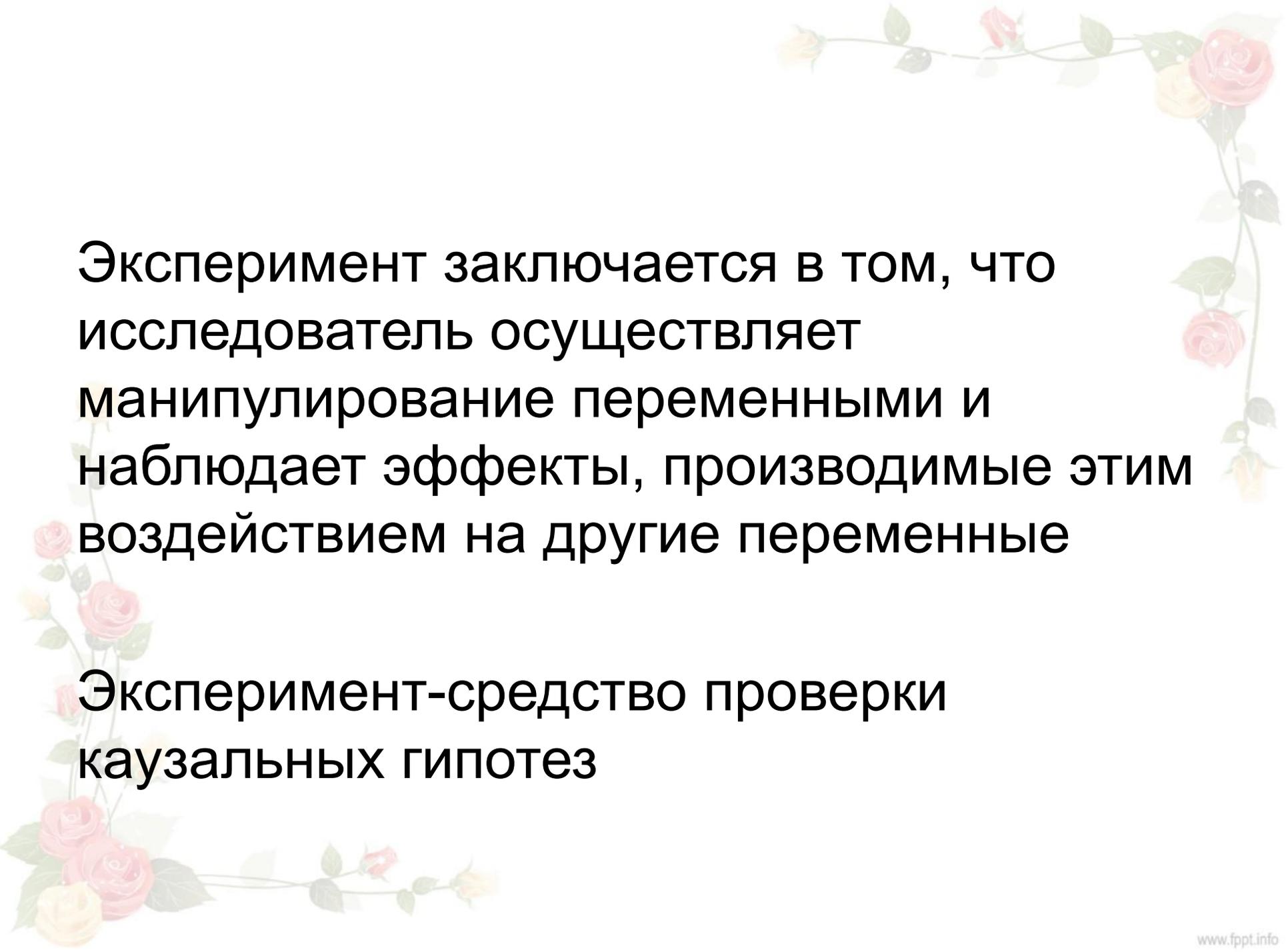


Экспериментальные планы



MARTIN

phillipmartin.info

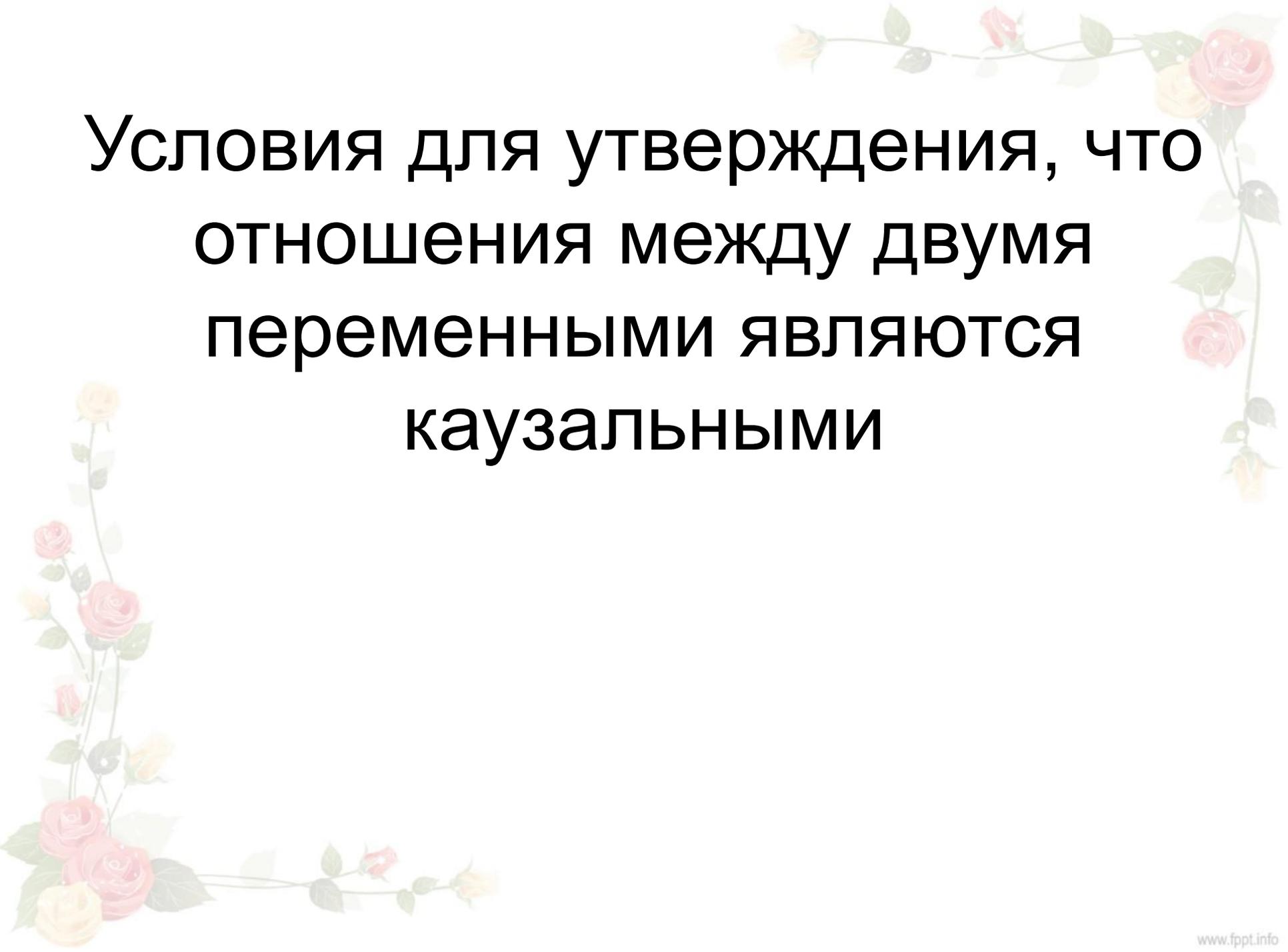


Эксперимент заключается в том, что исследователь осуществляет манипулирование переменными и наблюдает эффекты, производимые этим воздействием на другие переменные

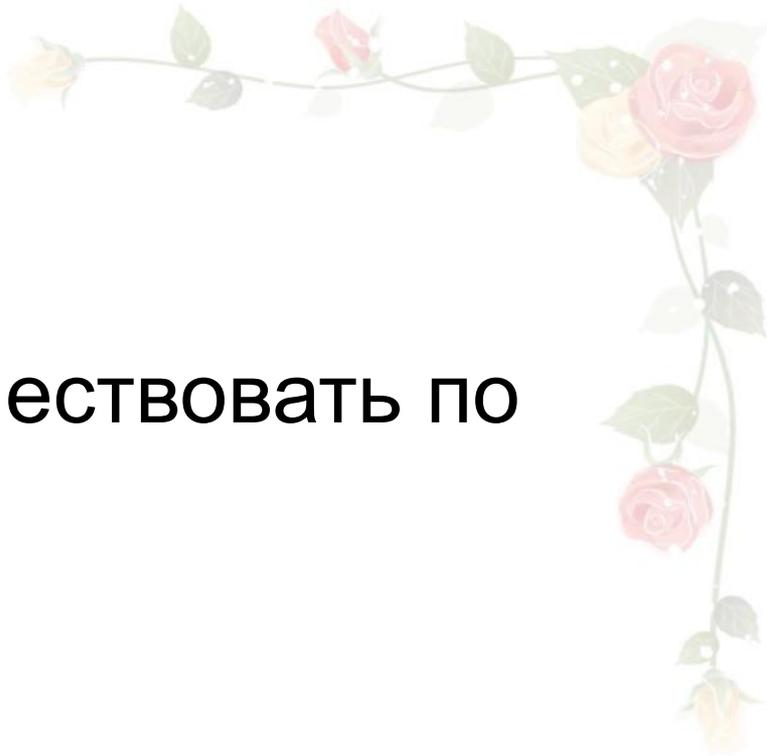
Эксперимент-средство проверки каузальных гипотез

Каузальные отношения
(причина-следственные –causal
не путать с casual)





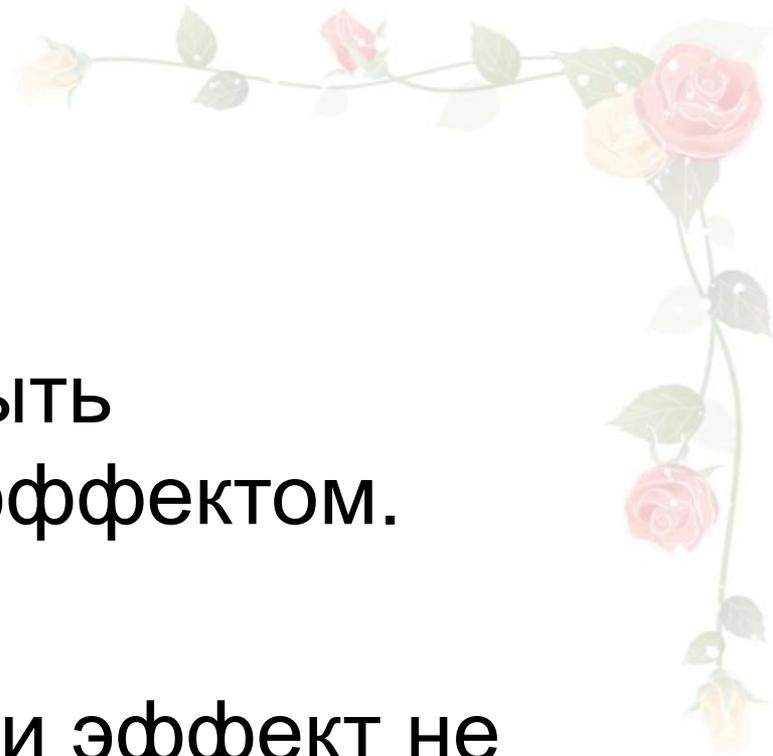
Условия для утверждения, что
отношения между двумя
переменными являются
каузальными



1. Причина должна предшествовать по времени эффекту.

