

**МИНИСТЕРСТВО ВЫСШЕГО И СРЕДНЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ УЗБЕКИСТАН**

Государственный Педагогический Университет им. Низами

**Кафедра психологии факультета “Педагогика и психологии” и подготовки
практических психологов**

НАЗЫРОВА Л.В.

Тексты лекций по предмету “ЭКСПЕРИМЕНТАЛЬНОЙ ПСИХОЛОГИИ”

(краткое содержание)

ТАШКЕНТ 2001 г.

Тема: Эксперименты, которые дублируют реальный мир

Лекция №1. 2 часа.

План:

1. Эксперименты и их отличия.
2. Простые эксперименты.
3. Условия проведения эксперимента.
4. Вычисление среднеарифметического.

Литература: Национальная программа по подготовке кадров.

Шванцара Й. Психодиагностика психического развития. П., 1978

С. Стивенс Экспериментальная психология. М. Мир, 1996

Годфруа Ж. Что такое психология?. М. Мир 1992, т. 2.

Артемьева А. Вероятностные методы в психологии. М. 1986.

Немов Р. С. Психодиагностика с элементами математической статистики. М. 1997.

Опорные слова: развитие, эксперимент, гипотеза, протокол, план.

Мы обсудили три эксперимента: использование наушников на ткацком производстве, сравнение двух методов заучивания фортепьянных пьес и оценка вкуса двух сортов томатного сока. Все это были эксперименты, дублирующие реальный мир. Данный тип эксперимента можно считать моделью общей стратегии, которая помогает вам делать выводы на базе ограниченной практики в нашей повседневной жизни. Как мы обобщаем свой жизненный опыт, так мы обобщаем и результаты эксперимента.

Мы начали с описания этого типа эксперимента, потому что он самый простой. Трудности, которые встречаются в более сложных экспериментах, например вопрос о возможности более широкого распространения экспериментальных выводов и необходимость понимания психологических теорий, нами не обсуждались. Подобные эксперименты читатель сразу же может проводить.

На примере эксперимента в ткацком цехе было показано, что неотъемлемым условием проведения эксперимента является его

организация. Без заранее подготовленного плана мы получим лишь отдельные корреляции между переменными. Но если мы организуем определенный порядок применения различных состояний независимой переменной, то сможем обнаружить ее возможные влияния на зависимую переменную. Полученное в эксперименте отношение между независимой и зависимой переменными служит для проверки экспериментальной гипотезы.

Следующее необходимое условие эксперимента — протоколирование. Оно значительно повышает качество полученных данных по сравнению с экспериментом, описанным плохо. Это показано на примере эксперимента Азиза Моцарта с заучиванием фортепьянных пьес. Если же экспериментатор не ведет протокола, он вынужден полагаться на свою память, которая никогда не бывает безупречной, а иногда может стать источником нежелательных предубеждений. Шаг за шагом мы описали все этапы работы Азиза по составлению документации эксперимента, а затем суммировали ее отдельные части.

Заключительной стадией описания Эксперимента является экспериментальный отчет. Его иллюстрация дана на примере эксперимента с томатным соком. Без отчета (который после публикации называется статьей) экспериментальные результаты остаются ничем не подтвержденными слухами. Описана типичная структура экспериментального отчета. Его основная цель — наиболее адекватное изложение эксперимента для читателя.

Вычисление среднеарифметического необходимо для количественного анализа полученного материала. Среднеарифметическое высчитывается путем сложения всех показателей и деления на количество этих показателей.

Например: $X_1+x_2+X_3/3$

ВОПРОСЫ.

1. Зачем проводятся эксперименты, дублирующие реальный мир?
2. Почему эти эксперименты обсуждаются в книге первыми?
3. Как осуществляется организация эксперимента?
4. Как протоколировать эксперимент?
5. Приведите примеры, раскрывающие содержание следующих терминов:

- а. независимая переменная;
 - б. зависимая переменная;
 - в. экспериментальные гипотезы.
6. Покажите на примере, как эти понятия связаны между собой.
 7. В чем состоят этапы протоколирования эксперимента?
 8. Что включается в основные разделы экспериментального отчета?

Тема: Основы планирования эксперимента

Лекция №2. 3 часа.

План:

1. План эксперимента и его регистрация.
2. Условия безупречного эксперимента.
3. Понятие зависимой и независимой переменной.
4. Факторы нарушения внутренней валидности эксперимента.
5. Составление схемы эксперимента.

Литература:

Введение в экспериментальную психологию личности. М.1980

В.М.Мельников Основы построения психологического эксперимента.
М.Просвещение, 1985

Опорные слова: переменная, данные, зависимая и независимая переменная, валидность, надежность эксперимента.

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. На этой лекции, мы пользовались этим образцом при сравнении экспериментов, описанных в теме 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 эксперименте — 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 смещение, когда каждое из условий независимой переменной неразрывно связано со своим уровнем одной из других переменных, и это нарушает внутреннюю валидность.

Нами 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 «той» 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 смещению.

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 « пробе, овязая 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 смещения 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 строб. 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 Саши 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 Азиза 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?

**ТЕМА: ОПРЕДЕЛЕНИЕ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ И
ПАРАМЕТРОВ НА ОСНОВЕ ВЫБОРОЧНЫХ СТАТИСТИК;
СРЕДНЕЕ И СТАНДАРТНОЕ ОТКЛОНЕНИЕ**

Лекция №3. 4 часа.

План:

1. Определение среднего в популяции (генеральной совокупности).
2. Основные пути и стратегии определения среднего и стандартного отклонения.
3. Необходимость определения среднего и стандартного отклонения при подготовке и проведении эксперимента.

Литература:

В.М.Ямпольский Введение в экспериментальную психологию личности. М.,1985
Дж.Гласс, Дж.Стенлли Математическая статистика в педагогике и психологии. М.1980.

Годфруа Ж. Что такое психология?. М.Мир 1992, т.2.

Артемьева А. Вероятностные методы в психологии. М.1986.

Немов Р.С. Психодиагностика с элементами математической статистики. М.1997.

Опорные слова: надежность, валидность, отклонение, стандартное отклонение, нормальная кривая распределения, генеральная совокупность, популяция, выборка.

В эксперименте по определению времени реакции, описанном в приложении к лекции 1, были взяты результаты действительного эксперимента. Предполагалось, что они представляют такие данные, которые могли бы быть получены в эксперименте с полной внутренней валидностью. Так, среднее время реакции на световой сигнал по 17 пробам представляло среднее, которое можно было бы получить в эксперименте с неограниченным числом проб.

Мы используем среднее для ограниченной выборки проб, чтобы сделать вывод о достаточно большой (вплоть до неограниченной) популяции проб. Такая популяция называется генеральной совокупностью. Среднее по генеральной совокупности таких, например, данных, как ВР, обозначается Mx . Такую характеристику генеральной совокупности называют параметром. Среднее, действительно вычисленное нами для данной выборки, называется статистикой, и обозначается Mx . Является ли статистика Mx наилучшей оценкой параметра Mx , которую мы можем получить на основе нашей выборки проб? Ответ — без доказательства — да. Но прежде чем вы решите, что это всегда так, давайте перейдем к стандартному отклонению, где дело обстоит иначе.

Вычисление стандартного отклонения.

Обычно помимо среднего значения оценок мы хотим знать еще кое-что, а именно, какова несистематическая вариация оценок от пробы к пробе. Наиболее распространенный способ измерения несистематической вариации состоит в вычислении стандартного отклонения.

Для этого, вы определяете, насколько каждая оценка (т. е. X) больше или меньше среднего (Mx). Затем вы возводите в квадрат каждую разность ($X - Mx$), и складываете их. Вслед за этим вы делите эту сумму на N число проб. Наконец, вы извлекаете квадратный корень из этого среднего.

Это вычисление представлено формулой с использованием символа σ_x (малая греческая буква сигма) для обозначения стандартного отклонения:

$$\sigma_x = \sqrt{\frac{\sum (X - Mx)^2}{N}}$$

(1)

Эту формулу можно сократить, введя маленькое x для обозначения ($X - Mx$). Тогда формула выглядит так:

$$\sigma_x = \sqrt{\frac{\sum x^2}{N}}$$

(2)

Давайте выпишем данные по условию А из приложения к лекции I и одновременно произведем по ним вычисления, указываемые формулой для σ_x

Проба	X	Mx	X—M _x или x	x ²
1	223	185	+38	+1444
2	184	185	—1	+1
3	209	185	+24	+576
4	183	185	-2	+4
5	180	185	—5	+25
6	168	185	—17	+289
7	215	185	+30	+900
8	172	185	—13	+169
9	200	185	+15	+225
10	191	185	+6	+36
11	197	185	+12	+144
12	188	185	+3	+9
13	174	185	—11	+121
14	176	185	—9	+81
15	155	185	-30	+900
16	165	185	-20	+400
17	163	185	—22	+484
			Σx ²	+5808

Поскольку

$$\sigma_A = \sqrt{\frac{\sum x_A^2}{N}}$$

то

$$\sigma_A = \sqrt{\frac{5808}{17}} = \sqrt{341,7} = 18,5 \text{ мс.}$$

Оценка стандартного отклонения генеральной совокупности.

