решила использовать в своем эксперименте только по две банки каждого сорта сока, и весь эксперимент занял четыре дня вместо 36.

3. Джек решил применить частичный метод заучивания к первым двум пьесам, а целостный — к двум следующим.

4. Или же, сохранив ту же последовательность методов, Джек выбрал для эксперимента короткие вальсы, а не более длинные пьесы, которые он разучивал обычно.

Мы довольно ясно чувствуем, что по сравнению с ранее описанными экспериментами все эти планы — неудачны. А если бы у нас был образец для сравнения, то мы совершенно точно могли бы сказать, почему именно первоначальные планы были лучше. Безупречный эксперимент служит таким образцом. В следующем разделе мы подробно обсудим его, а затем посмотрим, как он применяется для оценки наших экспериментов.

БЕЗУПРЕЧНЫЙ ЭКСПЕРИМЕНТ

Теперь у нас есть примеры удачно и неудачно спланированных экспериментов. А можно ли и дальше усовершенствовать хорошо спланированный эксперимент? И можно ли сделать эксперимент абсолютно безупречным? Ответ следующий: любой эксперимент можно улучшать до бесконечности, или — что то же самое — безупречный эксперимент провести нельзя. Реальные эксперименты совершенствуются по мере приближения к безупречному.

Идеальный эксперимент

Лучше всего определить безупречность с помощью понятия идеального эксперимента (Кеппел, 1973, с. 23). В идеальном эксперименте допускается изменение только независимой переменной (и, разумеется, зависимой переменной, которая принимает различные значения при разных условиях). Все прочее остается неизменным, и поэтому на зависимую переменную влияет только независимая. В наших трех хорошо спланированных экспериментах это, конечно, не так. Ткачихи носили наушники и работали без них в разное время — по четным или нечетным неделям. Пьесы, которые заучивал Джек с помощью целостного и частичного методов, тоже были разными. Йоко никогда не выпивала томатный сок обоих сортов в один и тот же день. В каждом случае помимо независимой переменной изменялось что-то еще. В последующих главах мы расскажем об экспериментах Другого типа, в которых для каждого из условий независимой переменной привлекаются разные испытуемые, что позволяет устранить временные изменения (вроде четных и нечетных недель) и различия с заданиях (заучиваемые пьесы). Но и они не отвечают всем требованиям идеального эксперимента,

ведь испытуемые тоже будут разными. Как вы скоро убедитесь, идеальный эксперимент невозможен. Однако сама идея является полезной, именно ею мы руководствуемся при совершенствовании реальных экспериментов.

В идеальном (неосуществимом) эксперименте ткачиха должна была бы работать с наушниками и без них в одно и то же время! Джек Моцарт одновременно заучивал бы одну и ту же пьесу целостным и частичным методами. В обоих этих случаях разница в значениях зависимой переменной была бы обусловлена только независимой переменной, различием ее условий. Иначе говоря, все побочные обстоятельства, все другие потенциальные переменные оставались бы на одном и том же неизменном уровне.

Бесконечный эксперимент

Бедная Йоко! В ее случае даже идеальный эксперимент не будет безупречным. Недаром она опасается, что в разных банках томатный сок одного и того же сорта различается по качеству. Даже если бы она провела идеальный эксперимент, ухитрившись одновременно из одного и того же стакана выпивать сок двух разных сортов, ее оценки все равно относились бы только к частным примерам каждого сорта. И тем не менее Йоко могла бы устранить влияния изменчивости качества сока в разных банках, совершив иной невозможный подвиг. «Все», что ей нужно, — это не прекращать свой эксперимент после 36 дней и продолжать его до бесконечности. Тогда она смогла бы усреднить не только изменчивость каждого из сортов сока, но и возможные колебания в собственных оценках его вкусовых качеств. Это и есть бесконечный эксперимент. Нетрудно видеть, что он не только невозможен, но и бессмыслен. Ведь общий смысл эксперимента заключается в том, чтобы на базе ограниченного количества данных делать выводы, имеющие более широкое приложение. Однако бесконечный эксперимент, подобно идеальному, также служит нам руководящей идеей.

Собственно говоря, Джеку Моцарту и авторам исследования в ткацком цехе тоже можно было бы предложить провести бесконечный эксперимент вместо идеального. Ведь даже если в идеальном эксперименте Джек обнаружит, что для данной конкретной пьесы частичный метод более эффективен, останется вопрос, сохранятся ли преимущества этого метода при разучивании других пьес. Те же сомнения вызывает и первый эксперимент: а что если ткачиха лучше работала с наушниками только во время его проведения? Однако их (и вас) нужно предупредить, что бесконечный эксперимент тоже имеет недостатки. Сам факт предъявления испытуемым одного из экспериментальных условий может сказываться (в период исследования) на их работе при другом условии. Возможно, частичный метод был более эффективным во время эксперимента только за счет контраста с целостным методом. А после эксперимента будет применяться единственный метод, и фактор контраста исчезнет. Все это доказывает, что полностью безупречными не являются ни идеальный, ни бесконечный эксперименты. К счастью, они имеют не только разные недостатки, но и разные преимущества и могут служить для оценки реальных экспериментов, весьма далеких от безупречного.

Эксперимент полного соответствия

Ни идеальный, ни бесконечный эксперименты не позволяют устранить недостатки неудачного варианта исследования Джека Моцарта — заучивания вальсов вместо сонат. В лучшем случае Джек мог бы провести блестящий эксперимент на вальсах — что, однако, не сделает их сонатами!

Чтобы совершенно исключить недостатки такого рода, нужен эксперимент полного соответствия. Этот эксперимент также является бессмысленным, хотя практически он осуществим. В своем исследовании Джек должен был бы заучивать те же самые пьесы, которые он будет разучивать и после него. Никакой пользы от такого эксперимента нет, как и от бесконечного. Но зато уж никто не сможет указать Джеку на несоответствие пьес, которые он разучивал в своем эксперименте.

Все три вида безупречного (почти) эксперимента нереальны. Идеальный эксперимент является невозможным, эксперимент полного соответствия — бессмысленным, а бесконечный — тем и другим вместе. Полезны они как эксперименты «мысленные». Они подсказывают нам, что нужно делать для создания эффективного эксперимента. Идеальный и бесконечный эксперименты показывают, как избежать посторонних влияний и тем самым добиться большей уверенности в том, что экспериментальные результаты действительно отражают связь независимой и зависимой переменных. Эксперимент полного соответствия напоминает о необходимости контроля других важных переменных эксперимента, которые мы сохраняем неизменными.

ОБОБЩЕНИЕ, РЕПРЕЗЕНТАТИВНОСТЬ И ВАЛИДНОСТЬ

Как мы установили в главе 1, цель любого экспериментального исследования — сделать так, чтобы выводы, основанные на ограниченном количестве данных, оставались достоверными за пределами эксперимента. Это называется обобщением. Выполненный нами анализ безупречного эксперимента показывает, что достоверность экспериментальных выводов определяется по крайней мере двумя требованиями. От них же зависит и правомерность возможных обобщений. Первое требование заключается в том, чтобы найденное в эксперименте отношение между независимой и зависимой переменными было свободным от влияния других переменных. Второе требование — чтобы постоянный уровень дополнительной переменной, задействованной в эксперименте, соответствовал ее уровню в более широкой области практики.

Репрезентативность

Мы уже знаем, что безупречный эксперимент невозможен, но он дает нам руководящие принципы для правильного планирования реальных экспериментов. Теперь мы можем задать вопрос о применении этих принципов. Ответ прост — нужно определить, насколько успешно реально проведенный эксперимент репрезентирует (представляет) эксперимент безупречный. Прежде всего посмотрим, в какой мере исключена в наших экспериментах возможность посторонних влияний на зависимую переменную.

В оригинальном исследовании, проведенном в ткацком цехе, испытуемая 13 недель работала с наушниками и 13 чередующихся с ними недель — без наушников. При «неудачном» пересмотре эксперимента она носила наушники в первые 13 недель, а следующие 13 работала без них. В идеальном эксперименте испытуемая должна была бы работать и с наушниками, и без них одновременно. Понятно, что схема чередования недель приближается к этому идеалу в, большей степени. Чередование двух условий, или АБАБАБАБАБ и т. д., более репрезентативно для их одновременного предъявления, чем последовательность, состоящая только из А и Б.