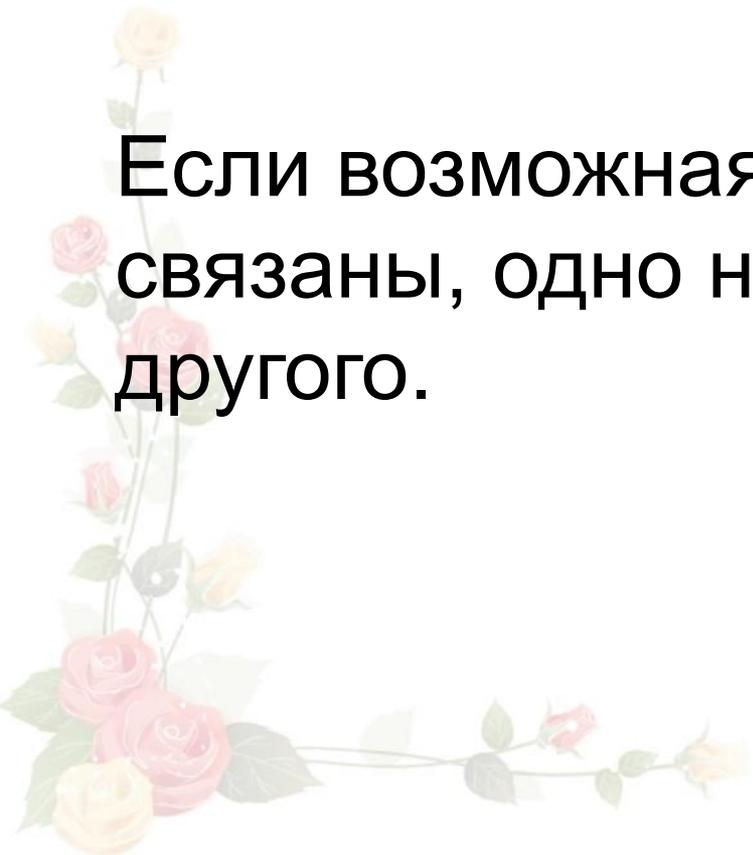


Воздействие А должно предшествовать эффекту Б.

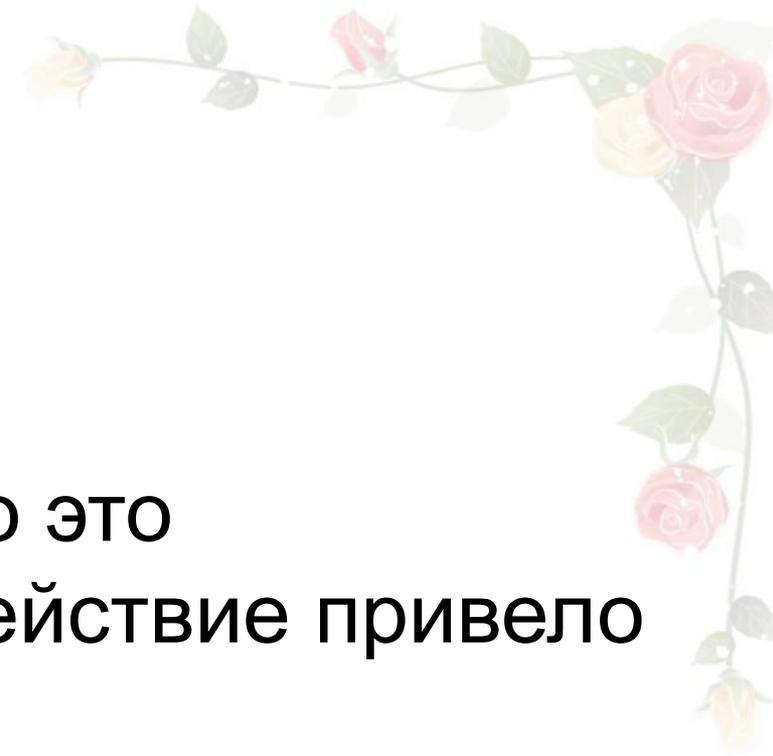


2. Воздействие должно быть статистически связано с эффектом.

Если возможная причина и эффект не связаны, одно не может быть причиной другого.

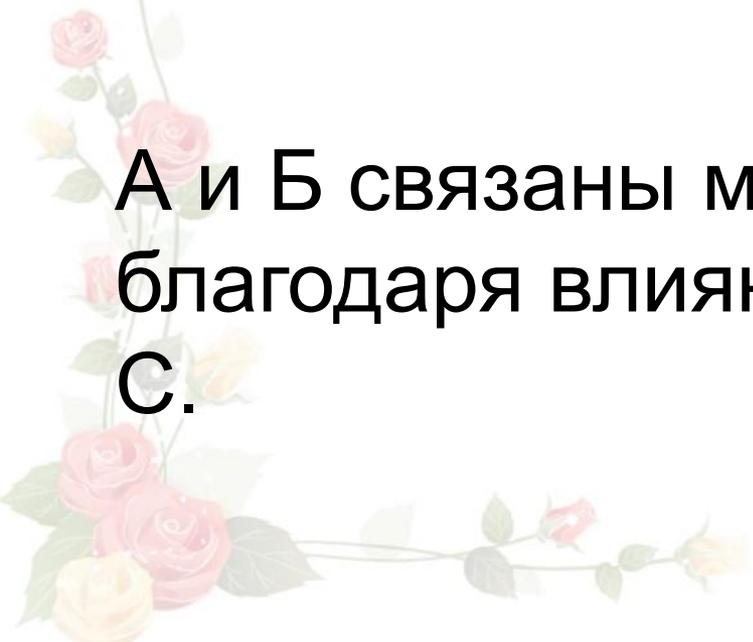


3. Не должно существовать правдоподобного альтернативного объяснения появления Б помимо А.



Внутренняя валидность

Действительно ли именно это экспериментальное воздействие привело к изменениям?

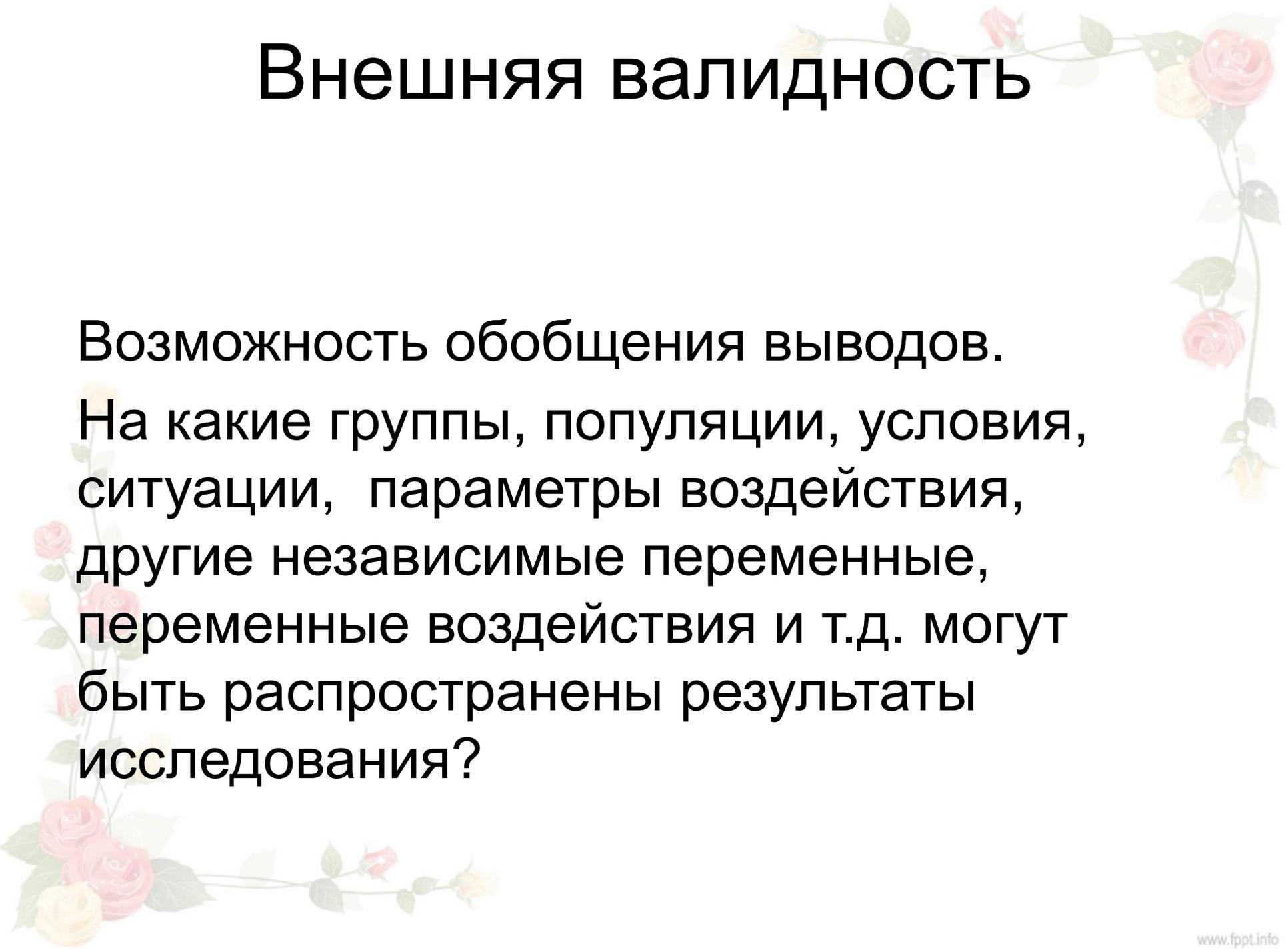


А и Б связаны между собой только благодаря влиянию третьей переменной С.

Конструктивная валидность



Внешняя валидность



Возможность обобщения выводов.
На какие группы, популяции, условия, ситуации, параметры воздействия, другие независимые переменные, переменные воздействия и т.д. могут быть распространены результаты исследования?