Для определения среднего генеральной совокупности, которое могло бы быть получено в бесконечном эксперименте, наилучшей оценкой фактически было среднее по выборке. Иначе обстоит дело со стандартным отклонением. В любом наборе реальных проб имеет место меньшее число результатов с очень высокими или очень низкими значениями, чем в генеральной совокупности. А поскольку стандартное отклонение является мерой разброса оценок, то его величина, определенная на основе выборки, всегда меньше параметра генеральной совокупности σ_x .

Более точная оценка стандартного отклонения для генеральной совокупности находится по формуле

$$S_x = \sqrt{\frac{\sum x^2}{N-1}}$$

(3)

или

$$\sigma_x = \sqrt{\frac{N}{N-1}}$$

(4)

Для наших числовых данных:

$$SA = \sqrt{\frac{5808}{16}} = \sqrt{363,0} = 19,1 \text{ мс}$$

В некоторых экспериментах высказывается гипотеза, что поведение в одном условии более вариативно, чем в другом. Тогда целесообразнее сравнивать стандартные отклонения, а не средние. Если для обоих условий N одно и то же, можно сравнивать между собой сигмы. Однако когда N различны, сигма для условия с меньшим N дает более заниженную оценку такого параметра генеральной совокупности, как стандартное отклонение. Поэтому следует сравнивать два S .

Таблица, которая приводится ниже, поможет вам запомнить эти положения и формулы.

Параметрические характеристики генеральной совокупности (г.с.) $\bar{M}_x =$

$$\frac{\sum X_{г.с.}}{N_{г.с.}} \quad \sigma_x = \sqrt{\frac{\sum X^2_{г.с.}}{N_{г.с.}}}$$

(5)

Статистические характеристики выборки $M_x =$

$$\frac{\sum x}{N} \quad \sigma_x = \sqrt{\frac{\sum x^2}{N}}$$

(6)

Оцениваемый параметр генеральной совокупности:

$$M_x = \frac{\sum X}{N} \quad S_x = \sigma_x \sqrt{\frac{N}{N-1}}$$

(7)

или

$$(8) \quad S_x = \sqrt{\frac{\sum x^2}{N-1}}$$

Задача : Вычислите σ_x и S_x для условий Б.

Ответ: $\sigma_B = 15,9$;

$\sigma_B = 16,4$;

ПРИЛОЖЕНИЕ:

Статистическая таблица 1

Числа от 1 до 100 с их квадратами
и корнями квадратными

1	1	1,000	19	3 61	4,359
2	4	1,414	20	4 00	4,472
3	9	1,732	21	4 41	4,583
4	16	2,000	22	4 84	4,690
5	25	2,236	23	5 29	4,796
6	36	2,449	24	5 76	4,899
7	49	2,646	25	6 25	5,000
8	64	2,828	26	6 76	5,099
9	81	3,000	27	7 29	5,196
10	100	3,162	28	7 84	5,292
11	1 21	3,317	29	8 41	5,385
12	1 44	3,464	30	9 00	5,477
13	1 60	3,606	31	9 61	5,568
14	1 96	3,742	32	10 24	5,657
15	2 25	3,873	33	10 89	5,745
16	2 56	4,000	34	11 56	5,831
17	2 89	4,123	35	12 25	5,916
18	3 24	4,243	36	12 96	6,000
37	13 69	6,083	69	47 61	8,307
38	14 44	6,164	70	49 00	8,367
39	15 21	6,245	71	50 41	8,426
40	16 00	6,325	72	51 84	8,485
41	16 81	6,403	73	53 29	8,544
42	17 64	6,481	74	54 76	8,602
43	18 49	6,557	75	56 25	8,660
44	19 36	6,633	76	57 76	8,718
45	20 25	6,708	77	59 29	8,775
46	21 16	6,782	78	60 84	8,832
47	22 09	6,856	79	62 41	8,888
48	23 04	6,928	80	64 00	8,944
49	24 01	7,000	81	65 61	9,000

50	25 00	7,071	82	67 24	9,055
51	26 01	7,141	83	68 89	9,110
52	27 04	7,211	84	70 56	9,165
53	28 09	7,280	85	72 25	9,220
54	29 16	7,348	86	73 96	9,274
55	30 25	7,416	87	75 69	9,327
56	31 36	7,483	88	77 44	9,381
57	32 49	7,550	89	79 21	9,434
58	33 64	7,616	90	81 00	9,487
59	34 81	7,681	91	82 81	9,539
60	36 00	7,746	92	84 64	9,592
61	37 21	7,810	93	86 49	9,644
62	38 44	7,874	94	88 36	9,695
63	39 69	77,93	95	90 25	9,747
64	40 96	8,000	96	32 16	9,798
65	42 25	8,062	97	94 09	9,849
66	43 56	8,124	98	96 04	9,899
67	44 89	8,185	99	98 01	9,950
68	46 24	8,246	100	1 00 00	10,000

ВОПРОСЫ:

1. Как определяется среднее отклонение?
2. Какие формулы Вы использовали для определения среднего стандартного отклонения?
3. Какие сведения можно получить при помощи определения среднего стандартного отклонения?

Тема: Эксперименты, которые «улучшают» реальный мир

Лекция №4. 2 часа.

План:

1. Разнообразие в построении эксперимента.
2. Искусственные эксперименты.
3. Проверка гипотезы.

4. Возможности искусственно созданного эксперимента.

Литература:

В.М.Ямпольский Введение в экспериментальную психологию личности. М.,1985

Дж.Гласс, Дж.Стенли Математическая статистика в педагогике и психологии. М.1980.

Годфруа Ж. Что такое психология?. М.Мир 1992, т.2.

Опорные слова: искусственный эксперимент, гипотеза, задачи, валидность, экспериментальной гипотезы: независимая, зависимая и дополнительные переменные.

На этой лекции мы обсуждали три эксперимента, которые не дублируют реальный мир, а «улучшают» его. В первом эксперименте проверялась гипотеза о том, что при посадках самолета на постепенно поднимающуюся территорию у пилотов возникает зрительная иллюзия, и поэтому они снижаются слишком резко. Если проводить этот эксперимент в настоящих аэропортах, неизбежно систематическое смещение независимой переменной с различными побочными факторами. Применение тренажера, имитирующего реальный мир, позволило устранить это смещение.

Гипотеза второго эксперимента состояла в том, что при поисковых операциях на море наблюдение невооруженным глазом даст более эффективные результаты, чем наблюдение с биноклем. Если бы этот эксперимент проводился во время реальных спасательных операций, то за любой практически приемлемый срок было бы собрано лишь небольшое количество данных. Надежность такого эксперимента была бы низкой. Благодаря использованию самодельных муляжей удалось за несколько недель провести эксперимент с достаточным числом проб, повысив тем самым его надежность.

В третьем эксперименте сравнивалось качество работы с двумя высотомерами. Гипотеза была следующей: использование новой унифицированной шкалы позволит сократить количество грубых ошибок при определении высоты полета по сравнению со старой шкалой (циферблат с двумя стрелками). Если бы этот эксперимент проводился в реальном полете, то на успешность работы испытуемого с высотомером влияло бы множество побочных факторов. Полученные данные имели бы большой разброс, и это вновь, понизило бы надежность. Применение задачи по снятию показаний с фотографий обеих шкал в установленном темпе позволило значительно сократить несистематическую изменчивость экспериментальных данных.

Таким образом, в искусственных экспериментах можно повысить внутреннюю валидность. Описаны три способа улучшения реального мира, позволяющие это сделать возможным. Первый из них — устранение систематического смещения. Второй — возможность получить необходимое количество данных за более короткий срок и тем самым повысить надежность эксперимента. И третий — сократить несистематическую изменчивость данных и, следовательно, их разброс, что также обеспечивает более высокую надежность.

Но именно потому, что искусственные эксперименты не дублируют реальный мир, возникает вопрос об их внешней валидности. Достаточно ли успешно представлена в этих экспериментах реальность, чтобы можно было считать полученные результаты адекватными? Поскольку искусственные эксперименты ставятся в тех случаях, когда эксперименты с простым дублированием реального мира страдают недостатком внутренней валидности, сравнение с последними не может служить критерием их адекватности. Внешняя валидность трех описанных экспериментов оценивалась путем сравнения с другими (но также искусственными) приемами их проведения. Каждая из составных частей экспериментальной гипотезы: независимая, зависимая и дополнительные переменные — была проверена нами на соответствие исследуемой реальности.

Достичь в эксперименте соответствия независимой переменной, как правило, довольно просто. Нужно только, чтобы вводимые условия были либо типичными для реальных ситуаций, либо вполне вероятными.

Соответствие зависимой переменной оценивалось по следующим трем пунктам.

- (1) Соответствует ли работа испытуемого в эксперименте его реальной деятельности?
- (2) Отражают ли измеряемые показатели наиболее важные аспекты этой деятельности?
- (3) Адекватен ли способ представления результатов измерений?

Самой важной проблемой оказывается здесь адекватный учет ошибочных ответов испытуемого, имеющих плюсовые и минусовые значения, как, например, переоценка и недооценка высоты полета. Чаще всего применяются два способа представления таких результатов. Один из них позволяет показать преимущественную направленность ошибочных ответов, а другой — определить величину их разброса.