В своем первоначальном эксперименте Джек Моцарт разучивал пьесы в следующем порядке: целостный метод — частичный — частичный — целостный. В «неудачном»

эксперимент последовательность была иной: 5целостный — целостный — частичный — частичный. В первом случае усредненные позиции целостного и частичного методов были одинаковы. Целостный метод занимал в последовательности позиции 1 и 4 среднее — 2,5. Позициями частичного метода были 2 и 3, среднее — 2,5. Напротив, в «неудачном» эксперименте целостный метод занимал позиции 1 и 2, среднее — 1,5, а частичный — 3 и 4, среднее — 3,5. Более репрезентативным для одновременного предъявления двух условий вновь оказался оригинальный эксперимент.

В первоначальном варианте своего эксперимента Йоко выпивала оба сорта сока — «Риттенхауз» и «БаддинБидл» — в случайном порядке в течение 36 дней. В "неудачно" измененном варианте она окончилась 4 днями. Ясно, что к бесконечности ближе 36, а не 4. Первоначальный план лучше репрезентирует бесконечный эксперимент, чем план измененный.

Эксперимент полного соответствия лучше представлен в оригинальном исследовании Джека, чем в его измененном варианте с вальсами. Хотя Джек и не разучивал все пьесы, которые он намеревался выучить в дальнейшем, он взял пьесы точно того же типа, т. е. выбрал соответствующий уровень дополнительной переменной. А вариант с вальсами оказывается «неадекватным», поскольку по своему уровню эти пьесы отличаются от тех, которые Джек разучивал бы в эксперименте полного соответствия.

Подводя итоги, можно сказать, что более надежную информацию об отношении между независимой и зависимой переменными дают те эксперименты, которые лучше представляют идеальный и бесконечный эксперименты. А чем ближе уровень значимой дополнительной переменной в проведенном эксперименте к ее уровню в эксперименте полного соответствия, тем лучше представлена в нем изучаемая реальная ситуация.

Валидность

В зависимости от того, насколько реальные эксперименты представляют безупречный, их называют более или менее валидными. Безупречный эксперимент позволил бы безошибочно отделить верную гипотезу от неверной. Если бы Джек Моцарт смог провести безупречный эксперимент, он бы совершенно точно знал, какая из его гипотез верна: частичный метод лучше или целостный метод лучше. Таким образом, говоря о валидности эксперимента, вы оцениваете качество той работы, которую предполагаете провести для определения справедливости одной из конкурирующих гипотез.

Внутренняя валидность. Всем трем описанным нами «неудачным» экспериментам недоставало внутренней валидности. Это значит, что они не позволяют рассматривать полученную картину отношений между независимой и зависимой переменными как достоверную. И виноваты в этом, как мы убедились, всевозможные посторонние влияния. Эксперимент, которому недостает внутренней валидности, не может быть использован для выяснения того, какая гипотеза о связи независимой и зависимой переменных истинна, а какая ложна. Например, если нам не ясно, почему ткачиха работала лучше: потому, что она носила наушники, или потому, что стояла хорошая погода, — мы не вправе считать результаты эксперимента достаточными для определения истинной и ложной гипотез о влиянии наушников на производительность труда.

Термин "внутренний" подчеркивает существо данного вида валидности. Можно сказать, что эксперимент, лишенный внутренней валидности, неудачен, так сказать, изнутри, по самой своей сути. Действительно, если он не позволяет убедиться в достоверности найденного отношения независимой и зависимой переменных, он просто бесполезен.

Внешняя валидность. «Неадекватный» эксперимент, который мог бы провести Джек, разучивая вальсы вместо сонат, не был бы неудачным в принципе. Это был бы вполне нормальный эксперимент по заучиванию вальсов. Бесполезным его считать нельзя. Джек мог бы использовать свои результаты, если бы задним числом решил, что на самом деле он искал наиболее эффективный метод заучивания вальсов. Однако этому эксперименту недостает внешней валидности. Он не обеспечивает достаточных оснований для определения верной и неверной гипотез о лучшем методе заучивания сонат.

Термин «внешний» относится к определению тематики проводимого эксперимента — чему именно он посвящен. В данном случае эксперимент не был внешне валидным потому, что "сонаты" — такая же необходимая составная часть проверяемой гипотезы, как независимая и зависимая переменные.

Общие определения. Понятия внешней и внутренней валидности являются центральными для всей нашей книги. Их применение в последующих главах в основных чертах определяется тем, что мы только что сказали. Сейчас мы приведем и более формальные определения этих понятий. Правда, все их значение вы поймете только тогда, когда познакомитесь с экспериментальными проблемами более высокого порядка. Но у вас уже будет основа для общего понимания и дальнейшего уточнения того, что такое валидность и два ее вида.

Начнем со схематического изображения экспериментальной гипотезы:

Независимая переменная ... Отношение ... Зависимая переменная ... Уровни других переменных. Итак, гипотеза включает в себя само отношение и обозначения обеих его сторон. Определение валидности эксперимента, как внутренней, так и внешней, состоит в следующем. Это степень правомерности вывода об экспериментальной гипотезе, которую обеспечивают результаты данного эксперимента по сравнению с результатами эксперимента, безупречного во всех трех аспектах.

Понятие внутренней валидности эксперимента касается только самого отношения и не затрагивает того, что именно соотносится. Отсюда внутренняя валидность — это степень правомерности вывода об экспериментальной гипотезе, основанного на результатах данного эксперимента, по сравнению с тем выводом, в основе которого лежат результаты идеального и бесконечного экспериментов, где изменения независимой и зависимой переменных происходят в одних и тех же условиях, а все другие, побочные факторы остаются неизменными.

Любой эксперимент сталкивается также с проблемой соответствия исследуемой ситуации — реальной. Вопрос о соответствии уровня дополнительной переменной, типа музыки, уже возникал. Несколько позже мы обсудим подобные вопросы для независимых и зависимых переменных. Понятно, что вопросы о соответствии касаются содержания того, что стоит по обе стороны изучаемого отношения. Это и есть вопросы внешней валидности. Ее можно определить как степень правомерности данного вывода об экспериментальной гипотезе по сравнению с тем выводом, который основан на результатах эксперимента с полным соответствием независимой, зависимой и уровнями всех дополнительных переменных.

В настоящей главе мы обсудим главным образом проблему внутренней валидности. В любом эксперименте вы с самого начала столкнетесь с этой проблемой; если внутренняя валидность не достигнута, рассматривать внешнюю не имеет смысла. Вспомните, что в главе 1 были представлены эксперименты такого типа, для которых вопросы внешней

валидности практически не обсуждаются. А в следующей главе мы рассмотрим эксперименты, в которых именно эти вопросы выступают на первый план.

Никаких гарантий. Мы можем сказать, что эксперимент валиден, не зная фактически, правильны ли выводы. Мы можем сказать, что он невалиден, не зная, что выводы ошибочны. Причина в том, что мы не можем знать наперед, какая из двух конкурирующих гипотез верна. Ведь если бы мы об этом знали, нам не надо было выпроводить эксперимент. Если бы Джек заранее знал, какая из двух его гипотез справедлива: (1) частичный метод лучше или (2) целостный метод лучше, — он мог бы не проводить свое исследование.

При определении валидности реальных экспериментов мы должны сравнивать сами процедуры их проведения с процедурами «проведения» безупречного эксперимента. Валидный эксперимент представляет безупречный эксперимент лучше, чем невалидный. Следовательно, в валидном эксперименте мы с большей вероятностью можем получить такие результаты, которых могли бы достичь в безупречном эксперименте. При этом важно помнить, что ограниченных — и всегда несовершенных — экспериментальных данных связано с риском. Даже самый высоковалидный эксперимент может дать неточную информацию о правильности экспериментальной гипотезы, а информация, полученная в невалидном эксперименте, может оказаться точной. Причины такого риска и его влияние на интерпретацию экспериментальных результатов мы обсудим в следующих главах, прежде всего — в главе 6 («Значимые результаты»).

ФАКТОРЫ, УГРОЖАЮЩИЕ ВНУТРЕННЕЙ ВАЛИДНОСТИ

Теперь мы можем применить понятие безупречного эксперимента (идеального и бесконечного) для описания того, что мешает достижению внутренней валидности в реальных экспериментах. Как мы увидим, некоторые из таких помех устранить нельзя; они необходимо связаны с процедурами проведения наших не-вполне-безупречных экспериментов. Скажем, если Джеку нужно разучить две пьесы, одну из них он неизбежно будет разучивать первой. Существуют, однако, и такие трудности, которые можно преодолеть, если заранее об этом позаботиться. Так, Джек уже знал, что не стоит применять частичный и целостный методы в разное время дня.