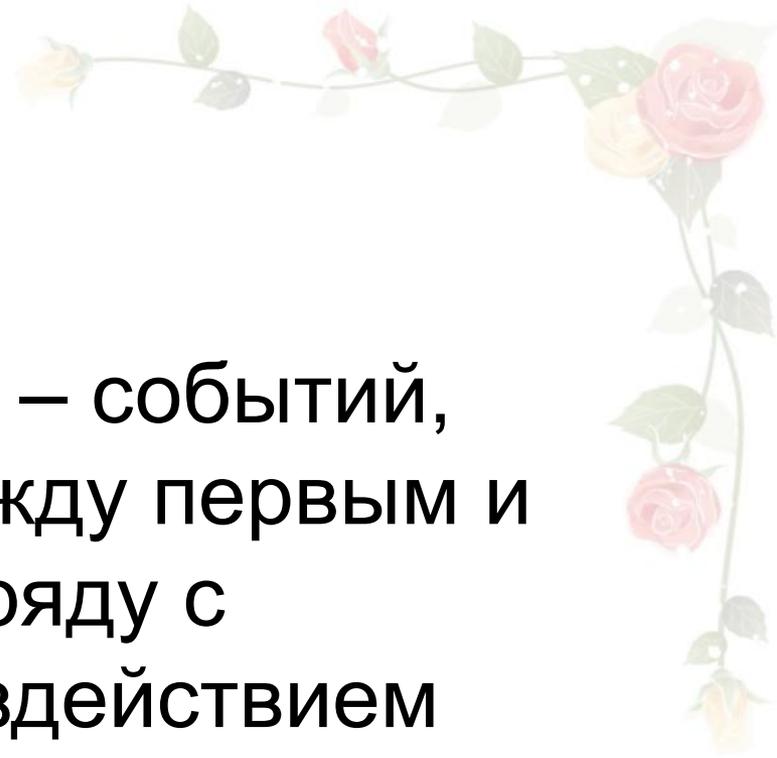
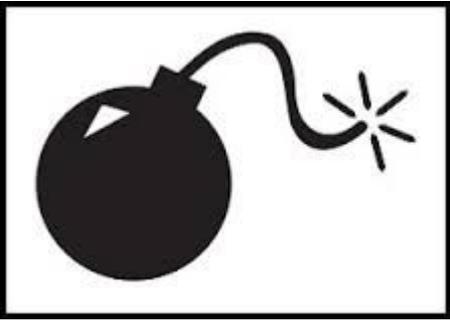
Факторы, угрожающие внутренней валидности



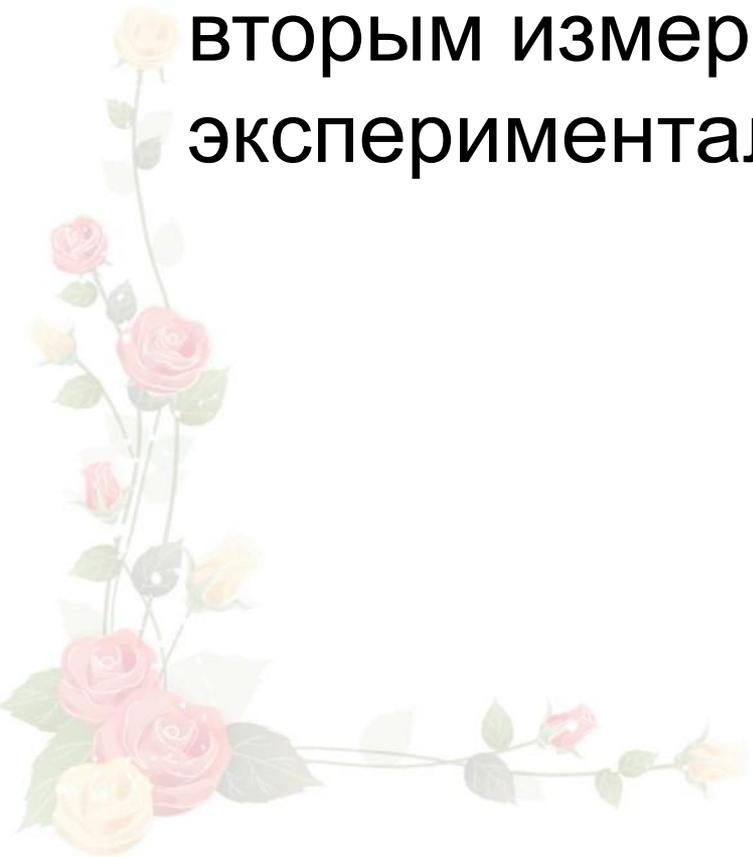
- Если эти переменные не контролируются в экспериментальном плане, то они могут давать эффекты, которые смешиваются с эффектом экспериментального воздействия

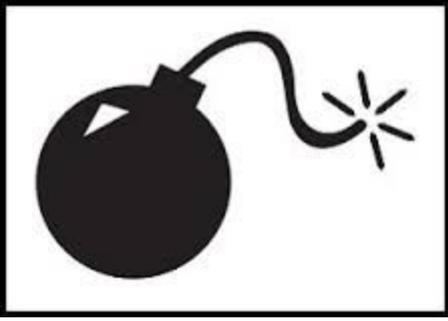
- O1X02





- **Влияние фона** (history) – событий, которые происходят между первым и вторым измерением наряду с экспериментальным воздействием

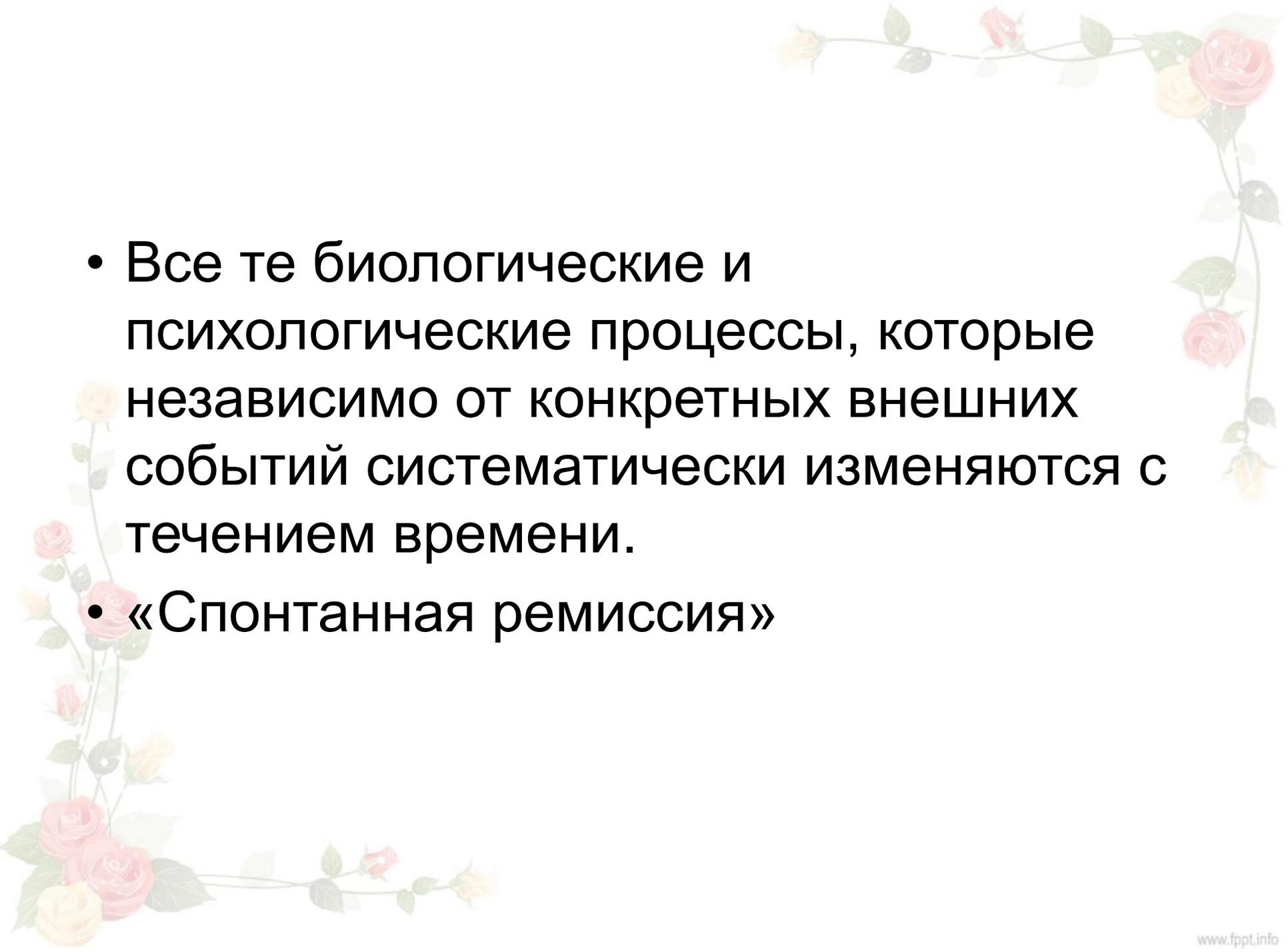


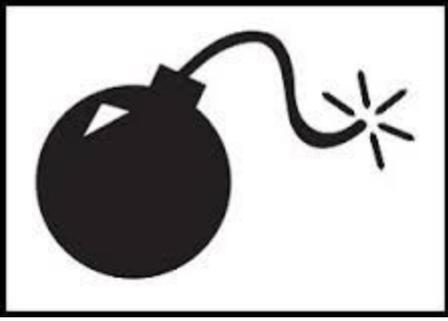


Влияние естественного развития
(maturation) – изменения испытуемых,
являющиеся следствием течения
времени.

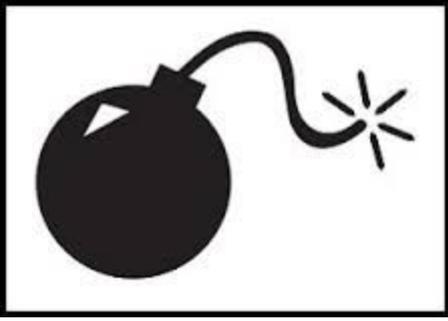
Взросление, усиление голода,
усталость....