В искусственных экспериментах возникают также вопросы о соответствии дополнительных переменных, стабильных по своему уровню. В целом ряде случаев такие переменные являются ключевыми, и их уровень должен соответствовать реальному миру.

Нужно стараться также воспроизводить в эксперименте те дополнительные (по отношению к основной задаче) действия которые в реальности выполняются одновременно с ней. Кроме того, следует выяснить, как скажется на внешней валидности искусственного эксперимента отсутствие эмоциональной напряженности (которая, как правило, характерна для соответствующих реальных ситуаций). И наконец, следует специально проанализировать последствия предъявления испытуемому всех экспериментальных проб за короткий (сжатый по сравнению с экспериментом, дублирующим реальность) период времени.

Но есть еще один вопрос - какова цена реализма? Каждый раз необходимо проверять, не слишком ли мы усердствуем добиваясь в экспериментах как можно более точного воспроизведения реальности и не страдает ли от этого их внутренняя валидность. Такая проверка также проведена путем сравнения с другими возможными приемами эксперимента.

В заключение был затронут вопрос о финансовой стоимости искусственного эксперимента с максимальной имитацией реального мира. Иногда эта стоимость так высока, что реальным может оказаться в действительности менее реалистичный эксперимент.

ВОПРОСЫ

1. Почему эксперимент с ночными посадками самолета не мог быть проведен в настоящих аэропортах?
2. в чем состоит основное преимущество эксперимента со спасательным поиском?
3. Приведите пример эксперимента, дублирующего реальный мир, в котором имела бы место чрезмерная несистематическая изменчивость получаемых данных
4. Перечислите, каким образом в экспериментах «улучшающих» реальный мир, добиваются большей внутренней валидности, чем в экспериментах, дублирующих реальность.

5. Как вопрос о внешней валидности связан с вопросом о видах безупречного эксперимента, описанных в главе 2.
6. Что имеют в виду, говоря, что решение вопроса о соответствии эксперимента — это проверка соответствия основных составляющих экспериментальной гипотезы?
7. Приведите конкретные примеры соответствующих и несоответствующих способов получения данных для оценки значения зависимой переменной.
8. В связи с тем, что в искусственных экспериментах редко воспроизводятся стрессовые условия реальной жизни, можно ли сделать вывод, что они не могут быть соответствующими?
9. Почему проблема «реалистичности» эксперимента обсуждалась нами на примерах столь различных в этом отношении исследований, как спасательный поиск и сравнение высотомеров?

Тема: Эксперименты на представительных выборках.

Лекция №5. 3,5 часа.

План:

1. Групповые эксперименты.
2. Преимущество группового эксперимента перед индивидуальным.
3. Подготовка и особенности проведения группового эксперимента.
4. Слабые и сильные стороны эксперимента на представительных выборках.

В.М.Ямпольский Введение в экспериментальную психологию личности. М.,1985
Дж.Гласс, Дж.Стенлли Математическая статистика в педагогике и психологии. М.1980.

Годфруа Ж. Что такое психология?. М.Мир 1992, т.2.

Артемьева А. Вероятностные методы в психологии. М.1986.

Немов Р.С. Психодиагностика с элементами математической статистики. М.1997.

Опорные слова: надежность, валидность, нормальная кривая распределения, генеральная совокупность, популяция, выборка.

На этом занятии, мы обсудим три новых эксперимента, для каждого из которых, в отличие от ранее описанных, привлекается большое число испытуемых.

Первый эксперимент посвящен способам информирования о ценах на продукты. Экспериментальная гипотеза состояла в том, что прямое сообщение о цене за единицу веса (путем маркировки пакетов) позволит покупателям выбрать более дешевую покупку точнее и быстрее, чем два другие способа:

- (1) сообщение о весе пакета и его общей цене (текущая практика)
- (2) И то же плюс использование специального вычислительного устройства.

Гипотеза подтвердилась.

Первое преимущество эксперимента данного типа по сравнению с индивидуальным экспериментом заключается в том, что его результаты, полученные на выборке испытуемых, распространяются на более обширную популяцию, а не только на поведение отдельного человека.

Во втором эксперименте проверяли гипотезу о том, поможет ли специальная мысленная тренировка улучшить навыки пилотов по выполнению посадочных операций. Было обнаружено, что курсанты, прошедшие такую тренировку, выполняли эти операции лучше тех, кто тренировался лишь по обычной программе.

Второе преимущество экспериментов нового типа также связано с наличием большого числа испытуемых. Становится возможным использование схемы эксперимента с межгрупповым сравнением. Эксперимент с мысленной тренировкой (в том виде, в каком он был представлен) требовал применения такой схемы.

Схема группы индивидуальных экспериментов, когда каждый испытуемый участвует во всех экспериментальных условиях, здесь просто неосуществима. Один и тот же человек не может овладеть одним и тем же навыком двумя различными путями. Но даже там, где в принципе можно провести ряд индивидуальных экспериментов, как в эксперименте с сообщением о ценах, преимущества остаются за схемой межгруппового сравнения.

Здесь устраняются такие источники нарушения внутренней валидности эксперимента, как эффекты последовательности и факторы задачи.

Детально описан третий эксперимент — с двумя методами обучения иностранному языку. Для его проведения было очень важно подобрать две группы

испытуемых, сходных по способности изучать иностранный язык (английский). И тогда стало ясно, что эксперимент с межгрупповым сравнением, устраняя некоторые источники нарушения внутренней валидности, порождает еще один. Это — индивидуальные различия испытуемых. Описаны два вида таких различий.

Во-первых, это такие очевидные характеристики, как пол, возраст испытуемых, их образование. Сюда же относятся различия, выявляемые с помощью тестов, скажем, на знание слов испанского языка. Кроме того, существуют несистематические различия испытуемых, которые - включают как изменчивость поведения каждого человека, так и отличия в выполнении задания разными людьми.

Во всех трёх экспериментах с межгрупповым сравнением использовались схемы, устраняющие систематическое смещение независимой переменной с индивидуальными различиями испытуемых. Но если группы, работающие в разных экспериментальных условиях, имеют явные различия, скажем, взяты из разных школ, то такое смещение неизбежно. Надежность эксперимента повышается за счет привлечения большого числа испытуемых и сокращения несистематических вариации.

Последнее обеспечивается тем, что условия работы каждого испытуемого в группе одинаковы, а также путем сокращения разброса данных каждого испытуемого, как описано в лекции 2.

Существуют три стратегии составления групп испытуемых, позволяющие уравнивать эти группы в отношении изучаемого вида деятельности:

1. случайная стратегия;
2. подбор пар;
3. случайная с выделением слоев.

Их иллюстрациями служат три схемы экспериментов с межгрупповым сравнением, в которых все имеющиеся в наличии испытуемые распределяются по группам для каждого из экспериментальных условий.

При использовании схемы случайного распределения групп каждый испытуемый может с равной вероятностью попасть в одну из них.

При попарном подборе групп испытуемые сначала располагаются по степени выраженности одной из характеристик, связанной с изучаемым видом деятельности. Испытуемые с одинаковой степенью выраженности этой характеристики попадают затем в разные группы.

Применение схемы случайного распределения слоев начинают с классификации испытуемых по некоторому значимому признаку, также связанному с деятельностью, которая исследуется в эксперименте. А затем внутри каждого класса, или слоя, проводится случайное распределение по экспериментальным группам.

Все три метода устраняют систематическое смещение независимой переменной с индивидуальными различиями испытуемых. Надежность повышается за счет увеличения числа испытуемых. Если основа подбора пар или выделения слоев действительно связана с исследуемой деятельностью, то при одном и том же количестве испытуемых соответствующие стратегии позволяют обеспечить более высокую надежность по сравнению с обычным случайным распределением.

Чтобы результаты эксперимента можно было распространить на интересующую нас популяцию, выборка испытуемых должна быть репрезентативной.

Эксперимент полного соответствия (неосуществимый на практике) потребовал бы участия в нем всей популяции. Внешнюю валидность эксперимента можно оценить потому, насколько близок к этой недостижимой цели метод отбора испытуемых. Хорошо представить популяцию в целом могла бы случайная выборка достаточно большого размера. Однако подходящие условия для такого выбора встречаются редко: часто исследуемая популяция целиком даже не существует в данный момент времени.

Самым удачным из описанных исследований был в этом смысле эксперимент с мысленной тренировкой. Здесь можно было применить стратегию случайного выбора к популяции курсантов. Поскольку эксперимент проводился по схеме межгруппового сравнения с исследованием двух условий независимой переменной, использовались две случайные выборки. Здесь случайная стратегия применялась для отбора испытуемых, а не для распределения по группам, это имеет место, когда все имеющиеся в наличии люди должны участвовать в эксперименте. Эта схема называется схемой случайно отобранных групп.

Хотя для эксперимента с мысленной тренировкой было отобрано не очень много испытуемых, выборка успешно представляла изучаемую субпопуляцию, так как индивидуальные различия между курсантами были не слишком велики. Если же исследователь хочет распространить результаты эксперимента за пределы данной популяции, скажем, на курсантов следующих лет, то проблемы внешней

валидности станут более серьезными. Возникнет вопрос о соответствии «уровней» субъектных характеристик изучаемой популяции и той популяции, к которой прилагаются полученные результаты.