Изменения во времени

Известные побочные факторы. В идеальном эксперименте различные состояния независимой переменной предъявляются испытуемому одновременно. Этого Джек сделать не мог, но он мог по крайней мере заниматься в одно и то же время дня. Время дня — это заранее известная побочная (т. е. отличная от независимой) переменная, которая может повлиять на эффективность занятия, и ее надо охранять неизменной. Если бы Джек был невнимателен, то в разные дни эксперимента он мог бы заниматься то при закрытых, то при открытых окнах. А уличный шум может сильно повлиять на эффективность занятий. Поэтому лучше сохранять его неизменным, держа окна закрытыми. В эксперименте с наушниками, который продолжался более шести месяцев, исследователи знали о возможных изменениях температуры и влажности в ткацком цехе. К сожалению, условия эксперимента не позволяли им исключить эти изменения. Но экспериментаторы фиксировали и старались учесть влияния названных факторов. И что самое главное — чередование двух условий независимой переменной снижало влияние этих факторов. Экспериментатор должен стараться заранее определить все возможные факторы, которые с течением времени могут изменяться. И главное, стараться удерживать их на постоянном уровне при каждой новой пробе.

Нестабильность во времени. Но даже стараясь изо-всех сил, экспериментатор не сумеет сделать одну пробу в точности (кроме отличия уровней независимой переменной) похожей на другие. Некоторая нестабильность во времени будет всегда. В эксперименте она проявляется в изменчивости побочных факторов, а также в некоторых вариациях самой независимой переменной. Наконец, всегда остаются совершенно неясные источники сильных колебаний в ответах испытуемых, приводящие к увеличению разброса экспериментальных данных. Давайте рассмотрим конкретные примеры каждой из этих трех форм нестабильности во времени.

Изменчивость побочных факторов. Часто бывает так, что экспериментатор знает о существовании посторонних факторов, влияющих на зависимую переменную, но не может управлять ими непосредственно. Какой-то день в работе ткачихи мог оказаться «не самым удачным» из-за того, что накануне она поздно легла спать. Конечно, экспериментатор мог бы попытаться убедить ее не делать этого, пока эксперимент не завершится. Но ведь эксперимент продолжался шесть месяцев! Поужинав накануне в ресторане, Джек неважно себя чувствовал во время разучивания одной из пьес — в другой раз ему следует быть осторожнее.

От пробы к пробе окружающие условия никогда не остаются неизменными. Описывая эксперимент в ткацком цехе, исследователи утверждают:

«Хорошо известно, что на производительность ткацкого труда могут влиять атмосферные условия. Так, с повышением температуры и относительной влажности уменьшается количество обрывов нити. С другой стороны, дальнейшее повышение того и другого, продолжая благоприятно воздействовать на физические свойства пряжи, неблагоприятно сказывается на физиологическом состоянии людей, работоспособность которых может снизиться так, что это сведет на нет любые положительные влияния» (Уэстон и Адаме, 1932, с. 56).

Следовательно, даже измеряя температуру и влажность, нельзя установить точно их влияние на производительность труда. Список побочных переменных можно было бы продолжать до бесконечности, включая в него и субъективные факторы, как, например, хорошее или плохое самочувствие испытуемого в течение эксперимента. Добросовестный экспериментатор может фиксировать некоторые из этих изменений, но не может их избежать. Теперь вам понятно, почему экспериментатор стремится уйти из реального мира в прекрасные звуконепроницаемые лаборатории и иметь дело с такими испытуемыми (белыми крысами), поведение которых он может контролировать 24 часа в сутки. Но даже там калориферы иногда остывают, бутылки с водой засоряются, и крысы подхватывают «насморк».

Само пребывание в ситуации эксперимента может вызвать продолжительные изменения в поведении испытуемого. Таков был главный вывод из знаменитых экспериментов «Хауторн», вывод, важный для всех экспериментальных психологов. На Западном электрозаводе в Хауторне (Иллинойс) было проведено исследование влияния освещения в цехах на производительность сборочных работ. Предварительные попытки установить какую-либо закономерность закончились неудачей. Тогда было предпринято систематическое исследование условий труда рабочих (Роетлисбергер и Диксон, 1946). Основной частью этого исследования были эксперименты с заданием по сборке переключателей. Оно представляло собой «сборку телефонных реле; это операция, которую обычно выполняют женщины: нужно соединить примерно 35 небольших деталей в «сборную арматуру» и закрепить ее четырьмя винтами» (с. 20).

Для эксперимента была оборудована специальная комната, чтобы исследователи могли контролировать условия работы и адекватно оценивать деятельность операторов. В качестве испытуемых в эксперименте приняли участие пять молодых женщин, вполне освоивших данный вид работы. Исследовались две независимые переменные: распределение периодов отдыха, а также длина рабочего дня и рабочей недели. Оплата труда производилась в соответствии с общим количеством переключателей, собранных бригадой из пяти человек.

Было установлено, что независимо от распределения периодов отдыха и длины рабочего дня и недели производительность труда продолжала расти в течение двух лет! Исследователи сообщают, во-первых, о «постепенном изменении социальных отношений в группе операторов в направлении групповой сплоченности и солидарности и, во-вторых, об изменении отношений между операторами и их контролерами. Организаторы эксперимента стремились создать среди девушек атмосферу взаимной поддержки и сотрудничества, избавить их от излишних волнений и тревог. Эти усилия по созданию необходимых условий эксперимента косвенным путем привели к изменению отношений между людьми» (с. 58—59).

Пользуясь нашей терминологией, эту ситуацию можно описать так. До эксперимента социальные условия работы испытуемых находились на одном уровне. В ситуации эксперимента данная «побочная переменная» перешла на другой уровень. Это привело к продолжительному изменению зависимой переменной — производительности труда, несмотря на то что объективно социальные условия в эксперименте оставались неизменными.

Независимая переменная. Рассчитывать на полную идентичность каждого из условий независимой переменной на всем протяжении эксперимента мы не можем. В какие-то дни или даже недели наушники могли быть надеты не так удобно, как в остальные. Несмотря на все старания Джека, он может по-разному относиться, например, к частичному методу, разучивая разные пьесы. И Йоко знала о вариациях каждого из условий своей независимой переменной. Сок одного и того же сорта в любых двух банках не бывает одинаков, и раз-яйца подчас очень велика. Какие-то изменения будут встречаться даже в тех экспериментах, в которых, казалось бы, достигнуто полное единообразие условий. Яркость электрического света (как стимула) будет меняться от перепадов напряжения в сети, а они случаются довольно часто. В течение эксперимента могут возникать и закономерные изменения, например с увеличением срока службы лампочки ее свет может становиться все менее ярким.

Зависимая переменная. При действии одной и той же независимой переменной испытуемый не всегда будет давать один и тот же ответ. Так будет даже в том случае, если экспериментатор необычайно искусен и пунктуален в устранении нестабильности побочных факторов и независимой переменной.

Нестабильность зависимой переменной очень эффектно представлена на графиках, отражающих результаты двух экспериментов. На рис. 2.1 показана недельная выработка испытуемой Д. в эксперименте с наушниками. Как видим, меньше всего ударов она пропускала с десятой по двенадцатую неделю и с восемнадцатой по двадцать вторую. А самые неудачные ее показатели — наибольшее количество пропущенных ударов — приходятся на четырнадцатую неделю и окончание эксперимента. И что особенно интересно — для обоих условий работы кривые поднимаются и опускаются совместно. Изменения в производительности труда, происходившие с течением времени, несомненно, более существенны, чем различия между использованием и неиспользованием наушников.

На рис. 2.2 показаны изменения ответов испытуемого в эксперименте на время реакции выбора. Пробы давались каждые шесть секунд; испытуемый должен был сдвигать рукоятку к себе или от себя и тем самым совмещать две световые точки. Разумеется, точки предъявлялись в случайном порядке. За 70 запланированных последовательных проб во времени реакции испытуемого наблюдались как непродолжительные колебания, так и более регулярные отклонения. Самое короткое время реакции было показано примерно между тридцатой и сороковой пробами, а самое большое — между шестидесятой и семидесятой. И это увеличение нельзя считать результатом усталости, ведь как раз перед сороковой пробой испытуемый отдыхал. В итоге самые большие показатели несколько превышали 400 мс, а самые малые — 200 мс, т. е. время реакции изменялось в отношении два к одному.