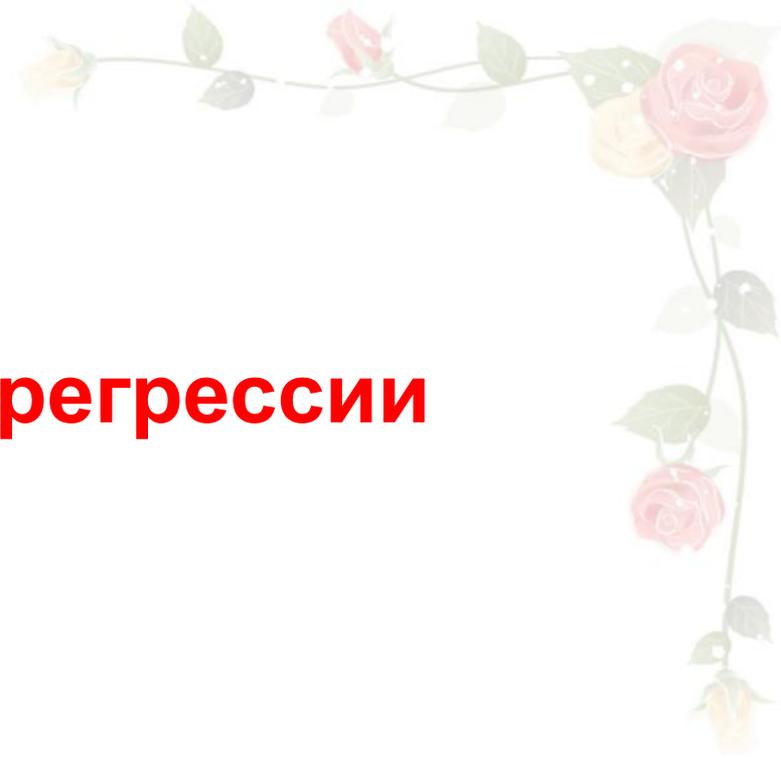
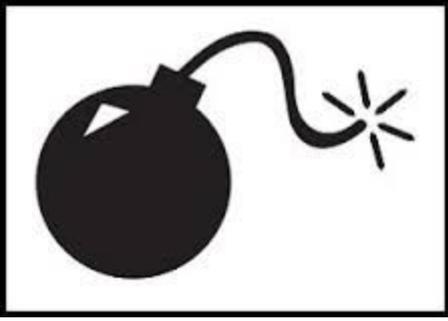
- 
- Все те биологические и психологические процессы, которые независимо от конкретных внешних событий систематически изменяются с течением времени.
 - «Спонтанная ремиссия»



Эффект тестирования(testing) –
влияние выполнения заданий,
применяемых для измерения, на
результаты повторного испытания.



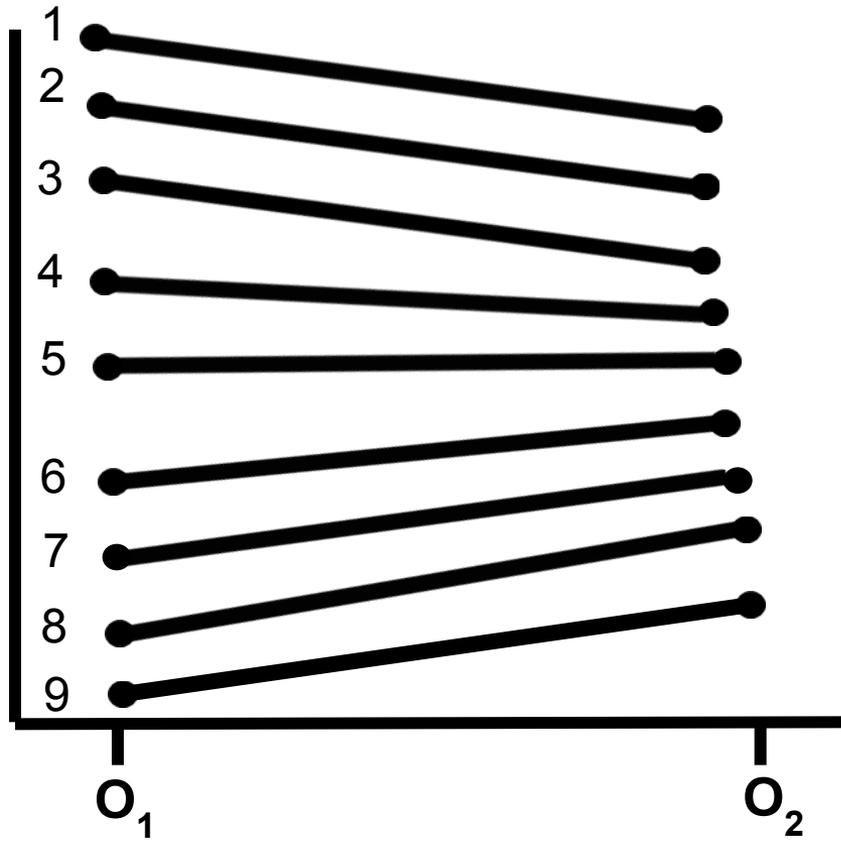
- **Фактор инструментальной погрешности**, нестабильности измерительного инструмента (instrumentation), при которой изменения в калибровке инструмента или изменения, характеризующие наблюдателя или оценочные показатели, могут вызвать изменения в результатах измерения

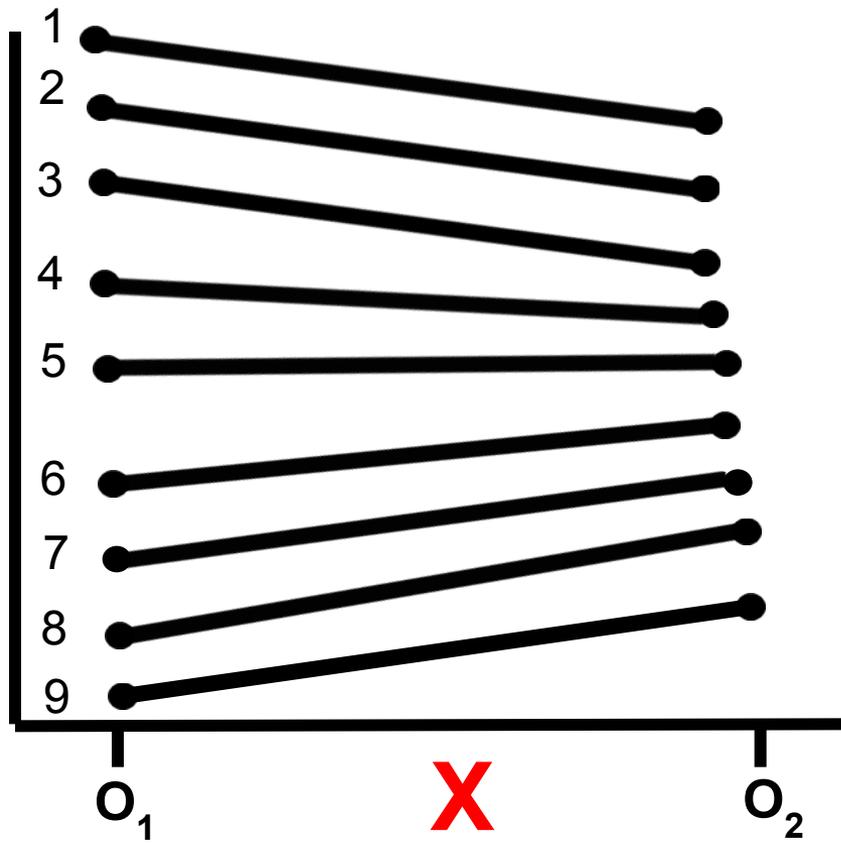


Фактор статистической регрессии
(statistical regression) –

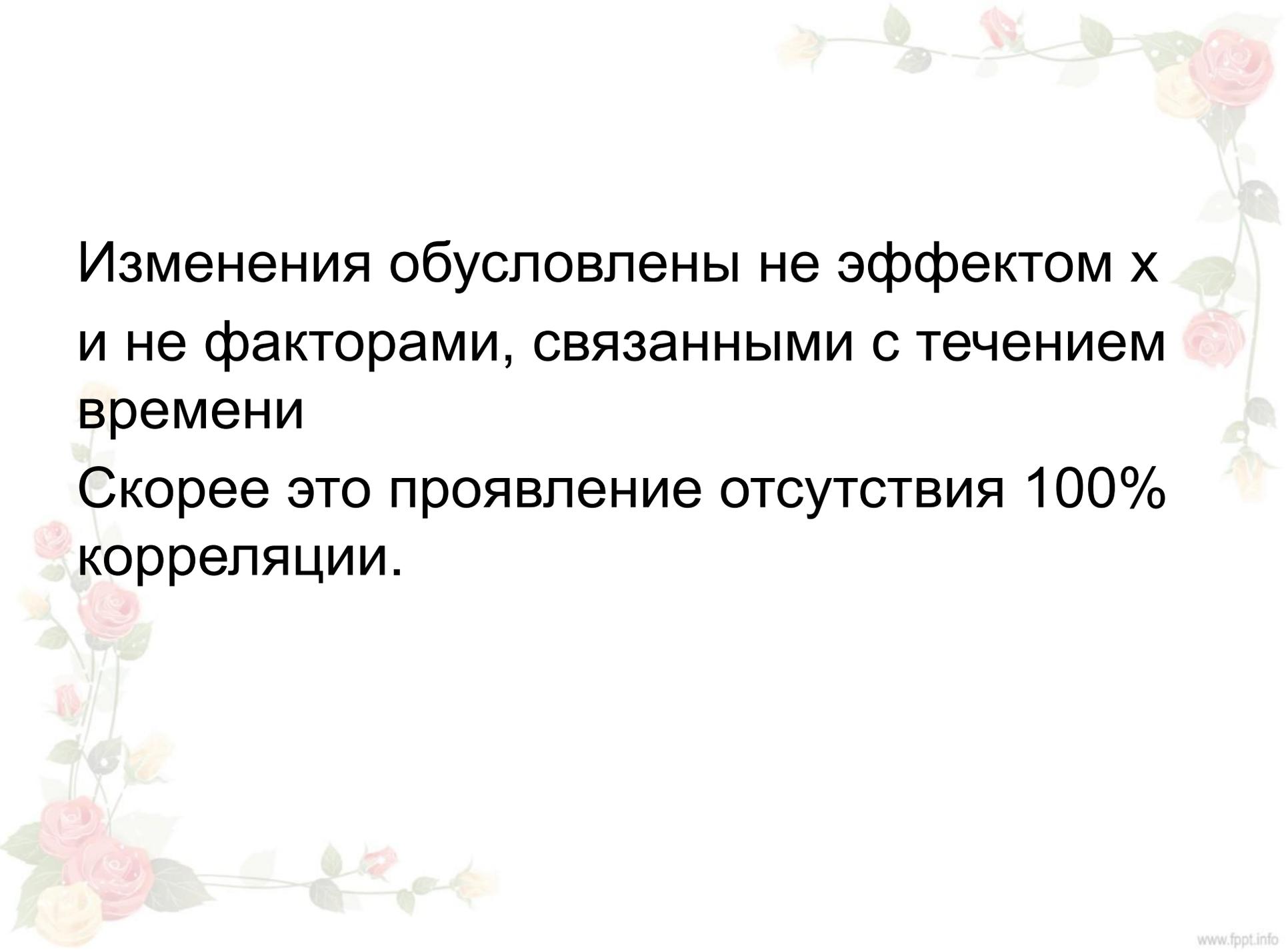


- Худшие результаты улучшаются при повторном тестировании, а лучшие ухудшаются.