Связь внешней и внутренней валидности была описана для экспериментов, в которых испытуемые отбираются из популяции. Если способ отбора испытуемых обеспечивает высокую внешнюю валидность (в отношении индивидуальных различий), то тем самым достигается достаточное подобие групп, т. е. повышается внутренняя валидность. Напротив, группы, хорошо уравненные между собой по индивидуальным характеристикам, могут не быть достаточно представительными для исследуемой популяции. Иначе говоря, высокая внутренняя валидность эксперимента не гарантирует его высокой внешней валидности. При использовании схемы индивидуального эксперимента связь между внешней и внутренней валидностью отсутствует, поскольку та и другая зависят от разных факторов.

Вместе со случайным отбором, так же как и со случайным распределением, можно использовать выделение слоев. Это дает пятую эффективную схему эксперимента с межгрупповым сравнением — послойный случайный отбор.

Сначала популяцию разделяют на классы, или слои, а затем внутри каждого слоя применяют случайный отбор. Выделение слоев дает то же преимущество, что и при случайном распределении, если его основание связано с изучаемым видом деятельности.

Когда для эксперимента привлекаются испытуемые, имеющиеся в наличии, будь то «заложники» или добровольцы, есть опасность, что выборка будет недостаточно репрезентативной. Для оценки репрезентативности выборки экспериментатор должен подробно проанализировать ситуацию, а также предвидеть последствия случаев, когда выборка недостаточно представляет популяцию.

ВОПРОСЫ

1. Чем отличаются возможности обобщения результатов экспериментов, описанных в этой главе, от экспериментов, рассмотренных ранее?
2. В каких случаях нужно проводить эксперимент с межгрупповым сравнением, а не интраиндивидуальный эксперимент?

3. Какие источники нарушения внутренней валидности устраняются в эксперименте с межгрупповым сравнением?
4. Каковы основные источники межиндивидуальных различий?
5. Почему неправильно проводить эксперимент на уже существующих группах испытуемых, например на ученических группах в разных школах?
6. Как можно повысить надежность в межгрупповом эксперименте?
7. В чем отличие двух межгрупповых схем: со случайным распределением по группам и со случайным отбором в группы?
8. Какова связь внешней и внутренней валидности в экспериментах с межгрупповыми схемами?
9. Приведите пример эксперимента, в котором хорошо представлена одна популяция и плохо—другая.

Тема: Выделение независимой переменной.

Лекция №6. 3 часа

План:

1. Понятие независимой переменной.
2. Определение независимой переменной и соответствующих ей показателей.
3. Раскрытие экспериментов по изучению механизмов поведения. Условия проведения эксперимента.
4. Выделение экспериментальной и контрольной группы.

Литература:

Годфруа Ж. Что такое психология?. М.Мир 1992, т.2.

Артемьева А. Вероятностные методы в психологии. М.1986.

Немов Р.С. Психодиагностика с элементами математической статистики. М.1997.

Опорные слова: надежность, валидность, поведение, выборка, контрольная и экспериментальная группа, естественное смешение.

Для научного понимания механизмов поведения необходимы эксперименты, в которых выделяется истинная независимая переменная. Проверяемая в них экспериментальная гипотеза касается уже единичной независимой переменной, в отличие от «группы факторов», которая характерна для экспериментов с немедленным

применением результатов. В этих новых исследованиях, к экспериментальной гипотезе предъявляются более высокие требования, чем это было раньше, поскольку описываемая независимая переменная должна быть «очищена» от других переменных. Эталоном для оценки внутренней валидности такого эксперимента, в котором делаются попытки выделить единичную независимую переменную, является чистый эксперимент. Этот эксперимент представляет собой прежде всего идеальный эксперимент, в котором испытуемому одновременно предъявляются два разных условия независимой переменной. Кроме того он требует, чтобы независимая переменная была единичной, а не сложной и чтобы другие переменные, которые могут влиять на поведение, сохранялись строго постоянными.

Прежде всего мы показали в эксперименте с индианками и в эксперименте по трудовой этике, что для выделения независимой переменной должны быть созданы специальные условия.

Например, если эксперимент проводится в реальной рабочей обстановке, то невозможно ограничить условия различием только по одной переменной. В некоторых экспериментах требуется большое мастерство, чтобы «очистить» независимую переменную. Например, при изучении направленного обоняния исследователь очень постарался, чтобы точно задать условия независимой переменной — разницу во времени — и ликвидировать другие переменные, т. е. щелчки реле.

Внутренняя валидность нарушается далее из-за некоторых новых типов систематического смещения. Они проистекают из того печального факта, что мы не можем делать лишь что-то одно. Во всех случаях эти типы смещения возникают от того, что активное условие независимой переменной сопровождается активным уровнем побочной переменной.

Таким образом, активный уровень независимой переменной связан с активным уровнем побочной переменной, а неактивный уровень независимой переменной связан с неактивным уровнем побочной переменной. В итоге независимая переменная смешивается с побочной переменной, и расчленив их влияние на поведение невозможно. В любом случае средством контроля такого сопутствующего смещения была схема, в которой неактивный уровень независимой переменной включал в себя активный уровень сопутствующей переменной.

Это называется контрольным условием; если же используется межгрупповая схема эксперимента, то такая комбинация представлена контрольной группой. Активный уровень независимой переменной в этом случае называется экспериментальным условием, а соответствующая группа — экспериментальной группой.

Первый тип обсуждавшегося сопутствующего смешения — смешение с артефактной переменной. Он рассматривался на примерах из психофизиологии. Если животные, у которых удалена часть мозга, сравниваются с животными, не оперированными вообще, то независимая переменная (удаление или неудаление части мозга) смешивается с артефактной переменной (повреждение других тканей). Поэтому необходимо ввести в эксперимент прооперированную контрольную группу животных для сравнения с оперированной экспериментальной группой.

Эта контрольная группа подвергается той же самой операции, исключая удаление исследуемой части мозга. Подобным же образом в эксперименте, в котором введение лекарства является экспериментальным условием, должно быть контрольное условие с введением плацебо.

Для иллюстрации других типов сопутствующего смешения использовался эксперимент, проверяющий гипотезу, что дети плачут из-за ухода близкого им человека (матери).

К этим типам относится, во-первых, смешение с расширенной переменной, когда вообще кто-нибудь уходит.

Во-вторых, смешение с вторичной переменной — ребенок находится с кем-то или совсем один.

В-третьих, смешение с другой вторичной переменной — с кем ребенок остается. Все это были иллюстрации сопутствующего смешения, называемого естественным смешением, в отличие от искусственного смешения, описанного выше.

Термин «естественное» означает только, что связь уровней независимой переменной и уровней побочной переменной заложена в «природе вещей», а не вызывается исключительно действиями экспериментатора. Способ контроля естественного смешения тот же, что и в случае искусственного смешения: создание контрольного условия, которое включает в себя активный уровень сопутствующей переменной. Если нарисовать диаграмму уровней, по которым

различаются условия, то можно увидеть, что контрольное условие делает возможным однолинейное сравнение только по независимой переменной.

Следующей была рассмотрена проблема операциональной валидности. Как и в случае с внешней валидностью, вопрос здесь тот же: правильный ли эксперимент проведен для проверки экспериментальной гипотезы. Переменные в экспериментах с выделением независимой переменной, как правило, выражены в довольно абстрактных понятиях; экспериментальные же ситуации — гораздо конкретнее. Поскольку в экспериментах приходится делать конкретные вещи, эксперимент полного соответствия невозможен. Можно лишь постараться, чтобы конкретные экспериментальные операции репрезентировали понятия, содержащиеся в экспериментальной гипотезе. Для иллюстрации приводилось исследование влияния социальной атмосферы на поведение людей. В отношении этого исследования рассматривался вопрос, насколько точно условия независимой переменной соответствуют общим понятиям, которые они по предположению репрезентируют.

Затем был проанализирован эксперимент по научению с точки зрения репрезентативности независимой переменной.

Мы рассмотрели вопрос об испытуемых, постепенно сужая их контингент. Некоторые эксперименты касаются всех животных, так что выбор конкретного вида зависит главным образом от соображений удобства. Если испытуемыми являются лабораторные крысы, то вопрос об общности результатов становится особенно актуальным. В отдельных случаях выбирается специальный вид животных из-за их характерных особенностей.

Для большого количества экспериментов реальный интерес представляет поведение людей. Использование студентов колледжа вполне допустимо, но вопрос об общности результатов необходимо проверять в контексте каждого конкретного эксперимента.

И все же необходимо проводить отбор или распределение испытуемых. Как правило, исследователь в эксперименте с выделением независимой переменной не пытается получить такую выборку испытуемых, которая представляла бы некоторую популяцию. Вместо этого он (или она) сосредоточивается на вопросах внутренней валидности и проблему обобщения результатов решает путем обследования разных групп испытуемых в отдельных экспериментах.

В межгрупповых схемах применяются две различные стратегии. Одна из них включает предварительные решения, которые обсуждались на примере

экспериментов с репрезентативными выборками. Другая является стратегией распределения испытуемых по группам. Если число испытуемых в группе невелико, то по мере появления новых испытуемых они распределяются по группам в последовательности регулярного чередования. При большем числе испытуемых лучше случайная последовательность распределения, так как она устраняет возможность нежелательных систематических влияний.