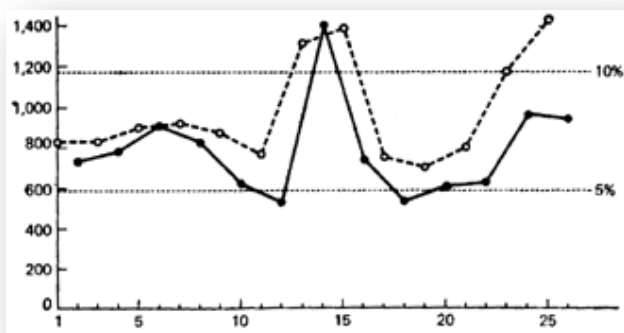


Рис. 2.1. Понедельная производительность труда испытуемой Д. Ось абсцисс — очередность недель проведения эксперимента. Ось ординат — количество пропущенных ударов (в среднем за час). Пунктирная линия — работа без наушников, сплошная — с наушниками

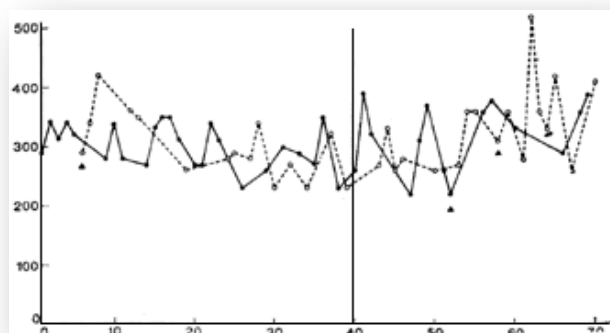


Рис. 2.2. Время реакции выбора для 70 последовательных проб. Ось абсцисс — номера проб (пунктиром отмечен период отдыха). Ось ординат — время реакции (в мс). Пунктирная линия — сдвиг рукоятки к себе, сплошная — от себя; треугольниками помечены ответы с ошибками

Таким образом, в исследовании времени реакции были обнаружены поминутные и даже посекундные его изменения. Они не связаны с усталостью, скорее их можно объяснить колебаниями внимания. На графике результатов ткачихи Д. отражены значительные колебания производительности ее труда. При этом взлеты и падения кривых, по-видимому, не зависят от температуры и влажности. Правда, увеличение числа пропущенных ударов к окончанию эксперимента можно объяснить применением искусственного (газового) освещения; оно было необходимым, так как эксперимент заканчивался осенью.

Даже тогда, когда сами ответы испытуемых постоянны, изменения могут вноситься процедурой их измерения. Счетчик фиксирует каждое передвижение челнока, совершающего новый удар. Однако приборы не всегда бывают исправны. А если измерения связаны с субъективными суждениями, то они непременно окажутся менее стабильными. Джек считал пьесу выученной окончательно после двух безошибочных исполнений наизусть. Однако в исполнении пьес было довольно много небольших погрешностей, почти ошибок. Иногда Джек мог посчитать их ошибками, а иногда нет. И объяснилось это вполне естественными колебаниями его субъективного состояния. Изменения в оценке исполнения пьес могли быть и регулярными. Например, в процессе эксперимента Джек мог все строже относиться к своим ошибкам.

Различия экспериментальных задач

Одну и ту же пьесу нельзя заучить (как в идеале) двумя различными методами одновременно. Но даже если методы следуют один за другим, их все равно нельзя применить к одной и той же пьесе. Если пьеса заучена, она заучена. Существуют эксперименты, в которых необходимо не только предъявлять разные экспериментальные условия в разное время, но и менять трудность задач. Это весьма значительное отличие от идеального эксперимента. Как Джек может убедиться в одинаковой трудности выбранных им пьес? А ведь в любом эксперименте по исследованию научения с участием одних и тех же испытуемых задачи для разных условий независимой переменной с необходимостью будут разными.

Эффекты последовательности

В неудачном варианте своего эксперимента Джек сначала заучивал две пьесы частичным методом, а затем две другие — целостным. Мы уже знаем, что на качество его игры могут повлиять любые (в том числе только что описанные) факторы, изменяющиеся во времени. Однако есть и другие влияния, связанные с положением каждого из условий независимой переменной в последовательности их предъявления. Влияния одного из условий на следующие за ним называются эффектами последовательности, эффектами порядка или эффектами переноса. Они могут быть положительными и отрицательными, иметь общий и специфический характер. Применение частичного метода могло положительно повлиять на дальнейшие занятия Джека по целостному методу за счет увеличения практики или привыкания к режиму эксперимента. Оно могло оказать и отрицательное действие: привычка заучивать пьесы короткими отрывками могла помешать заучиванию больших частей или Джек мог просто устать от занятий.

Предубеждения экспериментатора

Во времена возникновения автомобиля бытовал такой анекдот в форме загадки. Вопрос: Какой самый главный винтик в автомобиле? Ответ: Тот, что держит руль. В том же духе можем спросить и мы. Вопрос: Какой из факторов, угрожающих валидности эксперимента, самый опасный? Ответ: Экспериментатор.

Если у исследователя есть какие-либо ожидания относительно результатов эксперимента, особенно связанные с предпочтением одного из условий независимой переменной, то эти ожидания так или иначе проявятся при проведении опытов, Йоко хорошо знала, что главное — это составить случайную последовательность обоих сортов сока. Ей хотелось исключить любой намек на то, какой именно сорт она оценивала каждое утро. А вот Джек не проявил должной аккуратности. Сначала он подобрал пары пьес, которые казались ему одинаковыми по трудности (с тем, чтобы разучивать каждую из пьес разными методами), а затем сам расположил их в определенной последовательности. Но если при этом он рассчитывал на большую эффективность частичного метода, то мог невольно отобрать для целостного метода более трудные пьесы из каждой пары.

Кроме того, субъективные оценки качества исполнения пьес могли колебаться не случайным образом (как было показано выше). Джек мог невольно отдавать предпочтение одному из методов. Поэтому при оценке исполнения обеих пьес каждой пары Джеку не следует слишком верить в частичный метод, но и применяя целостный, тоже стараться достичь самых высоких результатов.

В эксперименте с наушниками исследователи, естественно, рассчитывали повысить с их помощью производительность труда и вполне могли передать свою уверенность участникам эксперимента. Поэтому, возможно, с наушниками ткачихи (в среднем) старались работать лучше.

Одно из самых коварных следствий предубежденности экспериментатора — это нежелание учитывать некоторые экспериментальные данные, как полученные якобы при нетипичных условиях, например при сильном уличном шуме. К сожалению, мнение экспериментатора о нетипичности условий часто весьма субъективно. Отсюда один и тот же уровень шума будет считаться нетипичным при одном состоянии независимой переменной, но вполне нормальным — при другом.

От предубеждения экспериментатора может зависеть даже точность регистрации данных. Показано, например, что в протоколах экспериментов по изучению экстрасенсорного восприятия встречаются ошибки в пользу наличия соответствующих феноменов, если протоколист верит в их существование. Те же, кто не верит в экстрасенсорное восприятие, не допускают подобных искажений (Кеннеди, 1939). Тщательный анализ этой проблемы в целом представлен в книге "Влияния экспериментатора в психологическом исследовании" (Розенталь, 1976).

ПОВЫШЕНИЕ ВНУТРЕННЕЙ ВАЛИДНОСТИ С ПОМОЩЬЮ ПЕРВИЧНОГО КОНТРОЛЯ

Термин «контроль» используется в нашей книге для обозначения любого способа усовершенствования экспериментов, который приближает их к безупречному. Здесь мы будем иметь дело с угрозами внутренней валидности эксперимента и мерами, которые позволяют их устранить. Если контроль в эксперименте недостаточен, то внутренняя валидность может нарушиться двояким образом. Она может пострадать от ненадежности и от систематического смещения. Ниже мы рассмотрим способы первичного контроля (применимые к любой экспериментальной схеме), которые могут повысить надежность эксперимента и сократить систематическое смещение.

Ненадежность

Предположим, что исследование в ткацком цехе проводилось лишь в течение 11-й и 14-й недель из 26. Для выбора, в какую из этих недель давать наушники, просто бросали монету (чтобы избежать предубеждений) Случилось так, что выбор пал на четырнадцатую неделю. А теперь посмотрите на рис 2.1. На этой неделе испытуемая Д. работала плохо: она пропускала более 1400 ударов за час. С другой стороны, одиннадцатая неделя была для нее удачной: менее 800 пропущенных ударов за час. Объективность такого эксперимента сомнений не вызывает, он просто слишком короткий. За две другие недели ткачиха покажет иные результаты и т.д. Двухнедельный эксперимент весьма далек от бесконечного, а 26-недельный гораздо ближе к нему, и у нас есть все основания надеяться, что другой эксперимент в 26 недель даст примерно те же результаты. Конечно, если разброс будет невелик, то большего количества проб не потребуются.