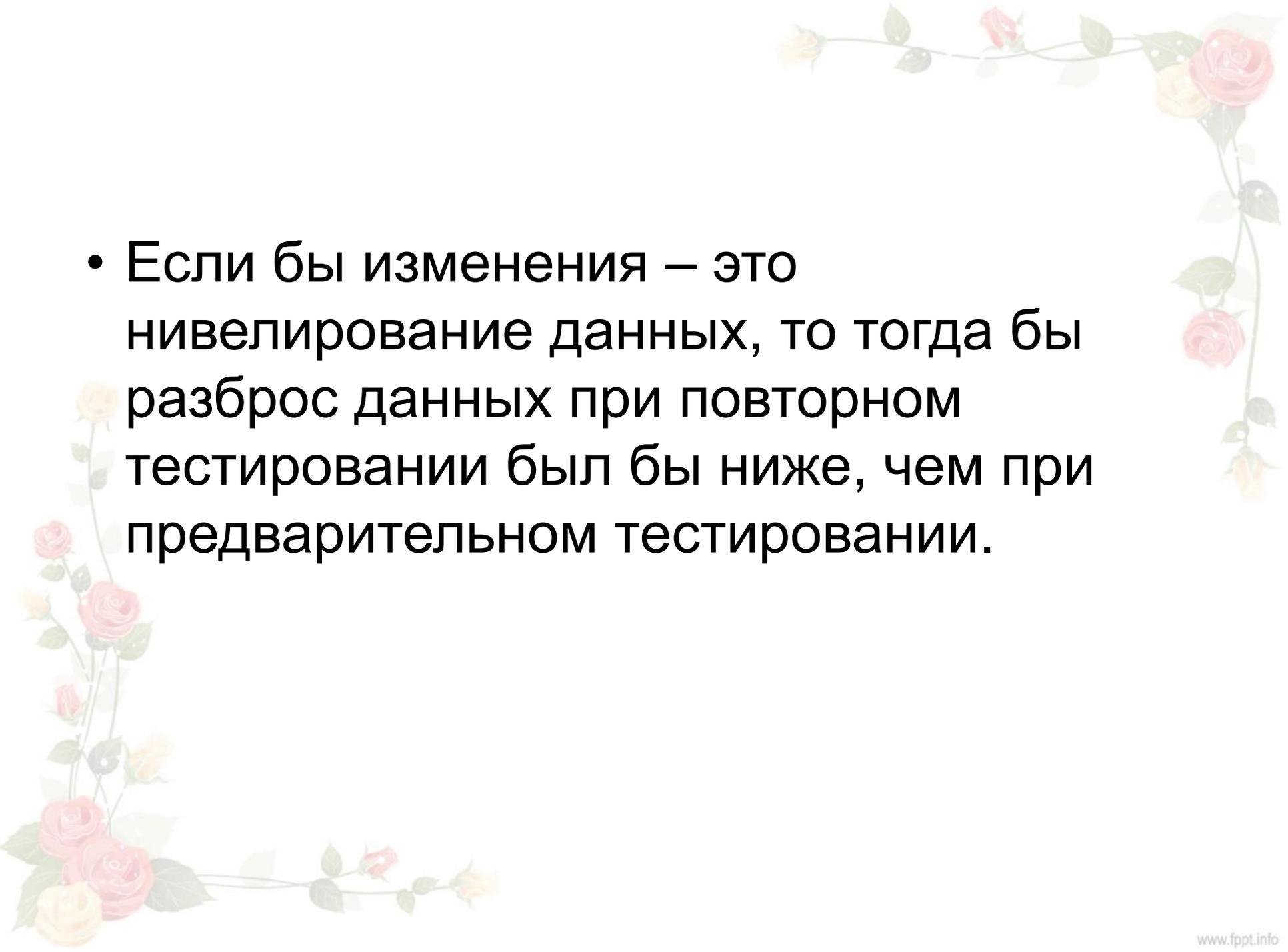


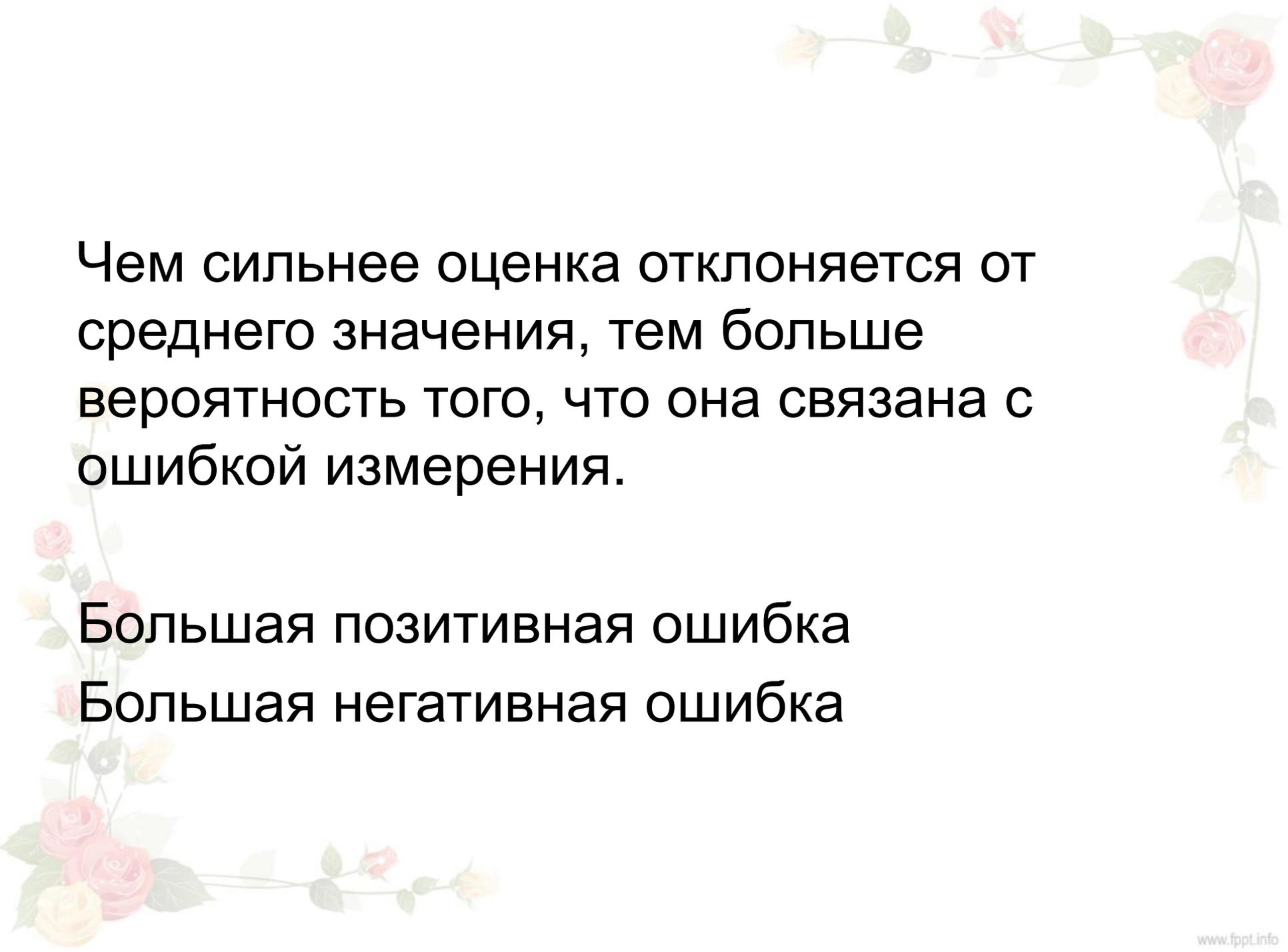




Изменения обусловлены не эффектом x
и не факторами, связанными с течением
времени

Скорее это проявление отсутствия 100%
корреляции.

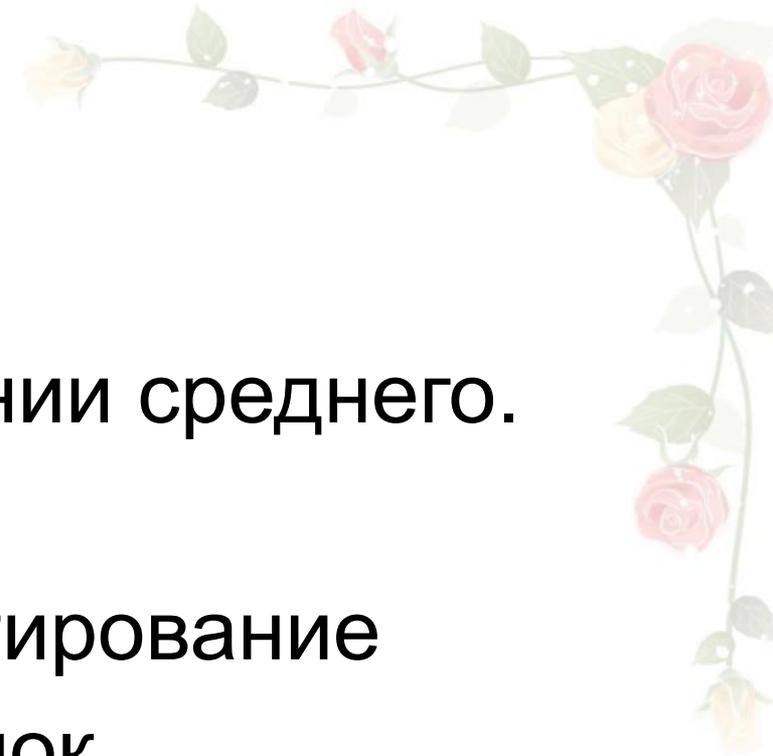
- 
- Если бы изменения – это нивелирование данных, то тогда бы разброс данных при повторном тестировании был бы ниже, чем при предварительном тестировании.



Чем сильнее оценка отклоняется от среднего значения, тем больше вероятность того, что она связана с ошибкой измерения.