Для улучшения результатов, основанных на сравнении групп, могут быть использованы особенно сходные индивиды. Для этого в сравниваемые условия могут, например, помещаться особи одного помета, а в случае с людьми — монозиготные близнецы.

В современных исследованиях схемы индивидуального эксперимента используются так же часто, как и межгрупповые. Небезынтересно отметить, что эта схема может даже вылиться в эксперимент с одним единственным испытуемым. В этом случае иногда возможен надежный контроль эффектов последовательности и временных эффектов.

Мы дали краткую характеристику направления нашего движения в предыдущих и последующих лекциях.

Было отмечено, что мы начали с весьма конкретных исследований, позволяющих производить минимальные обобщения. Затем мы перешли к более абстрактным переменным и к более широким обобщениям. В последующих лекциях будут описаны эксперименты, которые выявляют более полный, менее ограниченный набор факторов, влияющих на поведение.

ВОПРОСЫ

1. Что мы подразумеваем под контрольной группой в эксперименте, требующем удаления части мозга?
2. В чем разница между процедурным смешением, которое возникает из-за того, что эксперимент не может быть идеальным, и сопутствующим смешением как результатом того, что эксперимент не может быть чистым?
3. Приведите пример естественного сопутствующего смешения и покажите, каким образом оно может быть устранено. Составьте таблицу для пояснения ситуации.
4. Как сформулировать гипотезу, которую можно проверить экспериментальным путем?
5. Как используется идея эксперимента полного соответствия в экспериментах на выделение независимой переменной?

6. Как экспериментатор решает вопрос об использовании того или иного вида животных в эксперименте?
7. Что подразумевается под сериальными стратегиями распределения испытуемых по группам?
8. Всегда ли эксперименты на выделение независимой переменной проводятся на репрезентативных выборках? Хорошая ли это идея?

Тема: Значимые результаты

Лекция №7. 2,5 часа.

План:

1. Проверка экспериментальной гипотезы.
2. Статистическая проверка нуль-гипотезы.
3. Действие независимой переменной.

Литература:

Годфруа Ж. Что такое психология?. М.Мир 1992, т.2.

Артемьева А. Вероятностные методы в психологии. М.1986.

Немов Р.С. Психодиагностика с элементами математической статистики. М.1997.

Опорные слова: интерпритация, популяция, выборка, нуль-гипотеза, статистическая значимость, альфа- уровень, бета- уровень, надежность.

Большие различия в действии разных условий независимой переменной приводят экспериментатора к выводу о подтверждении экспериментальной гипотезы.

Меньшие различия интерпретируются как случайный результат. Основанием для таких различных выводов является статистическая значимость. Более конкретно это означает, что если бы в идеальном или бесконечном эксперименте различие отсутствовало, то было бы мало вероятно получить в конкретном эксперименте большое различие, не так невероятно — меньшее различие.

В научных экспериментах — в отличие от тех, где существует только два практических исхода — возможны три заключения из экспериментальных данных. В дополнение к подтверждению экспериментальной или противоположной ей гипотезы возможно заключение о не подтверждении ни одной из них. Какое

именно -из этих трех заключений будет сделано, зависит от статистического решения относительно нуль-гипотезы.

Если бы был проведен бесконечный эксперимент и нуль-гипотеза оказалась верной, то среднее различие между условиями было бы равно нулю. Однако в отдельных конкретных экспериментах различия могут быть как в пользу одного условия, так и в пользу другого. Если различие настолько велико, что очень редко могло бы быть получено в бесконечном эксперименте, нуль-гипотеза отвергается. Однако если вероятность появления различия, подобного полученному, достаточно высока, нуль-гипотеза не отвергается. Когда нуль - гипотеза отвергается, делается вывод о подтверждении экспериментальной гипотезы (или противоположной ей гипотезы, если различие оказалось с обратным знаком). Когда нуль-гипотеза не отвергается, ни экспериментальная, ни противоположная ей гипотезы не подтверждаются. Это последнее заключение может означать одно из двух. Если данные ненадежны, заключение будет состоять в том, что действие независимой переменной просто не удалось выявить. При надежных данных экспериментатор может быть уверен, что условия не оказывают различного действия.

Величина различия между условиями, необходимая для отвержения нуль-гипотезы, определяется двумя факторами.

Первое — это надежность. Чем больше надежность, тем меньше различие, допускающее отвержение.

Второй фактор — вероятность того, что экспериментатор рискнет ошибочно отвергнуть нуль-гипотезу, когда она верна. Он называется альфа-уровнем правила его решения. Ошибка, которая будет увеличиваться с возрастанием этого риска, называется ошибкой I типа. Так, риск ошибки I типа в пять раз выше при альфа-уровне 0,05 по сравнению с альфа-уровнем 0,01.

Однако при уменьшении альфа-уровня увеличивается риск противоположной ошибки. Это риск не отвергнуть нуль-гипотезу, а верна некоторая другая гипотеза (и, конечно, нуль-гипотеза ошибочна). Это называют ошибкой II типа. Для любого конкретного набора данных эта вероятность (называемая бетой) увеличивается с уменьшением альфа-уровня. Однако, увеличивая надежность эксперимента, можно найти приемлемую величину бета даже при строгом альфа-уровне. Говорят, что статистическая проверка имеет силу в той мере, в какой низка вероятность бета и в которой может быть выявлено истинное различие.

Использование строгого альфа-уровня (например, 0,01) рекомендуется в тех случаях, когда различие между условиями должно подтвердить новую гипотезу, противоречащую общепринятому мнению. Эта строгость нужна для того, чтобы не засорять науку слишком большим числом артефактов. 5 ложных утверждений из 20 - это слишком тяжелое бремя для науки. С другой стороны, если результаты показали влияние независимой переменной, его нельзя сбрасывать со счетов только потому, что различие не достигло уровня значимости 0,01.

Назначение проверок на значимость — повышение внутренней валидности. Ведь внутренняя валидность и проверка нуль-гипотезы могут быть описаны через бесконечный эксперимент. В бесконечном эксперименте, состоящем из множества отдельных экспериментов (таких, какие проводятся реально), общее среднее различие между условиями будет равно нулю, если верна нуль-гипотеза. Однако различия, обнаруживаемые в отдельных экспериментах, не будут равны нулю, а лишь только распределятся вокруг нуля. Экспериментатор может выяснить это распределение. Он соотнесет полученное различие с его вариабельностью, но не будет делать вывода о различии только на том основании, что много отдельных экспериментов дает достаточно большое различие.

Если же верна нуль-гипотеза, экспериментатор также хотел бы обосновать и этот вывод. Но даже чтобы иметь возможность сделать вывод о правильности какой-то другой гипотезы, экспериментатор вынужден идти на некоторый риск. Экспериментатор хочет иметь заключение о верности экспериментальной гипотезы с такой степенью обоснованности, как если бы ожидаемое различие было получено в бесконечном эксперименте. Положение, которое он в конце концов занимает между ошибками I и II типа, отражает его оценку относительной валидности обоих типов обоснованности.

На пути к окончательным выводам остаются три трудные проблемы.

Первая состоит в том, что только одного значимого различия недостаточно, если ожидается сильное влияние независимой переменной. Статистическая проверка наиболее пригодна в тех случаях, когда действие исследуемого фактора «зашумлено» другими случайными факторами.

Вторая проблема заключается в том, что использование слишком большого числа испытуемых обнаруживает действие определенных дополнительных факторов.

Третья проблема касается универсальности результатов. Можно ли отнести выводы ко всей соответствующей популяции, если они справедливы даже не для всех исследовавшихся испытуемых? Причем не только по причине случайных изменений.

Наконец, было показано, что мы не можем принять экспериментальные выводы только на основе постоянных и достаточно сильных выявленных различий между условиями.

Эксперименту будет не хватать внешней валидности, если он не будет удовлетворять хотя бы одному из целого ряда условий. Более того, он не будет обладать даже внутренней валидностью, если не организовать достаточный контроль за систематическим смещением.

ВОПРОСЫ

1. Почему Флинер и Керне заключили, что старшие дети больше плачут при уходе матери, чем при уходе ассистентки, а у младших детей такого различия нет?
2. Что такое нуль-гипотеза?
3. Что показывает диаграмма, иллюстрирующая: различие между средними для каждого условия, статистическое решение и заключение об экспериментальной гипотезе?
4. Как влияет уменьшение надежности на величину различия между средними, требуемую для отвержения нуль-гипотезы?
5. Как влияет альфа-уровень в правиле решения на величину различия между средними, требуемую для отвержения нуль-гипотезы?
6. Соотнесите альфа-уровень с риском ошибок I и II типов.
7. Когда особенно важно избегать ошибки I типа?
8. Опишите три фактора, влияющие на вероятность бета. Что это означает в отношении риска ошибки II типа?
9. При каких условиях экспериментатор может заключить, что независимая переменная не оказывает действия?
10. Почему говорят, что разумное использование правила статистического решения способствует внутренней валидности?
11. Может ли быть в эксперименте слишком много испытуемых?
12. Если в эксперименте получены надежные данные и высоко значимые различия между условиями, обеспечивает ли это полностью валидность вывода?

Тема: Многоуровневые эксперименты

Лекция № 8 лекция. 3,5 часа.