Когда мы говорим о надежности среднего значения зависимой переменной для каждого из условий независимой переменной, то имеем в виду его устойчивость при повторении эксперимента второй, третий, четвертый и т.д. раз. Итак, если надежность высокая, то при многократном повторении эксперимента будет воспроизводиться примерно одно и то же значение зависимой переменной.

Выбор адекватного количества проб. Для достижения высокой надежности эксперимента количество проб должно соответствовать изменчивости изучаемого поведения. В эксперименте с временем реакции, как можно судить по рис. 2.2, требуется много проб — 50, 100 или даже больше. Ведь экспериментальные данные сильно варьируют от пробы к пробе. Джеку Моцарту в эксперименте с заучиванием фортепьянных пьес проб нужно гораздо меньше, поскольку каждая проба включает довольно много отдельных действий, и результат измерения оказывается более стабильным. Таким образом, необходимое количество проб зависит от разброса результатов измерений.

Сокращение изменчивости во времени. Увеличение количества проб — лишь один из способов повышения надежности. Другой — сокращение самой изменчивости изучаемого поведения. Достигается оно прежде всего с помощью организации хода эксперимента и его протоколирования, уже обсуждавшихся в главе 1. Процедура эксперимента должна строго соблюдаться, а значимая информация о ней (учитывая возможность ошибок при воспроизведении по памяти) — фиксироваться документально. Кроме последовательного выполнения запланированных действий уменьшения разброса экспериментальных данных можно достичь, соблюдая точность в эксперименте.

Любые меры, которые позволяют сделать эксперимент более точным, сокращают изменчивость поведения во времени. Так, чтобы не увеличивать эту изменчивость за счет поведения экспериментатора, везде, где это возможно, применяются автоматические методы. Например, в эксперименте с временем реакции выбора экспериментатор сам мог бы давать испытуемому команду «Приготовьтесь!» примерно за секунду до предъявления сигнала. Разумеется, время подачи такой команды, громкость голоса экспериментатора не могут не изменяться. Поэтому гораздо лучше использовать для предупреждения испытуемого прибор-автомат, зажигающий стандартную световую точку. Далее, экспериментатор мог бы измерять время реакции с помощью обычного секундомера: включать его при предъявлении сигнала и останавливать, увидев, что испытуемый дал ответ. Но это, несомненно, увеличило бы различие между пробами. Использование же электрических или электронных часов, которые автоматически включаются вместе с сигналом и останавливаются с выдачей ответа, позволяет практически предотвратить подобные вариации

Из сказанного вполне понятно, что любые способы стабилизации побочных переменных повышают надежность эксперимента. Джеку Моцарту, конечно, не следует во время своих занятий то закрывать окна, то широко раскрывать их. Если у него есть возможность поддерживать в комнате нормальную температуру, то окна лучше всегда держать закрытыми, чтобы уменьшить шум транспорта и, что более важно, сократить колебания в уровне этого шума. Мы помним, как внимательно следила Йоко за тем, чтобы томатный сок, который она пила, имел постоянную температуру.

Систематическое смещение

Однако на самом деле ненадежность — еще не самый худший вариант нарушения внутренней валидности. В принципе этот недостаток всегда можно исправить путем увеличения числа проб. В тех же случаях, когда этого сделать нельзя, мы используем термин «систематическое смещение». План «неудачного» эксперимента с ткачихами — работа в первые 13 недель с наушниками и следующие 13 без них — не позволяет улучшить эксперимент путем увеличения количества проб. Можно повторить эксперимент полностью, но и тогда останется сомнение, не пришелся ли один из этих 13-недельных периодов на времена, когда зависимая переменная испытывала благоприятные (или неблагоприятные) влияния. Напротив, с помощью плана чередования

недель, который применялся в действительности, можно продлить эксперимент настолько, насколько потребуется для достижения надежности. Следовательно, «неудачный» план приводит к систематическому смещению Независимой переменной (использование или неиспользование наушников) с другими факторами, изменяющимися во времени (в первые или вторые 13 недель).

Систематическое смещение означает, что независимая переменная постоянно сопровождается некоторыми побочными переменными. Использование наушников сопровождается изменениями в состоянии ткачихи за первые 13 недель, их неиспользование — изменениями в ее состоянии за вторые 13 недель. По данным такого эксперимента мы не сможем судить о том, была ли высокая производительность труда (за первые 13 недель) результатом воздействия независимой переменной (использования заглушек) или каких-то других факторов.

Теперь нам понятно, почему описанный эффект (в отличие от ненадежности) назван систематическим. Надежность можно повысить, увеличивая число проб, ведь в течение долгого времени любые благоприятные (или неблагоприятные) влияния окажутся размытыми. Если же эксперимент страдает от систематического смещения, то — с увеличением проб этот недостаток лишь — усугубится.

Довольно много подобных влияний может возникнуть в случае, если не принять необходимых мер против предубеждений экспериментатора. Если экспериментаторы надеялись на повышение производительности труда с помощью наушников, то они могли передать свою веру рабочим, а быть может, и ошибаться в снятии показаний счетчика ударов. Все это явно нарушило бы равновесие между двумя состояниями независимой переменной. Одно из этих состояний (использование наушников) сопровождалось бы одним уровнем побочной переменной — благоприятным предубеждением экспериментатора, а второе состояние (неиспользование наушников) — другим ее уровнем — неблагоприятным предубеждением. При этом экспериментатору вовсе не обязательно проявлять свои склонности постоянно. Даже эпизодические случаи будут порождать систематические влияния, приводящие к неравенству двух состояний независимой переменной.

Довольно много подобных влияний может возникнуть в случае, если не принять необходимых мер против предубеждений экспериментатора. Если экспериментаторы надеялись на повышение производительности труда с помощью наушников, то они могли передать свою веру рабочим, а быть может, и ошибаться в снятии показаний счетчика ударов. Все это явно нарушило бы равновесие между двумя состояниями независимой переменной. Одно из этих состояний (использование наушников) сопровождалось бы одним уровнем побочной переменной — благоприятным предубеждением экспериментатора, а второе состояние (неиспользование наушников) — другим ее уровнем — неблагоприятным предубеждением. При этом экспериментатору вовсе не обязательно проявлять свои склонности постоянно. Даже эпизодические случаи будут порождать систематические влияния, приводящие к неравенству двух состояний независимой переменной.

СХЕМЫ КОНТРОЛЯ ЗА ФАКТОРАМИ ВРЕМЕНИ В ЭКСПЕРИМЕНТЕ С ОДНИМ ИСПЫТУЕМЫМ

В описанных нами исследованиях разные состояния, или условия, независимой переменной давались одному и тому же испытуемому. Каждый раз реальный эксперимент отличался от идеального тем, что условия давались испытуемому в разное время.

Следовательно, в эксперименте нужно каким-то образом контролировать довольно большое количество факторов, которые можно объединить под названием «факторы времени». Для этой цели есть практически только три схемы последовательного предъявления условий, которые нам нужно подробно обсудить. Такие схемы, как предъявление сначала всех проб одного условия, а затем — всех проб второго условия, вроде «неудачного» эксперимента с наушниками, нас вообще не интересуют. Три экспериментальные схемы — это те, что применялись в трех экспериментах из главы 1: случайная последовательность условия эксперимент с томатным соком), их регулярное чередование (наушники) и позиционно уравненная последовательность (заучивание фортепьянных пьес). Сейчас мы разберем основания для применения каждой из этих схем и обсудим, насколько успешно они позволяют контролировать факторы времени.

Схема случайной последовательности

Эта схема особенно хороша для тех экспериментов, где сами пробы достаточно коротки, но в интересах надежности их используется довольно много. В тех случаях, когда испытуемый не должен знать о состоянии независимой переменной в каждой данной пробе, как в эксперименте Йоко, схема случайной последовательности единственно возможна.

Как следует из самого названия схемы, состояния, или условия, независимой переменной предъявляются в случайном порядке. Бросают, скажем, игральную кость и смотрят, на какое число точек (четное или нечетное) она выпала. Если сравниваются два условия, то выпадению нечетного числа может соответствовать условие А, а четного — условие Б. Если количество проб для каждого из условий должно быть одинаковым, то можно распределить их так, как это сделала Йоко, — воспользоваться равным количеством бумажных бланков для того и другого условия. Более строгий способ составления случайной последовательности будет изложен в главе 4.