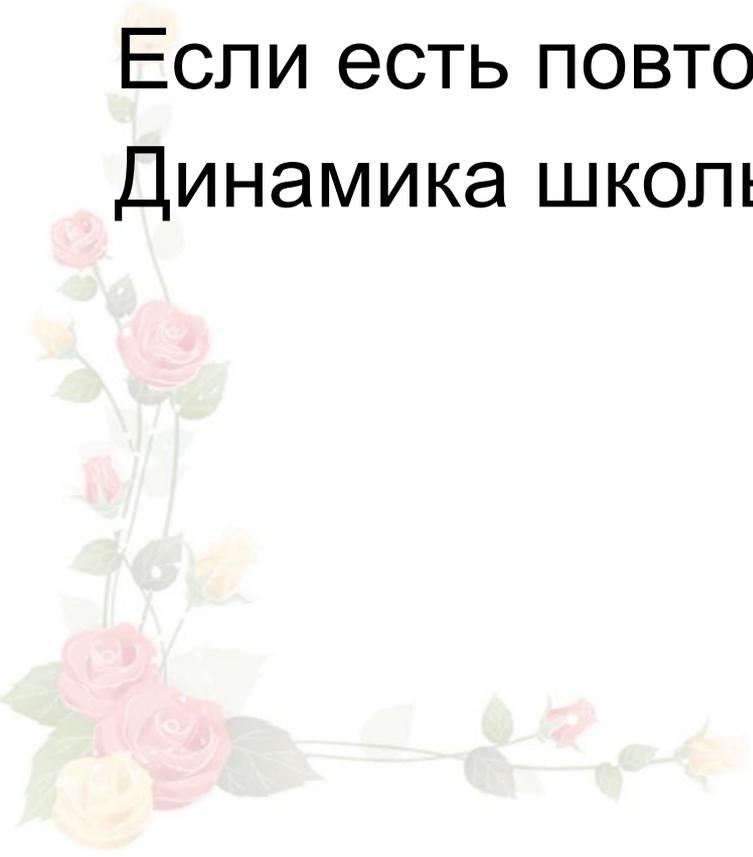
Большая позитивная ошибка

Большая негативная ошибка



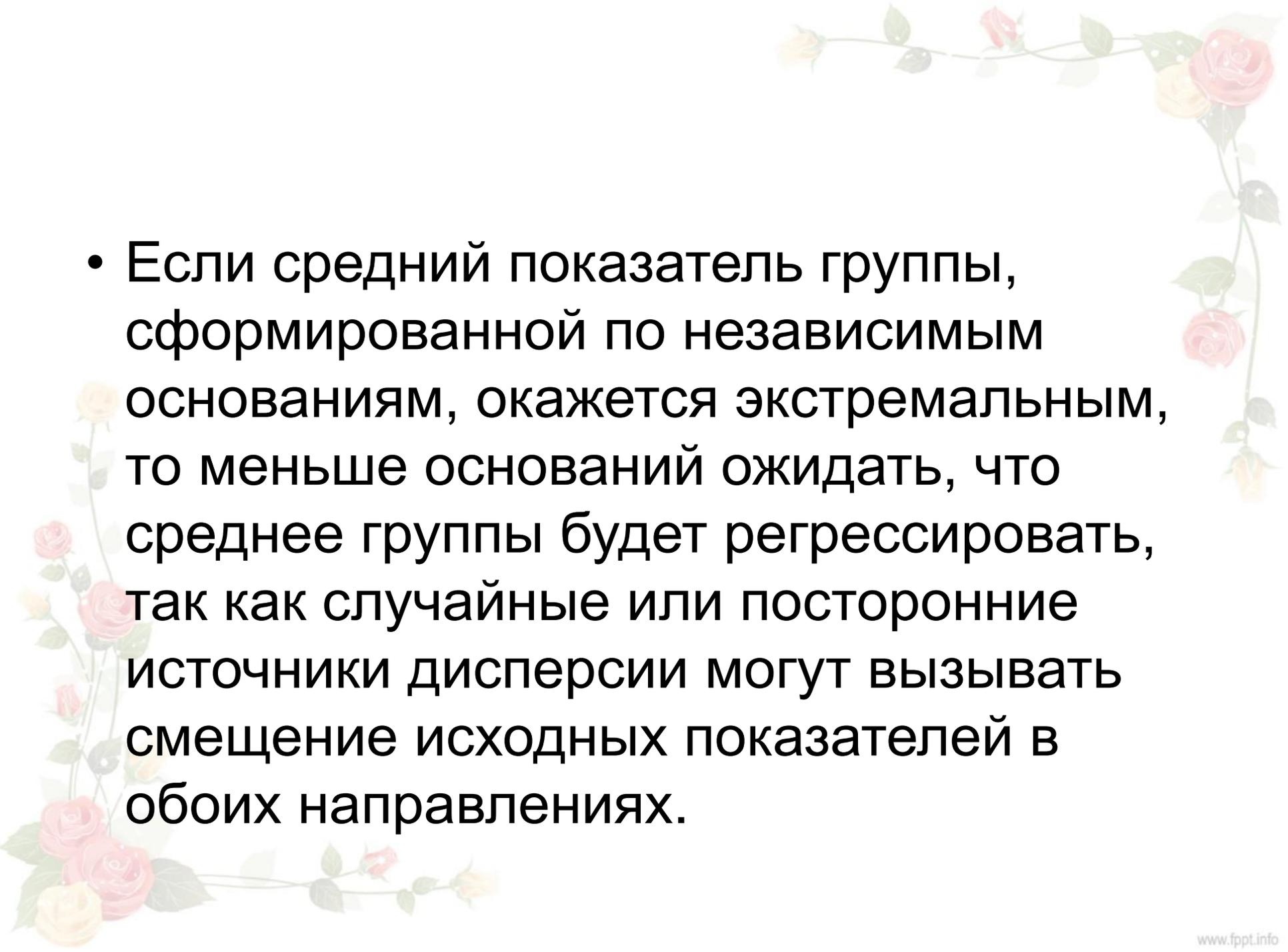
РЕГРЕССИЯ в направлении среднего.

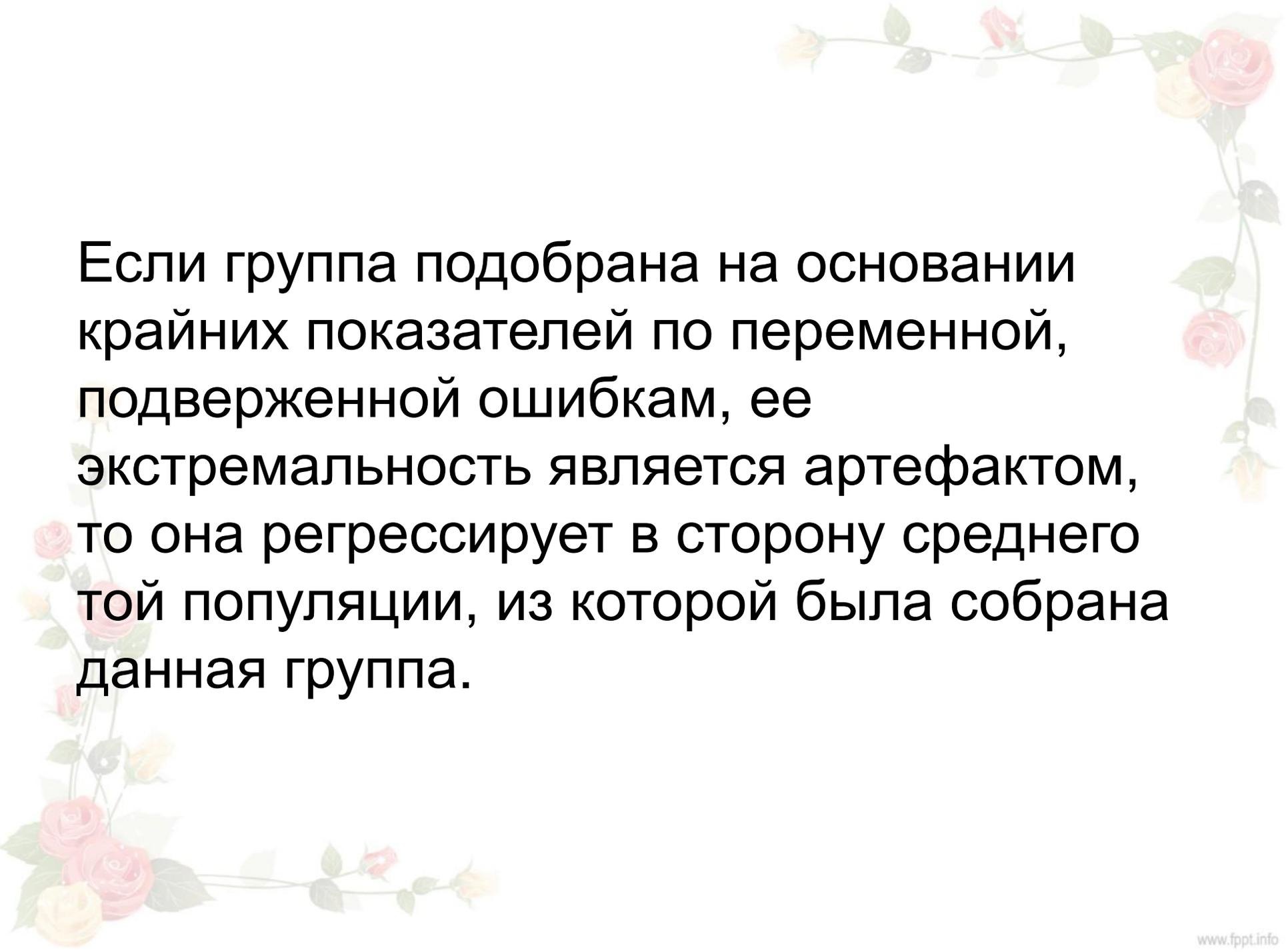
Если есть повторное тестирование
Динамика школьных оценок



Чем ниже корреляция, тем в большей степени выражена регрессия к среднему.

- Если группы отобраны по контрасту, на основе крайних показателей

- 
- Если средний показатель группы, сформированной по независимым основаниям, окажется экстремальным, то меньше оснований ожидать, что среднее группы будет регрессировать, так как случайные или посторонние источники дисперсии могут вызывать смещение исходных показателей в обоих направлениях.



Если группа подобрана на основании крайних показателей по переменной, подверженной ошибкам, ее экстремальность является артефактом, то она регрессирует в сторону среднего той популяции, из которой была собрана данная группа.