План:

1. Многоуровневые эксперименты.
2. Условия построения.
3. Определение внешней и внутренней валидности.
4. Положительные и отрицательные стороны эксперимента.

Литература:

В.М.Ямпольский Введение в экспериментальную психологию личности. М.,1985
Дж.Гласс, Дж.Стенлли Математическая статистика в педагогике и психологии. М.1980.

Опорные слова: многоуровневые эксперименты, внутренняя валидность, активный уровень независимой переменной., качественная переменная, эффект ряда, неоднородный перенос.

Было рассмотрено три возможных двухуровневых эксперимента, которые оказались совершенно неадекватными по сравнению с аналогичным реально проведенным многоуровневым экспериментом. На этом примере были разобраны преимущества многоуровневого эксперимента.

Во-первых, он обеспечивает большую внутреннюю валидность, чем простые эксперименты, описанные в предыдущих главах. В некоторых экспериментах, где фигурируют только два уровня, независимая переменная по существу является количественной. Использование только двух уровней такой переменной может не выявить истинного отношения, которое могло бы быть найдено в идеальном количественном эксперименте, где используется неограниченное число уровней. Чем больше уровней независимой переменной мы используем, тем больше мы приближаемся к этому невозможному эксперименту и тем больше становится внутренняя валидность.

Кроме того, лучше оказывается контроль за сопутствующим смещением. Если активный уровень независимой переменной сравнивается с ее нулевым (или неактивным) уровнем, это может косвенно ввести активный уровень вторичной переменной. В качестве примеров приводилось осознание действия лекарства и установление контакта с экспериментатором. В то же время ступенчатое изменение независимой переменной, приводящее к ступенчатому изменению зависимой переменной, делает маловероятным такое смещение.

Многоуровневые эксперименты превосходят более простые эксперименты и в другом отношении. В них могут проверяться гипотезы, которые ведут к более тонкому пониманию механизмов поведения.

Во-первых, в сравнении с экспериментами, использующими качественные независимые переменные, здесь можно лучше выделить единичную переменную. Качественная переменная — такая, как чтение — в отличие от прослушивания — может быть только нерасчлененным комплексом факторов. Возможностей для введения количественных переменных — сколько угодно, включая шкалирование поведения.

Дальнейшие преимущества определяются возможностью проверки более тонких гипотез об отношении между независимой и зависимой переменными. Часто наиболее правильной оказывается гипотеза о максимальной (или минимальной) величине зависимой переменной при некотором промежуточном уровне независимой переменной. Она может следовать из теории двух процессов, связанных противоположным образом с уровнем независимой переменной. Например, может случиться, что негативный процесс берет верх над позитивным только на очень высоких уровнях независимой переменной. Именно так был проанализирован воображаемый многоуровневый эксперимент по трудовой этике. В качестве основных здесь были предположены «стремление к активности» и чувство «неприязни» к нажиманию на рычаг. Другим примером служила теория, согласно которой в основе поведения лежат два позитивных процесса, на которые увеличение уровня независимой переменной влияет противоположным образом. Именно так был проанализирован эксперимент по запоминанию списков студентами колледжа с варьированием интервалов между элементами запоминаемого списка и эксперимент, связывающий величину удара током с перцептивным различием у танцующих мышей. Переменные, лежащие в основе поведения в этих двух экспериментах, представляли собой различие стимулов и образование ассоциаций.

Многоуровневые независимые переменные позволяют проверять более детализованные экспериментальные гипотезы. Последние создаются на основе моделей и теорий, объясняющих, каким образом ступенчатые изменения независимой переменной приводят к изменению зависимой переменной. Так, на основе сканирующей модели мнемического поиска была выдвинута гипотеза о том, что одинаковые по абсолютной величине приросты объема запоминаемого материала будут сопровождаться примерно равными абсолютными приростами

времени поиска: гипотеза «абсолютно-абсолютных» отношений. При исследовании связи между числом альтернатив и временем реакции проверялась гипотеза «относительно-абсолютного» отношения: при каждом увеличении количества альтернатив в одно и то же число раз будет наблюдаться увеличение времени реакции на одну и ту же абсолютную величину. Эта гипотеза была основана на модели наиболее эффективного способа принятия решения.

На основе теоретического представления о том, как сенсорные органы превращают физическую энергию стимула в нервное возбуждение, была предсказана гипотеза «относительно-относительного» отношения между величиной поднимаемого веса и субъективным ощущением тяжести. Во всех перечисленных случаях результаты подтверждали гипотезу: при выборе на осях соответствующих шкал получалась линейная зависимость между независимой и зависимой переменной.

В многоуровневых экспериментах могут быть использованы и ранее описанные экспериментальные схемы. Для межгрупповой схемы существует практическая трудность: она состоит в необходимости привлекать слишком большое количество испытуемых. Внутрииндивидуальный контроль наиболее пригоден в случаях, когда предъявляются в случайном порядке короткие пробы на различных уровнях в большом наборе проб. Когда же каждая проба длительна, как это обычно и бывает при использовании внутрииндивидуального уравнивания, в многоуровневом эксперименте возникает практическая трудность: необходимость затраты слишком большого времени на каждого испытуемого.

Эти практические трудности можно преодолеть путем использования кросс-индивидуального реверсивного уравнивания. Однако эта конкретная схема не обеспечивает контроль эффектов *неоднородного переноса* от предыдущей пробы к следующей. Такой контроль обеспечивает схема полного позиционного уравнивания, но она требует слишком большого числа различных последовательностей (и групп испытуемых), чтобы быть практически удобной. Типичным методом внутрииндивидуального уравнивания, который также контролирует неоднородный перенос, является латинский квадрат. В этой схеме каждый уровень независимой переменной появляется однажды в каждой позиции последовательности. Более тщательный контроль достигается путем использования только сбалансированных квадратов, в которых каждому уровню

независимой переменной только один раз предшествует каждый из остальных уровней.

И все же ни одна схема кросс-индивидуального уравнивания не обеспечивает контроль *эффектов ряда*. В любой последовательности низким уровням чаще предшествуют более высокие, чем более низкие, уровни, а высоким — низкие, и это порождает угрозу *асимметричного переноса*.

Другим описанным эффектом ряда является эффект центрации. Он возникает в связи с тем, что только! уровням, близким к середине ряда, могут в равной мере предшествовать и высокие и низкие уровни. Более благоприятное положение средних уровней было показано в эксперименте с обработкой деталей.

Был рассмотрен ряд полезных советов. Так, при кросс-индивидуальной схеме вместо реверсивного уравнивания лучше использовать латинский квадрат; для того чтобы избежать влияния утомления, необходимо давать достаточный отдых между пробами; и, наконец, для контроля за эффектами ряда необходимо использовать перекрывающиеся диапазоны уровней независимой переменной.

Независимо от экспериментальной схемы при проверке точных гипотез все-таки остаются две угрозы внутренней валидности. Одна из них состоит в том, что форма кривой, полученной на грук-пе испытуемых, может не представлять индивидуальные кривые ни одного испытуемого. В идеальном эксперименте один и тот же испытуемый должен был бы проверяться одновременно по всем уровням. Поэтому возможность неверной представленности истинного отношения в полученном отношении является источником внутренней невалидности. Особенно подвержены такой опасности межгрупповые схемы. Опасность может быть уменьшена благодаря уравниванию испытуемых и использованию однородных групп.

Другим источником неверной представленности отношения между независимой и зависимой переменными может быть искажающее действие измерительных приборов и шкал, с помощью которых измеряют зависимую и независимую переменные. С наибольшей вероятностью такое искажение возникает в двух типах измерения. Первый случай — когда необходимо усиление малых физических величин, второй — когда используется субъективное шкалирование.

На предыдущих лекциях в качестве угроз внутренней валидности описывались ненадежность и систематическое смещение.

На этой лекции показана новая угроза — неверно установленное отношение между независимой и зависимой переменными. Оно может быть следствием использования усредненных кривых, которые не представляют индивидуальные, затем — применения слишком малого числа уровней независимой переменной, наконец, проведения неверных измерений.

ВОПРОСЫ

1. Чем многоуровневый эксперимент отличается от экспериментов, описанных на предыдущих лекциях?
2. Что означает утверждение, что многоуровневые эксперименты обеспечивают контроль для проверки экспериментальных гипотез, которые могли бы быть проверены и в двухуровневом эксперименте?
3. Сравните с теоретической точки зрения результаты эксперимента с количественным изменением независимой переменной и эксперимента с условиями, отличающимися только качественно.
4. Что подразумевается под экспериментальной гипотезой максимума или минимума?
5. Почему к эксперименту Стернберга по исследованию памяти приложим термин «абсолютно-абсолютного» отношения? Что лежало в основе этой экспериментальной гипотезы?
6. Определите различие между количественными экспериментальными гипотезами Хика (1922) о времени реакции и Харпера и Стивенса (1948) о субъективной тяжести.
7. Каковы практические причины использования позиционного уравнивания по всем испытуемым, а не межгрупповой схемы или интраиндивидуального позиционного уравнивания?
8. Что такое латинский квадрат?
9. Может ли предохранить полное позиционное уравнивание от эффектов неоднородного переноса? От эффектов ряда?
10. Какие угрозы внутренней валидности остаются при использовании любых схем проверки гипотезы точного отношения между независимой и зависимой переменными?
11. Понятие идеального эксперимента было вновь введено в связи с угрозой внутренней валидности, отличающейся от ненадежности и систематического смещения. Как это было сделано? Как бы вы в таком случае определили внутреннюю валидность?