При использовании этой схемы исключается всякая возможность систематического смещения независимой переменной с факторами времени, поскольку в случайной последовательности никакой системы не существует. С увеличением проб повышается надежность эксперимента.

Схема регулярного чередования

Нам понятно, почему в исследовании с наушниками экспериментаторы использовали регулярное чередование условий, а не их случайную последовательность. Они не знали, сколько времени фабрика будет выпускать один и тот же вид ткани, и хотели, чтобы число недель работы с наушниками и без них было одинаковым. Если распределять условия случайно, работа с наушниками могла начаться, скажем, лишь на 10-ю неделю из первых 15. Кроме того, если с течением времени в самочувствии оператора, состоянии станка, влажности воздуха и т. п. происходили систематические изменения, то одно из условий, более часто попадающее в первую (или вторую) половину эксперимента, могло иметь преимущество перед другим. Поэтому для достижения надежности здесь нужно регулярное чередование, а не случайная последовательность.

Данную схему можно использовать при проведении эксперимента с большим числом проб, если только эти пробы не обязательно распределять случайно (чтобы усреднить влияния факторов времени). А вот в эксперименте Йоко никакой регулярной схемы применить было нельзя: нужно было избежать влияния пристрастности испытуемой к одному из экспериментальных условий.

У схемы регулярного чередования есть только один небольшой недостаток. Что если бы в каждую вторую неделю эксперимента происходило некоторое событие, скажем, очередная наладка станка? Это привело бы к систематическому смещению независимой переменной, которое благоприятствовало бы одному из ее условий, следующему сразу же за этим событием. Напротив, при случайном распределении условий любые влияния факторов времени можно компенсировать, увеличивая продолжительность эксперимента. Ведь для этой схемы более актуален вопрос о ненадежности, чем о систематическом смещении с побочными факторами.

Схема позиционно уравненной последовательности

Позиционно уравненная последовательность условий применяется в тех случаях, когда экспериментатор по той или иной причине не использует большого количества проб (или блоков проб). Именно таким был случай Джека Моцарта. Джек хотел найти более эффективный метод заучивания достаточно быстро, не разучивая в эксперименте слишком много пьес. Поэтому здесь позиционно уравненная последовательность состояла только из четырех проб — АББА (целостный—частичный—частичный—целостный). Аналогично, если экспериментатору нужно сравнить время реакции на появление звуковых и световых сигналов, он мог бы составить последовательность из четырех блоков по 50 проб в каждом. Во всех пробах каждого блока применялся бы либо только звук, либо только свет. Тогда в позиционно уравненной последовательности АББА: звук—свет—свет—звук — каждую позицию занимала бы не отдельная проба, а целый их блок.

При использовании данной схемы, так же как и двух других, надежность эксперимента зависит от количества проб. Мы имеем в виду либо количество отдельных действий, входящих в одну пробу, либо число самих проб в блоке. А как обстоит дело с систематическим смещением независимой переменной? Если в эксперименте применяется схема позиционного уравнивания, то оба условия следуют в среднем через одинаковые временные промежутки. Это позволяет контролировать любые изменения во времени, которые имеют линейный характер. Как отмечалось ранее, в последовательности из четырех проб в эксперименте Джека каждое условие занимало в среднем позицию 2,5. Если предположить, что все возможные изменения нарастают по прямолинейной функции, то данная схема позволяет их хорошо контролировать. Например, если бы с исполнением каждой новой пьесы Джек становился все более сосредоточенным и уровень его внимания возрастал каждый раз на 4 условные единицы, то влияния данного фактора времени оказались бы контролируемы. Скажем, при разучивании первой пьесы уровень внимания Джека имел значение 90, при разучивании второй—94, третьей—98 и четвертой—102. Среднее значение для условия А (пробы 1 и 4) составило бы 96 и среднее для условия Б (пробы 2 и 3) —тоже 96.

Для проверки предположения о прямолинейном характере происходящих изменений нам стоит еще раз посмотреть на рисунки 2.1 и 2.2. Эти изменения совершенно бессистемны и непредсказуемы. Из них можно заключить, что сосредоточенность Джека возрастала от пробы к пробе: на 5 единиц, на 2 и на 1. Тогда ее значения для каждой из четырех проб будут следующими: 90, 95, 97 и 98. Среднее для условия А (90 и 98) составит 94, в то время как для условия Б (95 и 97)— 96. Указанная схема не позволит уравнивать влияния факторов времени для обоих экспериментальных условий. В результате возникнет систематическое смещение независимой переменной с этими побочными влияниями.

КАК КОНТРОЛИРОВАТЬ ФАКТОРЫ ЗАДАЧИ

Вы помните, что если бы Джек смог провести идеальный эксперимент, он заучивал бы одну и ту же пьесу двумя разными методами. Поскольку это невозможно, то самое лучшее — найти пару пьес, одинаковых по трудности. Такая проблема возникает в любом эксперименте, где из-за влияния научения для разных экспериментальных условий нужно использовать разный материал, т. е. разные задачи. Давайте посмотрим, как можно уравнивать влияния факторов, связанных с различием задач (или, короче, факторов задачи), с помощью трех указанных схем, в том числе схемы позиционного уравнивания, которой пользовался Джек.

Случайная последовательность

Начнем с того, что здесь понадобились бы не четыре пьесы, а гораздо больше, пожалуй, даже слишком много, чтобы реализовать эту схему на практике. Предположим, однако, что такой эксперимент можно осуществить. Тогда есть две стратегии подбора различных пьес (задач). Первая — это выбрать для заучивания 30 или 40 пьес и затем расположить их в случайном порядке. Название каждой пьесы можно записать на бланке, положить бланки в коробку, а затем выбирать их случайным образом. Иначе говоря, случайную последовательность пьес можно получить так же, как и случайную последовательность условий независимой переменной. Другая стратегия заключается в разделении пьес на пары по степени их трудности. Если всего отобрано 30 пьес, то сначала составляют пару из двух самых трудных пьес, затем из двух, следующих за ними по трудности, и т. д., получая таким образом 15 пар. В каждой паре путем случайного выбора, т. е. бросая монету, определяют, какая пьеса будет заучиваться каждым из методов. Затем пары можно расположить в порядке возрастания трудности и приступить к заучиванию, начиная с самой легкой пары. Пьесы, отобранные для каждого метода, можно распределить и случайно. Однако при этом влияние разного уровня трудности задач усилится за счет влияния временных изменений, точнее, факторы задачи просто войдут в состав факторов времени. Если пьесы сильно различаются по трудности, то для достижения надежности эксперимента потребуется большое количество проб, но зато не будет никакого систематического смещения.

Регулярное чередование

Поскольку эта последовательность используется в экспериментах с несколько меньшим количеством проб, то следует применять не простое случайное распределение по каждому из методов, а с предварительным разделением пьес на пары. Разучивание лучше всего начинать с двух самых легких пьес, затем переходить к двум несколько более трудным и т. д. При слишком большом различии заданий, снижающем надежность эксперимента, также следует увеличить количество проб. Здесь систематическое смещение независимой переменной с факторами задачи будет отсутствовать.

Позиционно уравненная последовательность

Если экспериментатор использует только одну серию позиционно уравненной последовательности АББА, как это было у Джека, решающую роль играет подобие пьес в каждой паре. Сначала Джек попытался отобрать для каждого метода две длинные и две короткие пьесы. Затем он составил пары пьес, практически одинаковых по трудности. И все же ему не стоило самому принимать решение о том, каким методом заучивать ту или другую пьесу. Для каждого метода он должен был случайно выбрать по пьесе из обеих пар. Тогда он избежал бы невольного отбора более легких пьес для предпочитаемого им

метода. Однако, принимая свое решение, Джек не мог иметь точного представления о трудности пьес, поэтому его выбор можно считать удовлетворительным. Но так или иначе, систематического смещения независимой переменной с факторами задачи в данном случае избежать нельзя. Насколько оно серьезно — зависит от того, в какой мере подтвердится предположение о подобии пьес. Экспериментатору всегда стоит стремиться отбирать такие задачи, которые можно как-то измерить. А это очень непросто. Было показано, например, что даже наборы бессмысленных слогов могут различаться по своей семантике и трудности заучивания.