Косвенные эффекты регрессии

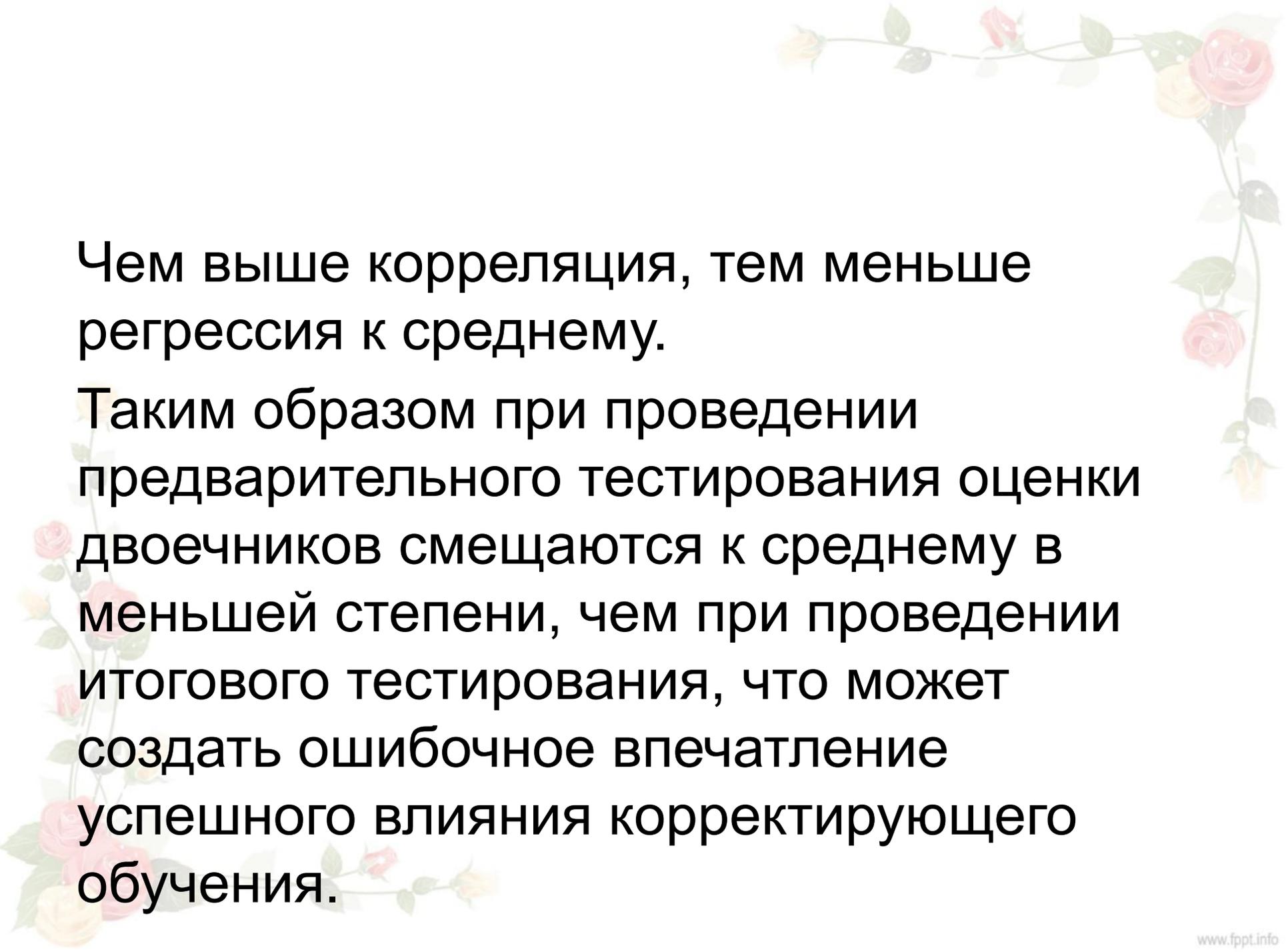
- Связаны с отбором по другим экспериментальным показателям, а не по данным предварительного тестирования.

Провалившие тестирование
(контрольная работа + тест:форма А)
Экспериментальное обучение (тест: форма Б)



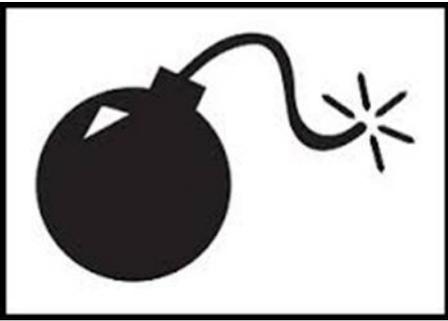
- Результаты контрольной лучше коррелируют с формой А, чем с результатами формы Б, которая применяется через несколько месяцев.





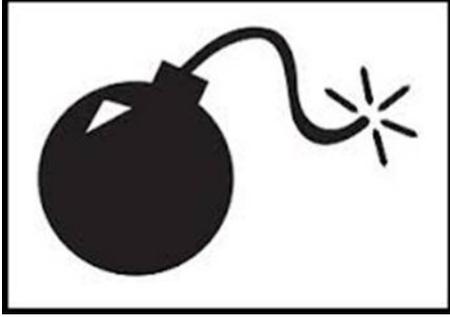
Чем выше корреляция, тем меньше регрессия к среднему.

Таким образом при проведении предварительного тестирования оценки двоечников смещаются к среднему в меньшей степени, чем при проведении итогового тестирования, что может создать ошибочное впечатление успешного влияния корректирующего обучения.



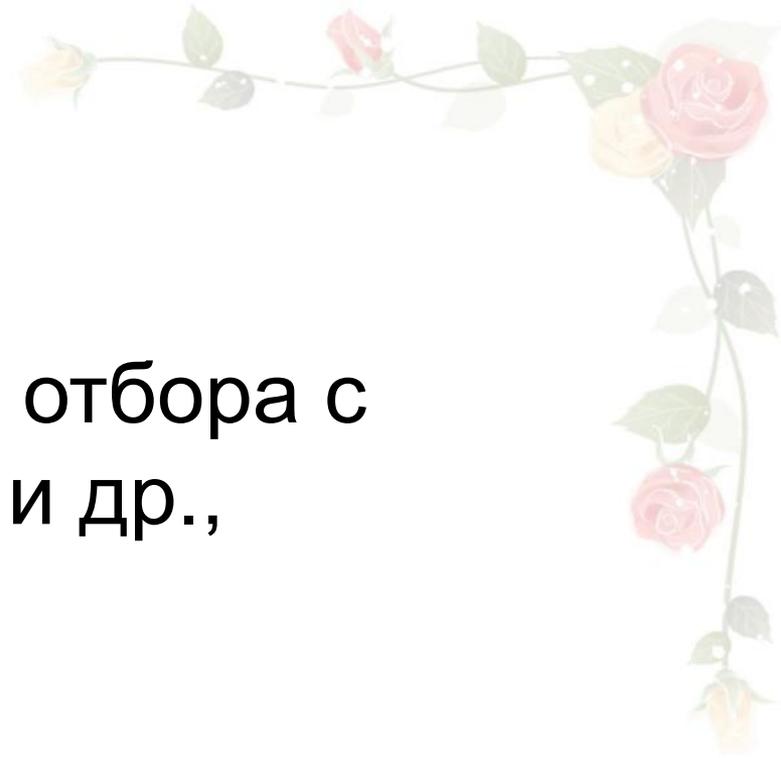
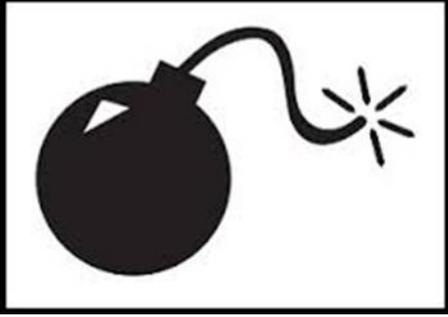
Факторы отбора испытуемых (selection)

—
неэквивалентности групп по составу,
вызывающей появление систематической
ошибки в результатах.



Факторы отсева в ходе эксперимента
(experimental mortality) – неравномерность
выбывания испытуемых из сравниваемых
групп



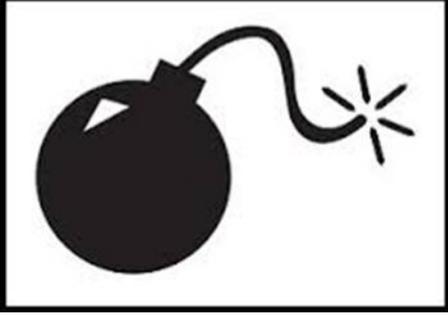


Взаимодействия фактора отбора с
естественным развитием и др.,



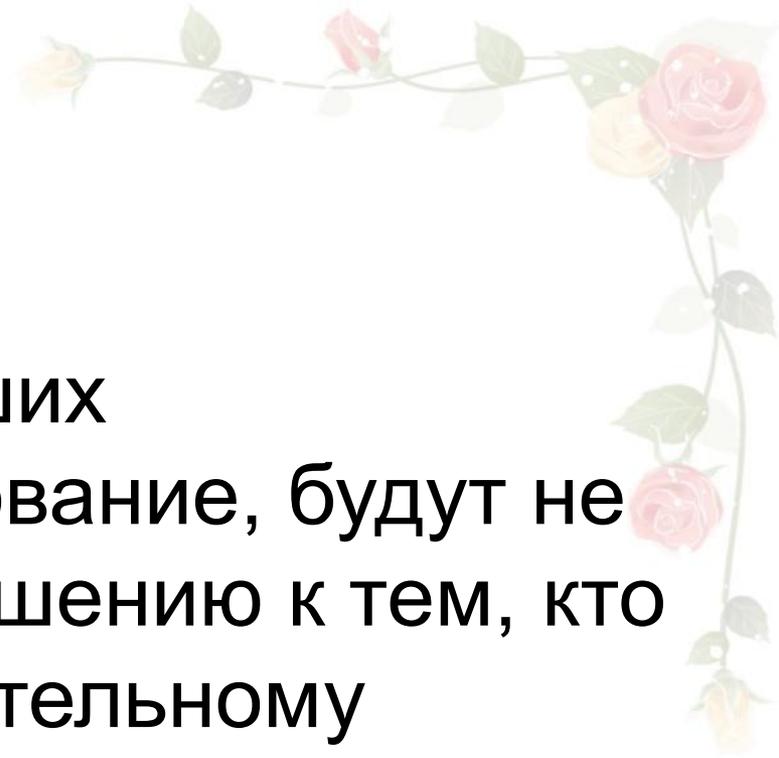
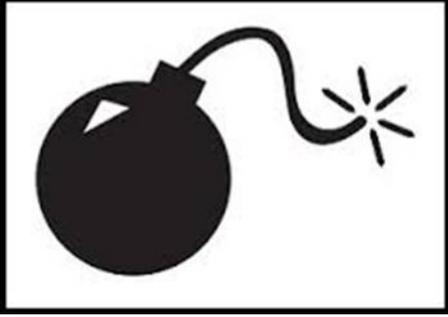
Факторы, угрожающие внешней валидности