Тема: Факторные эксперименты.

Лекция № 9. 3,5 часа.

План:

1. Отличительные особенности факторного эксперимента.
2. Положительные и отрицательные стороны факторного эксперимента.
3. Взаимодействие независимых переменных в факторном эксперименте.
4. Эксперименты Йоркса-Додсона. Построение факторной схемы. Эксперименты на животных.

Литература:

Годфруа Ж. Что такое психология?. М. Мир 1992, т.2.

Артемьева А. Вероятностные методы в психологии. М. 1986.

Немов Р.С. Психодиагностика с элементами математической статистики. М. 1997.

Пиажже Ж. Экспериментальная психология. М. 1979

Хрестоматия по истории психологии. М. 1980.

Оорные слова: фактор, факторный эксперимент, взаимодействие, идеальный эксперимент, межгрупповое сравнение, интра- или кросс-индивидуальное сравнение.

Для проверки гипотезы о воздействии рассечения свода у обезьяны на ее способность узнавания нужно было провести два факторных эксперимента: один — чтобы определить, нарушается ли при этом именно память, а другой — чтобы установить специфичность ее нарушения. В каждом из этих экспериментов помимо первичной, изучаемой переменной — состояния свода — привлекалась вторая независимая переменная, факторным называется эксперимент с использованием по крайней мере двух независимых переменных.

В итоге факторного эксперимента определяются основной результат действия каждой из независимых переменных и взаимодействие между ними. На примере эксперимента с рассеченным сводом мы раскрыли конкретное содержание основных результатов и взаимодействия и показали способы их измерения (когда две независимые переменные имеют по два уровня). Основной результат действия независимой переменной — это разность между общими средними значениями зависимой переменной по каждому из уровней.

Взаимодействие — это разность между двумя разностями по каждому из уровней второй независимой переменной.

Основные виды взаимодействия, которые можно получить в факторных экспериментах, иллюстрировались графически. Здесь же представлены и способы измерения основных результатов и взаимодействий. Одна из независимых переменных помещается на горизонтальной оси, и затем по соответствующим значениям зависимой переменной строятся отдельные графики для каждого из уровней другой независимой переменной. Если отрезки, полученные таким образом для второй независимой переменной, параллельны, взаимодействие отсутствует [т.е. нулевое]. Если отрезки расходятся вправо или влево, взаимодействие называют расходящимся. А если они пересекают друг друга, то это пересекающееся взаимодействие.

Факторные эксперименты часто применяются для проверки гипотез с одним отношением, т. е. для определения результата влияния одной независимой переменной на поведение. Именно такие гипотезы обсуждались нами в предыдущих главах. Центральная задача в таких экспериментах — устранить сопутствующее смешение, как естественное, так и искусственное. Если, согласно гипотезе, независимая переменная действует на определенную базисную переменную, то с помощью факторного эксперимента можно обеспечить контроль за выделением этого воздействия. Однако при том условии, которое организуется для определения воздействий различных уровней изучаемой независимой переменной, последняя может по-разному влиять и на вторую базисную переменную. Это показано на примерах экспериментов Гаффана (1974) с рассечением свода у обезьян. Чтобы устранить смешение собственно памяти с восприятием, между пробой-образцом и пробой-выбором давались не только длинные, но и короткие интервалы отсрочки. При коротком интервале для правильного выбора не нужно запоминать предмет надолго. Смещение способности к узнаванию и запоминания по ассоциации устранялось путем использования чисто ассоциативных задач, для решения которых не нужно опознавать предъявляемые предметы. В каждом из этих экспериментов второй уровень контрольной переменной не оказывал никакого воздействия на предлагаемую базисную переменную. Показателем успешности разделения предполагаемой и сопутствующей переменных служил факт расходящегося взаимодействия между ними.

Таким образом, применение факторной схемы позволило получить результаты, близкие к результатам идеального (чистого) эксперимента, в котором

можно было бы исследовать только базисную переменную. Тем самым повышается внутренняя валидность эксперимента.

Использование факторных схем позволяет повысить и внешнюю валидность. Прежде всего в тех случаях, когда нужно проверить гипотезу о непосредственном простом действии независимой переменной. Согласно альтернативной гипотезе, воздействие является комбинированным, т. е. активный уровень независимой переменной дает результат только благодаря сочетанию с активным уровнем дополнительной переменной. Так, Вайз и Даусон (1974), проводя опыты на крысах, предполагали, что лекарственный препарат диазепам непосредственно приводит к увеличению количества съеданной пищи, повышая аппетит. Согласно же комбинированной гипотезе, лекарство действует не прямо, а снимая состояние тревожности. В опытах, проводившихся как в новой обстановке, вызывающей у крыс подобное состояние, так и в домашних клетках, было обнаружено значительное различие в количестве съеденной пищи между животными с инъекцией диазепама и с инъекцией плацебо. Тем самым удалось показать, что действие диазепама является непосредственным, оно не зависит от уровня тревожности. А если бы результат действия зависел от уровня дополнительной переменной, то обобщать его было бы нельзя. При проверке обобщения результатов следует „учитывать не только факторы, вызванные участием экспериментатора, но и индивидуальные различия испытуемых. Для того и привлекались девочки-индианки в экспериментах по изучению трудовой этики, чтобы распространить полученные результаты на людей, принадлежащих к разным культурным группам.

При использовании схем позиционного уравнивания эксперимент автоматически становится факторным. Помимо изучаемой независимой переменной здесь появляется переменная местоположения каждой пробы в порядке их предъявления. В кросс-индивидуальном эксперименте с позиционным уравниванием определяется/ среднее значение по данным всех испытуемых для первой пробы, затем для второй и т. д. Однако главное преимущество факторных экспериментов — это возможность проверять настоящие комбинированные гипотезы. Например, для проверки гипотезы, известной теперь как закон Иеркса — Додсона, оптимальная сила напряжения, способствующего хорошему изучению, должна понижаться с увеличением трудности задач — нужно было использовать задачи различных уровней трудности различения. Далее обсуждаются факторные эксперименты, в каждом из

которых, был получен один из основных видов взаимодействия, описанных нами выше. Нулевое взаимодействие было установлено в эксперименте Стернберга (1969) с изменением четкости опознаваемого тестового стимула и размера ранее предъявленного набора знаков. В эксперименте Дененберга и Мортонна (1962) о использовании двух независимых переменных - условий содержания крыс во время и после вскармливания — было обнаружено расходящееся взаимодействие. А пересекающееся взаимодействие было получено в эксперименте по изучению времени реакции, проведенном Саймоном и Руделлом (1967). Одной независимой переменной был тип команды («левая» и «правая»), а другой — ухо, на которое она подавалась (левое и правое). Выяснилось, что значительное влияние на время реакции оказывает то, с какой стороны поступает сигнал (а эта переменная была нерелевантной). В тех случаях, когда «ухо» не совпадало с типом команды, время реакции возрастало.

В каждом из упомянутых экспериментов для проверки комбинированных гипотез нужно было использовать факторную схему. Оценить справедливость соответствующих теорий и моделей, привлекая только одну независимую переменную, было бы невозможно. Взаимодействие, полученное Иерксом и Додсоном, подтвердило теорию о двух базисных переменных — различия стимулов и ассоциирования, на каждую из которых сила электроудара оказывала противоположное воздействие. Результаты Дененберга и Мортонна можно объяснить тем, что приручение в период вскармливания делает крыс неспособными к использованию всех преимуществ свободного передвижения для дальнейшего успешного научения в лабиринте. Результаты, полученные Стернбергом, подтвердили предложенную им модель разделения стадий процесса обработки информации — кодирование стимула и сканирование ранее предъявленного набора знаков при опознании одного из них. А обнаруженная Саймоном и Руделлом прочная связь между стимулируемым ухом и типом команды остается пока интригующей загадкой.

Существуют эксперименты, где изучаемых независимых переменных — больше двух, в них проверяются более сложные комбинированные гипотезы. Взаимодействие между двумя переменными называется взаимодействием первого порядка, между тремя — взаимодействием второго порядка и т.д.

Все обсуждавшиеся нами экспериментальные схемы можно классифицировать по трем основным параметрам.

Первый параметр — это способ сравнения условий, или уровней, независимой переменной: сравнение по каждому испытуемому, межгрупповое сравнение или сравнение по всем испытуемым.

Второй — это характер изменения независимой переменной: качественный или количественный.

И третий — число независимых переменных: либо эксперимент с одной переменной, либо факторный эксперимент.

Схемы не обязательно должны быть одинаковы (гомогенны) в отношении каждой из независимых переменных. Например, для одной независимой переменной может применяться межгрупповое сравнение, а для другой — интра- или кросс-индивидуальное.