ПРОБЛЕМА ЭФФЕКТОВ ПОСЛЕДОВАТЕЛЬНОСТИ

Важность этой проблемы все более осознается в экспериментальной психологии (Поултон и Фримен, 1966). Сейчас мы обсудим, почему *эффекты последовательности* — 80это основной фактор, нарушающий внутреннюю валидность в эксперименте с одним испытуемым. Речь идет о тех характеристиках ответа испытуемого в данной пробе, которые вызваны самим фактом предъявления ему предыдущих проб. Эти влияния могут быть положительными и отрицательными. Они могут иметь глобальный характер, как, например, адаптация к режиму эксперимента или усталость испытуемого. Они могут быть и специфическими, скажем, умение предсказывать тот момент, когда в челноке окончится нить. Они бывают недолгими, распространяющимися лишь на одну последующую пробу, и продолжительными, причем их действие от пробы к пробе может накапливаться. Такие влияния обычно называются эффектами переноса.

Однородные и неоднородные эффекты

Рассмотрим эффект переноса, который накапливается по мере предъявления проб. Предположим, что в течение четырех экспериментальных проб (целостный метод—частичный—частичный—целостный) Джек Моцарт постепенно втягивался в режим эксперимента. Если подобное влияние на каждую последующую пробу является однородным, то ни один из методов заучивания не получит преимущества перед другим. Например, если величина положительного переноса каждой предыдущей пробы (или переноса от пробы к пробе) равна 2 «единицам», то значения «помощи» испытуемому были бы следующими: первая проба целостного метода — никакой помощи, первая частичного метода — 2 единицы, вторая частичного метода — 4 единицы, вторая целостного метода — 6 единиц. В итоге на каждый метод, целостный и частичный, пришлось бы по 6 единиц. Таким образом, эффект однородного переноса оказывается уравновешенным.

Однако чаще случается так, что процесс научения сначала протекает более интенсивно, а затем замедляется. Поэтому в нашем примере лучше предположить, что перенос первой пробы на вторую равнялся 3 единицам, от второй к третьей пробе он возростал на 2 единицы, 81а от третьей к четвертой—только на 1. При таком неоднородном переносе каждой пробе отвечали бы следующие значения: первая проба целостного метода— никакой помощи, первая частичного метода — 3 единицы, вторая частичного метода — 5 единиц, вторая целостного метода — 6 единиц. Теперь целостный метод по-прежнему получает в итоге 6 единиц, в то время как частичный — 8.

При использовании позиционно уравненной последовательности АББА условия независимой переменной (А или Б) оказываются связанными с ранним-или-поздним переносом. Условие А связано с поздним переносом, поскольку оно получает «помощь» только на четвертой пробе, а условие Б—с ранним, на второй и третьей пробах. Внутренняя валидность эксперимента пострадает в той мере, в какой предположение об

однородности переноса окажется неверным. Приведенное объяснение справедливо не только в случае положительного, но и отрицательного эффекта последовательности, например по причине усталости испытуемого. Только в этом случае преимущество получает условие А.

При использовании схем случайной последовательности и регулярного чередования, когда число проб достаточно велико, проблема неоднородных влияний далеко не так существенна, как при позиционно уравненной последовательности. Ведь каждое из условий независимой переменной проходит в эксперименте по нескольку раз, как в ранних, так и в поздних пробах. Гораздо большую опасность для достижения внутренней валидности при использовании любой из трех рассмотренных схем представляют асимметричные влияния. Обратимся к их описанию.

Симметричные и асимметричные эффекты

Если мы уверены в том, что в нашем эксперименте с одним испытуемым эффекты последовательности являются симметричными, то связанные с ними трудности можно преодолеть. Посмотрим, почему это так и что означает в данном контексте слово «симметричные». Оно означает, что влияние условия А на последующее условие Б является точно таким же, как и влияние условия Б на последующее условие А. Скажем, использование частичного метода занятий в эксперименте Джека Моцарта точно так же влияет на последующее применение целостного метода, как использование целостного метода на последующее применение частичного.

Предположим, что между двумя методами существует своего рода антагонизм, т. е. негативный перенос, равный 5 единицам. При последовательности методов: целостный—частичный—частичный—целостный — его эффект обнаружится на второй и на четвертой пробах, т. е. на второй (частичный метод) и четвертой (целостный) пробах эффективность заучивания пьес будет снижаться на 5 единиц. Таким образом, при позиционно уравненной последовательности, которой воспользовался Джек, эти симметричные влияния взаимно компенсируются. При большем количестве проб (в случайной или чередующейся схемах) условие А предшествует условию Б, а Б предшествует А примерно одинаковое число раз, и поэтому их взаимовлияния вновь будут уравнены.

Но если перенос условия А на условие Б отличается от влияния Б на А, то экспериментатор оказывается в весьма затруднительном положении. Предположим — и это самый худший случай,—что практика, получаемая при использовании целостного метода, облегчает Джеку разучивание пьес с помощью частичного метода, а практика, получаемая при частичном методе, мешает заучиванию с помощью целостного. Пусть, как и раньше, эти влияния равны 5 единицам. При последовательности АББА качество исполнения пьес повысится на 5 единиц на второй пробе (частичный метод) и понизится на 5 единиц на четвертой пробе (целостный метод). Ясно, что эффекты переноса скомпенсированы не будут, и частичный метод получит преимущество. Для того чтобы это произошло, вовсе не обязательно разнонаправленные влияния, им достаточно быть просто неодинаковыми по величине. В данном случае мы имеем дело с систематическим смещением независимой переменной (метод заучивания) с другой переменной — последовательностью проб: либо условие А—условие Б, либо условие Б—условие А. Одна из проб условия А сопровождается влиянием Б на А, а одна из проб условия Б — влиянием А на Б. И беда в том, что экспериментатор не знает, какой вид влияния имеет Место. Все, что у него есть, — это четыре показателя качества исполнения пьес, на каждое из которых воздействуют к тому же факторы времени, а иногда (как в данном эксперименте) —еще и факторы задачи,.

Не слишком изменится эта ситуация и при регулярном чередовании проб. Каждая проба условия Б следует за пробой условия А и наоборот. Если влияния асимметричны, то систематическое смещение независимой переменной будет не в половине проб, как в схеме позиционного уравнивания, а в каждой пробе (кроме первой). И вновь нет практически никаких средств для определения асимметричности этого переноса.

При использовании случайной последовательности примерно половина проб одного условия предшествует пробам другого условия. Возникает хоть какая-то возможность определить само наличие влияний последовательности и их асимметричный характер. Например, для каждого из следующих сочетаний проб: условие А предшествует условию Б, А не предшествует Б, Б предшествует А, Б не предшествует А — можно получить отдельное значение зависимой переменной. Различие между первыми двумя значениями позволит обнаружить величину влияния условия А на условие Б, а различие между двумя вторыми значениями — величину влияния Б на А. Зная эти величины, можно позаботиться об устранении систематического смещения: определяя значения зависимой переменной при каждом из условий, нужно вычитать соответствующие величины эффектов последовательности.

Следует заметить, что наше обсуждение не дает полного представления о последствиях предъявления обоих условий независимой переменной одному и тому же испытуемому. Здесь возможны влияния более общего характера. Например, целостный метод заучивания может становиться менее эффективным только в сочетании с частичным. По контрасту он может показаться испытуемому слишком утомительным. А если бы применялся один метод, контраста бы не было. Для определения подобных влияний также нет никаких практических средств. Кроме того, в эксперименте с одним испытуемым ни одна из указанных схем не устраняет возможности асимметричного переноса. Именно эти влияния следует признать самым серьезным источником систематического смещения независимой переменной.

Более того, он является и наиболее общим видом систематического смещения. Если между двумя условиями независимой переменной существуют асимметричные эффекты последовательности, то они скажутся в любом эксперименте, сравнивающем эти условия. Влияние предубеждений экспериментатора будет сказываться только в данном конкретном эксперименте, в другом эксперименте оно может радикально измениться, если новый экспериментатор имеет противоположные предубеждения. Точно так же смещение с факторами времени и факторами задачи при использовании короткой позиционно уравненной последовательности будет существенно меняться от эксперимента к эксперименту, как и неоднородные эффекты последовательности.

КРАТКОЕ ИЗЛОЖЕНИЕ

Мало иметь заранее подготовленный план эксперимента и хорошо регистрировать его ход. Не всякий план является удачным. Реальный эксперимент можно оценить по его близости к эксперименту *безупречному*. Конечно, на практике безупречность недостижима. Значение идеи безупречного эксперимента состоит в том, что она дает образец для оценки реальных экспериментов и точного определения их недостатков. В данной главе мы пользовались этим образцом при сравнении экспериментов, описанных в главе 1, с менее удачными планами проведения тех же экспериментов.