ВОПРОСЫ

- I. С какой целью в эксперименте с рассечением свода использовалось несколько временных интервалов между пробой-образцом и пробой-выбором?
 1. 2. Приведите примеры основных результатов и взаимодействия в факторных экспериментах.
 2. Как измеряются основные результаты действия независимых переменных на графике?
 3. Опишите, как выглядят на графиках основные виды взаимодействия двух независимых переменных (если у каждой из них — по два уровня).
 4. Почему в экспериментах по проверке гипотез о воздействии независимой переменной на определенную базовую переменную часто требуется контрольная переменная?
 5. Расскажите, как использовать несколько уровней второй независимой переменной для повышения внешней валидности.
 6. Почему для проверки закона Иеркса — Додсона был необходим факторный эксперимент?
 7. Каким образом вытекает конкретная комбинированная гипотеза, проверенная в эксперименте Стернберга, из предложенной им модели процесса обработки информации?
 8. Изобразите на графике такие экспериментальные результаты, которые могли бы быть получены, если бы приручение крыс в период вскармливания не

влияло на успешность их научения в лабиринте, а содержание в просторных ящиках после вскармливания оказывало большую помощь, чем содержание в тесных клетках.

9. Придумайте комбинированную гипотезу, которая подтверждалась бы наличием пересекающегося взаимодействия. (Постарайтесь отвлечься от исследования Саймона и Руделла.)
10. Какое место (и почему) занимает эксперимент Иеркса и Додсона в системе экспериментальных схем?

Тема: Корреляционные исследования.

Лекция №10. 3 часа.

План:

1. Особенности построения корреляционных исследований.
2. Положительные и отрицательные стороны корреляционных исследований.
3. Схемы построения корреляционных исследований.
4. Методы контроля в корреляционных исследованиях.

Литература:

Годфруа Ж. Что такое психология?. М.Мир 1992, т.2.

Артемьева А. Вероятностные методы в психологии. М.1986.

Немов Р.С. Психодиагностика с элементами математической статистики. М.1997.

Опорные слова: интерпритация, популяция, выборка, корреляция, дисперсия, мера связи.

В тех случаях, когда для проверки некоторых гипотез относительно поведения нельзя использовать активный эксперимент, проводят корреляционные исследования.

В двух рассмотренных нами примерах — исследованиях предпосылок оптимальной приспособленности и влияния порядка рождения на интеллект — активный эксперимент был невозможен. В исследовании же по отбору контролеров просто трудно себе представить, какие активные изменения можно было бы внести.

В результате сравнения хорошо и плохо приспособленных испытуемых по ряду параметров было обнаружено, что главный фактор для хорошо приспособленных — это хорошие методы воспитания в детстве. Однако здесь присутствовали также другие переменные, смешивающиеся с методами воспитания. Двумя из них были социоэкономический статус и интеллект

родителей. Нельзя сказать, влияли ли они в действительности, так как соответствующий контроль не был организован. Другим источником возможного смещения были субъективные оценки. Поскольку оценки производились не «вслепую», оставалась возможность для предубежденности экспериментатора. Поскольку материал для оценок получился из бесед с испытуемыми, существовала также опасность предубеждения испытуемого — хорошо приспособленные испытуемые могли быть склонны к более благоприятным воспоминаниям.

В исследовании, проведенном почти на 400000 молодых людях, было обнаружено, что показатель интеллекта снижается вместе с порядком рождения. Этот эффект сохранился даже при нивелировании возможного влияния социального положения и размера семьи. В общем, контроль в этом исследовании был достаточно хорошим. Однако было отмечено, что социальные слои состоят из довольно разнообразных групп, которые могут различаться по значимым побочным переменным. Корреляция между показателем интеллекта и порядком рождения может быть объяснена разными причинами; их можно разделить на пренатальные и постнатальные. В этом исследовании не были проконтролированы две переменные: первая — возраст родителей, особенно возраст матери, который может оказаться очень важным; вторая — брачный стаж родителей к моменту рождения испытуемого.

Методы контроля в корреляционных исследованиях могут быть двух видов.

Первый — индивидуальный подбор пар испытуемых. Если сравниваются две группы, то каждому индивиду первой группы подбирается индивид второй группы, который имеет тот же уровень определенных побочных переменных. Данный метод связан с двумя трудностями.

Первая: чем больше переменных, по которым уравниваются испытуемые, тем меньше испытуемых оказывается в распоряжении исследователя. Отсюда понятно, что много побочных переменных проконтролировать невозможно.

Вторая: оставшиеся испытуемые — не совсем типичны для сравниваемых групп, и полученные отношения оказываются непредставительными.

Таким образом, имеется опасность как слишком слабого, так и слишком строгого контроля.

Другим методом является выделение однородных подгрупп, как это делалось в исследовании порядка рождения. Например, была выделена подгруппа испытуемых с семьей из пяти человек, и уже внутри этой подгруппы выяснялось

влияние порядка рождения. Число различных подгрупп совпадает с числом уровней побочной переменной.

Таким образом, оказывается возможным обнаружить, как и в случае с двумя независимыми переменными в активном эксперименте, существует ли значимое взаимодействие между рассматриваемыми переменными. В корреляционных исследованиях это может также помочь пониманию отношений между переменными. Например, небольшое понижение линии тестовых оценок интеллекта для работников сельского хозяйства может быть следствием большего желания завести еще одного ребенка. Если бы имелись данные о таком желании, может удалась бы показать, что оно лежит в основе влияния порядка рождения. Это следовало бы, во-первых, из понижения линии тестовых оценок интеллекта в зависимости от порядка рождения для однородных подгрупп как городских, так и сельских жителей, не желающих следующего ребенка, и во-вторых, из отсутствия такого понижения для подгрупп, желающих ребенка.

Однако поскольку доказательство оставалось бы корреляционным, нельзя было настаивать на том, что желание иметь еще одного ребенка — основа рассматриваемого эффекта. Например, желание могло просто коррелировать с физической энергией родителей, и эта последняя могла быть истинной причиной основного эффекта. Таковы корреляционные исследования. Здесь могут быть проверены интересные гипотезы, предложено много разных идей. И все-таки контроль здесь никогда не будет таким же хорошим, как в активном эксперименте.

Последнее исследование приводилось как искусственный пример того, как можно использовать тест на способности, чтобы произвести отбор контролеров. Группа испытуемых проводилась через тест на специальные способности. Затем оценивалось качество их работы. Отношение между этими двумя переменными — тестовыми оценками и оценками работы — было представлено на диаграмме разброса. Форма этой диаграммы может быть отражена в коэффициенте корреляции. Положительные значения коэффициента находятся в пределах от 0 до 1. С помощью диаграмм разброса представляющих разные значения коэффициента корреляции, было показано, как можно отобрать претендентов согласно заданному правилу решения, например, чтобы, по крайней мере, 80% из всех выбранных оказались удовлетворительно работающими. Чем выше коэффициент корреляции, тем ниже может быть расположена так называемая секущая оценка для теста на способности, что позволяет увеличить число отобранных индивидов при данном правиле решения. Корреляция позволяет также

предсказать с помощью линии предсказания, связывающей средние рабочие оценки с тестовыми оценками, какой рабочей оценки может достичь индивид. Чем выше коэффициент корреляции, тем круче эта линия и тем меньше ошибка предсказания.

Были рассмотрены общие черты и различия разных корреляционных исследований. Все эти исследования сходны в том, что переменные в них уже существуют, в отличие от активных экспериментов, где условия независимой переменной активно организуются для выявления влияния последней на зависимую переменную. Различаются же они между собой по многим параметрам.

Прежде всего одна из переменных в них в разной степени приближается к независимой переменной. Например, в исследовании психологической приспособленности факторы — предпосылки вполне могут быть расценены как независимые переменные. Другая крайность представлена исследованием по отбору контролеров, где ни одна переменная не похожа на независимую. Поэтому однонаправленность предсказания определяется только практическими целями.

Во-вторых, корреляционные исследования различаются тем, вычисляется ли коэффициент корреляции или нет. Представление степени связи с помощью коэффициента корреляции имеет наибольший смысл в том случае, если значения каждой переменной образуют непрерывное колоколообразное распределение. Этому условию почти всегда удовлетворяют два множества тестовых оценок. Кроме того, такие переменные, как доход или балльные оценки свойств личности, так-же зачастую достаточно хорошо соответствуют этим условиям.

В-третьих, корреляционные исследования различаются по цели: имеет ли оно чисто познавательную цель или предполагает немедленное практическое приложение. Тот факт, что в практическом исследовании, приведенном в качестве примера в этой главе, использовался коэффициент корреляции и соотносились две оценки поведения для каждого испытуемого, вовсе не означает, что исследование такого типа не может служить целям расширения нашего познания.

ВОПРОСЫ

1. Почему исследование, в котором сравниваются хорошо и плохо приспособленные группы, называется корреляционным?
2. Почему сопутствующее смещение всегда присутствует в корреляционном исследовании и только иногда — в активном эксперименте?
3. Как можно использовать идеальный эксперимент в качестве эталона внутренней валидности в корреляционном исследовании?

4. С какими трудностями связан контроль смешения путем индивидуального подбора пар? Приведите пример.
5. Дайте пример того, как использование однородных подгрупп может обеспечить сведения о взаимодействии.
6. Какие смешивающиеся переменные были упущены исследователями при определении влияния порядка рождения?
7. Почему даже самое лучшее корреляционное исследование ограничено в выделении переменных, влияющих на поведение?
8. Почему высокий коэффициент корреляции позволяет отобрать пропорционально большее число индивидов, работающих качественно?
9. По каким параметрам различаются корреляционные исследования?