Одним из видов безупречного эксперимента является *идеальный* эксперимент, в котором одному и тому же испытуемому в одно и то же время предъявляются разные условия независимой переменной. Главное в нем — это обеспечение неизменности всех побочных

факторов. Другим видом безупречного эксперимента является *бесконечный*, т. е. постоянно продолжающийся, эксперимент. Центральное значение имеет здесь достаточно большое количество данных. Третий вид безупречного эксперимента — эксперимент *полного соответствия*, все обстоятельства проведения которого суть те же самые обстоятельства, на которые будут распространяться его выводы. В нем важно то, что дополнительные факторы нужно сохранять не просто на неизменном, но и на адекватном уровне. При сравнении первоначально описанных экспериментов с их менее удачными вариантами было обнаружено, что во всех случаях оригинальные планы были ближе к одному из видов безупречного эксперимента.

В любом реальном эксперименте не все полученные данные можно считать достоверными. Однако нужно стремиться получить результаты (основу для будущих выводов), как можно более близкие к результатам безупречного эксперимента. Чем ближе реальный эксперимент к безупречному, безошибочному и по плану, и по процедурам, тем лучше он репрезентирует, или *представляет* его. Во всех случаях сравнения оригинальных экспериментов с их менее удачно спланированными вариантами оказалось, что именно оригинальные лучше представляют безупречный эксперимент. В зависимости от того, насколько реальные эксперименты репрезентируют безупречный, они бывают более или менее валидными. Различают два вида *валидности*. Первый называется *внутренней* валидностью. Речь идет о таком планировании эксперимента, при котором можно получить то же отношение между независимой и зависимой переменными, что и в идеальном или бесконечном экспериментах, т. е. об устранении побочных влияний. Если же эксперимент по своему проекту позволяет получить те же результаты, что и эксперимент полного соответствия, то он обладает *внешней* валидностью. Эксперимент, страдающий недостатком внутренней валидности, можно назвать *неудачным*, несостоятельным, а эксперимент, которому недостает внешней валидности, — *неадекватным*. Примером последнего служит эксперимент, при проведении которого уровень значимой Дополнительной переменной не соответствует ее реальному уровню. Было установлено, наконец, что даже когда эксперимент удачно спланирован и успешно проведен, нет полной гарантии того, что полученные в нем результаты подобны результатам безупречного эксперимента. И наоборот, неудачно спланированный эксперимент может дать корректные результаты. Хотя, конечно, у хорошо спланированного эксперимента таких шансов больше.

Особое внимание в этой главе мы уделили внутренней валидности — главному требованию к любому эксперименту. Существует Целый ряд факторов, затрудняющих достижение внутренней валидности. В эксперименте с одним испытуемым это прежде всего всевозможные изменения, происходящие с течением времени. Идеальный эксперимент неосуществим, нельзя одновременно предъявить одному и тому же испытуемому различные условия независимой переменной. На практике двух идентичных проб не существует, независимо от того, предъявляются они одному испытуемому или разным. Это связано, в частности, с теми побочными факторами, уровни которых могут изменяться. Некоторые из этих факторов можно зафиксировать и проконтролировать. Однако в большинстве случаев связанная с ними нестабильность результатов от пробы к пробе неизбежна. Как продолжительные, так и кратковременные колебания такого рода являются скорее правилом, чем исключением. Некоторые из них связаны с непостоянством побочных влияний, которые можно определить, но нельзя проконтролировать. Даже сама независимая переменная не всегда остается неизменной в различных пробах. Непостоянство же зависимой переменной может быть связано также с особенностями изучаемого поведения и с его измерениями в эксперименте. Все эти факторы мы объединили под названием «*факторы времени*».

Помимо них мы описали еще три источника нарушения внутренней валидности. Некоторые эксперименты, особенно связанные с научением, требуют применения различных задач для разных условий, и тогда «*факторы задачи*» становятся одним из таких источников. Далее, в любом эксперименте, где разные условия предъявляются одному и тому же испытуемому, существуют *эффекты последовательности*, т. е. влияния ранее предъявленного условия на предъявленное позже. И еще один источник нарушения внутренней валидности — это *предубеждение экспериментатора* о преимуществе одного из условий независимой переменной.

Все эти угрозы внутренней валидности в случае их неустранения приведут к одному из двух следствий. Первое — *ненадежность* эксперимента. Она возникает в том случае, если при большом разбросе данных проведено слишком мало проб. Здесь у нас нет полной уверенности в том, что при повторении эксперимента будут получены те же самые результаты. Второе — *систематическое смешение*, когда каждое из условий независимой переменной неразрывно связано со своим уровнем одной из других переменных, и это нарушает внутреннюю валидность.

В главе были обсуждены различные схемы эксперимента с одним испытуемым и то, насколько успешно позволяют они устранить указанные недостатки. Какая бы схема ни применялась, неотъемлемой частью эксперимента является первичный контроль за побочными факторами. Это — организация и протоколирование эксперимента, стабилизация известных побочных переменных, точность экспериментальных процедур и необходимое количество проб.

Схема случайной последовательности особенно удобна в тех экспериментах, где для каждого из условий можно применить большое количество проб. Валидность зависит от числа проб, необходимого для достижения высокой надежности эксперимента. Использование этой схемы исключает возможность всех эффектов последовательности, за исключением асимметричных влияний. Впрочем, последние сохраняются и при использовании других схем.

Если в эксперименте можно использовать не так много проб, то лучше предъявлять их не в случайном порядке, а применить схему регулярного чередования. Надежность эксперимента вновь будет зависеть от соответствия количества проб разбросу показателей. Экспериментатору следует обратить особое внимание на возможные влияния событий, происходящих через равные промежутки времени. Эти влияния могут давать определенное преимущество одному из исследуемых условий и, таким образом, приводить к систематическому смешению.

Схема позиционно уравненной последовательности применяется при относительно небольшом количестве проб (или блоков проб). Надежность зависит от адекватности выбора проб или их блоков изучаемому поведению. Здесь контроль за систематическими влияниями факторов, изменяющимися от пробы к пробе, связан с предположением о линейном характере этих изменений.

Схемы различаются также и по тому, насколько успешно они позволяют преодолеть различия задач. При использовании позиционно уравненной последовательности все задачи разделяют на пары, стараясь подобрать в каждую пару задачи примерно одинаковой трудности. Если это не удастся полностью, то внутренняя валидность эксперимента пострадает из-за неизбежного для данной схемы систематического смешения независимой переменной с факторами задачи. Если используются случайная последовательность или регулярное чередование условий, т. е. число проб достаточно

велико, то задачи (или пары задач) можно предъявлять в случайном порядке. Тогда систематического смещения с факторами задачи не происходит. Однако высокая степень изменчивости этих факторов снижает надежность эксперимента.

К систематическому смещению приводят также эффекты последовательности проб. При использовании схемы позиционного уравнивания внутренней валидности угрожает неоднородность влияний ранних и поздних экспериментальных проб. Труднее всего устранить такие случаи систематического смещения, когда влияния последовательности взаимно асимметричны, т. е. влияние условия А на условие В отличается от влияния условия В на условие А. И это может произойти в любом эксперименте, где различные условия независимой переменной предъявляются одному и тому же испытуемому.

ВОПРОСЫ

1. Как еще можно оценить эксперимент, кроме оценки качества организации его хода и протоколирования?
2. Покажите, как с помощью понятия идеального эксперимента можно определить, что один способ проведения эксперимента лучше другого.
3. Почему для опытов Йоко с томатным соком образцом безупречного эксперимента служит бесконечный, а не идеальный эксперимент?
4. Как внутренняя валидность связана с безупречным экспериментом?
5. Можно ли считать хорошим эксперимент, который не вполне репрезентирует эксперимент полного соответствия?
6. Опишите основные факторы, затрудняющие достижение внутренней валидности эксперимента.
7. Почему вопрос о различии задач касался прежде всего эксперимента Джека с заучиванием фортепьянных пьес, а не двух других экспериментов?
8. Покажите различие между ненадежностью и систематическим смещением.
9. Как вы определите, что в эксперименте лучше применить схему регулярного чередования, чем схему случайной последовательности?
10. Сравните возможность систематического смещения при использовании схемы позиционного уравнивания и двух других схем.
11. В чем различие между систематическим смещением, которое может меняться от эксперимента к эксперименту при исследовании какой-то проблемы, и систематическим смещением, которое может произойти во всех экспериментах, направленных на решение данной проблемы?