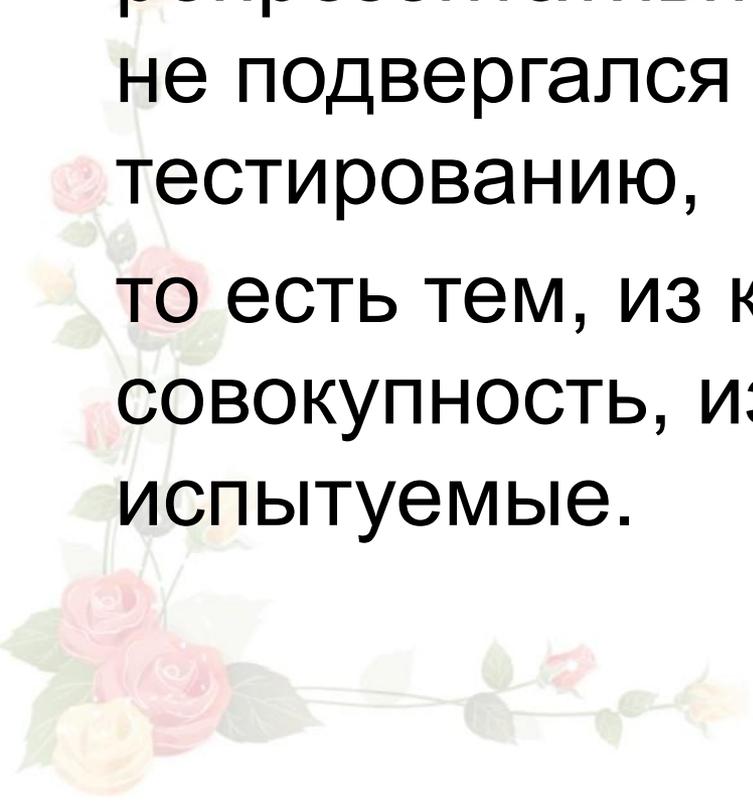


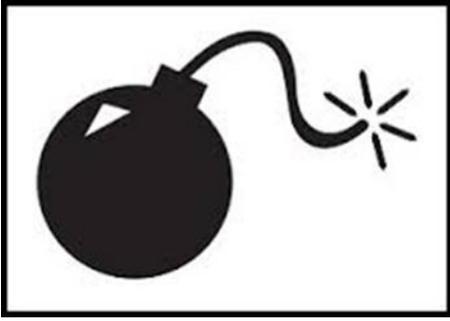
Реактивный эффект, или эффект взаимодействия тестирования, -

Возможное уменьшение или увеличение сензитивности или восприимчивости испытуемых к экспериментальному воздействию под влиянием предварительного тестирования

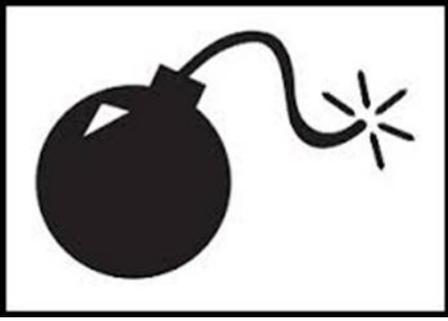


Результаты лиц, прошедших предварительное тестирование, будут не репрезентативны по отношению к тем, кто не подвергался предварительному тестированию, то есть тем, из кого состоит генеральная совокупность, из которых были отобраны испытуемые.

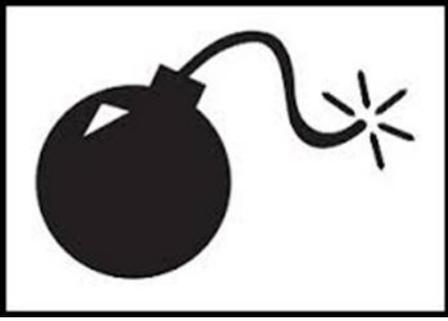




Эффекты взаимодействия фактора отбора и экспериментального воздействия.



- Условия организации эксперимента, вызывающие реакцию испытуемых на эксперимент, которая не позволяет распространить полученные данные о влиянии экспериментальной переменной на лиц, подвергающихся такому же воздействию в неэкспериментальных условиях.



- Взаимная интерференция экспериментальных воздействий, возникающая когда одни и те же испытуемые подвергаются нескольким воздействиям, поскольку влияние более ранних воздействий не исчезает.

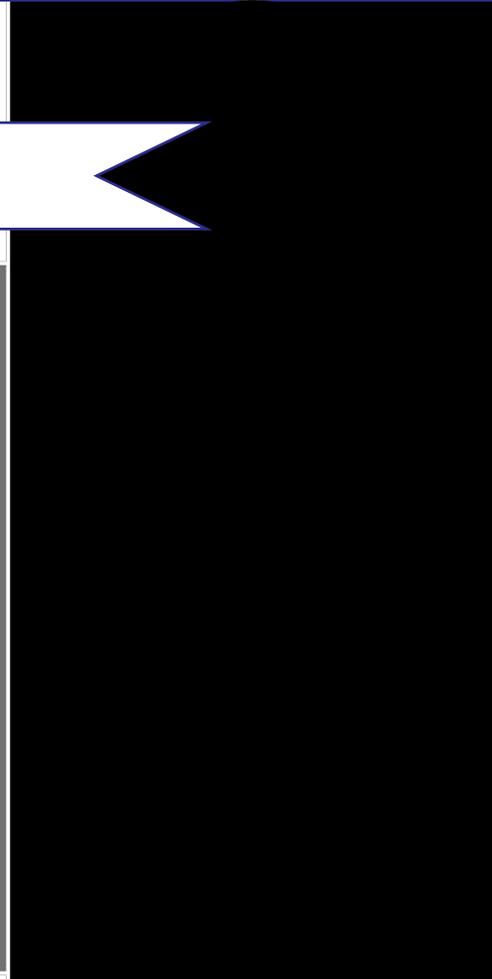
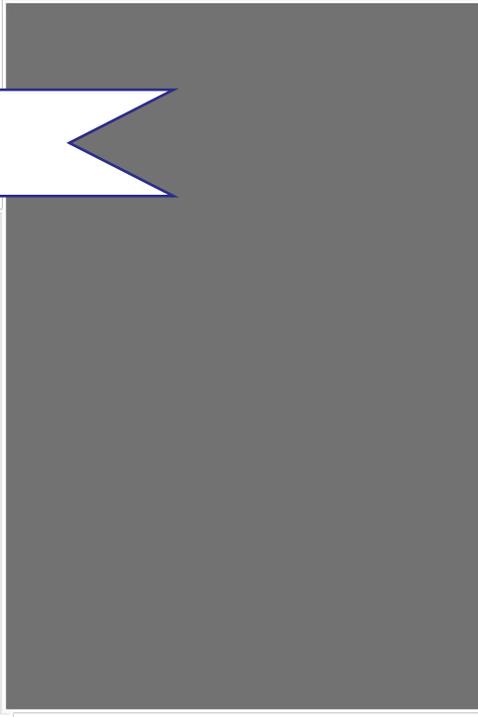
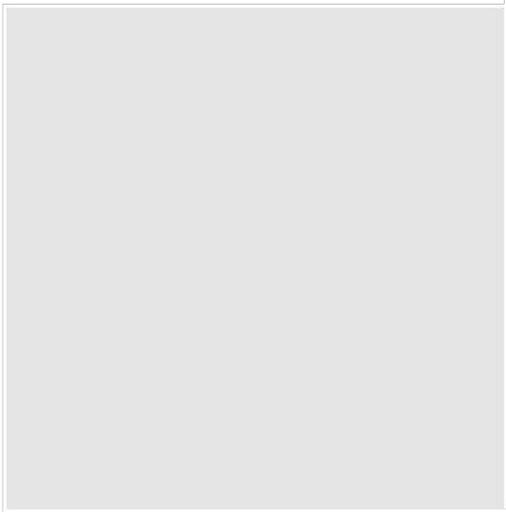
Критерий истинного эксперимента...

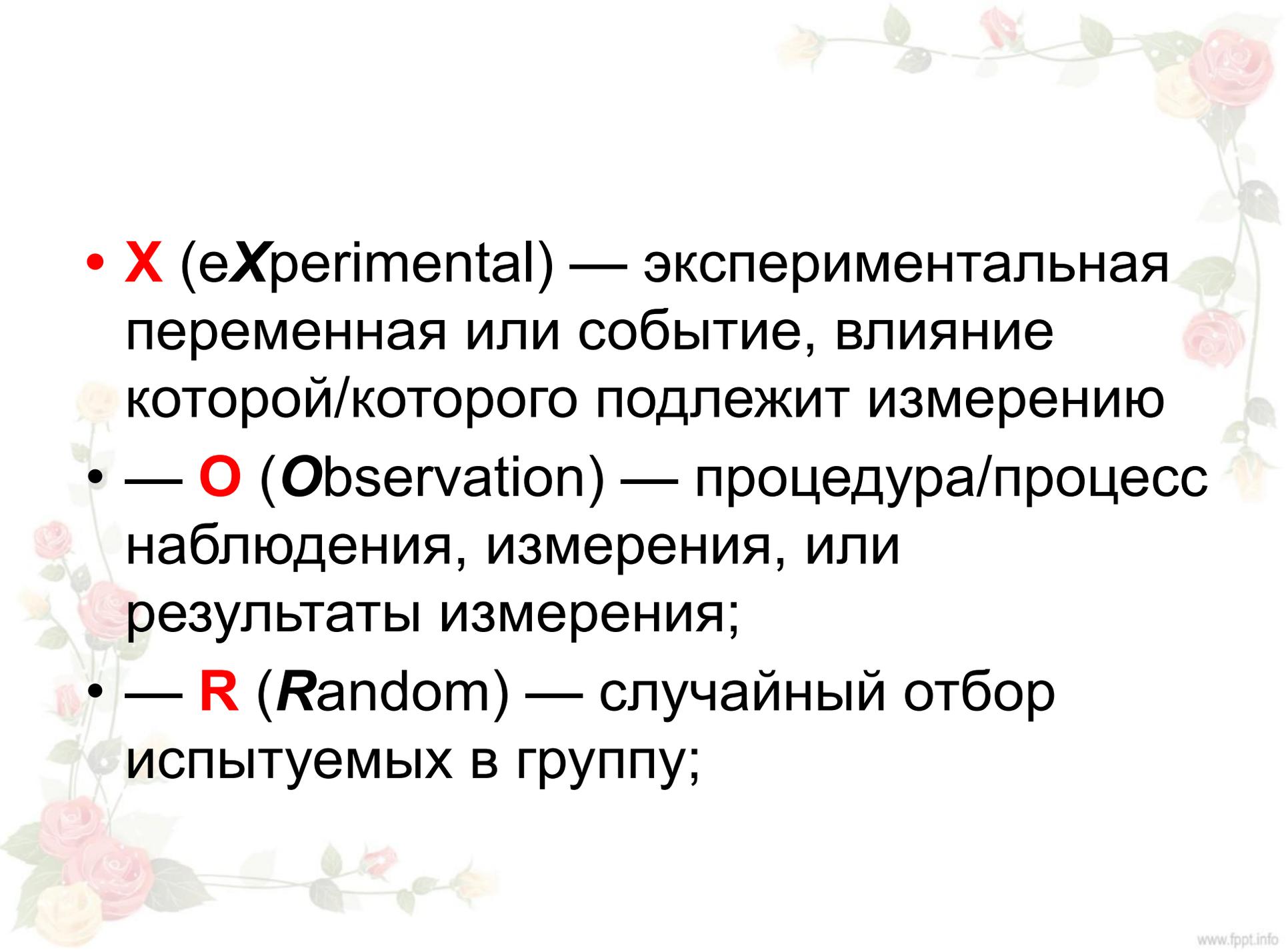


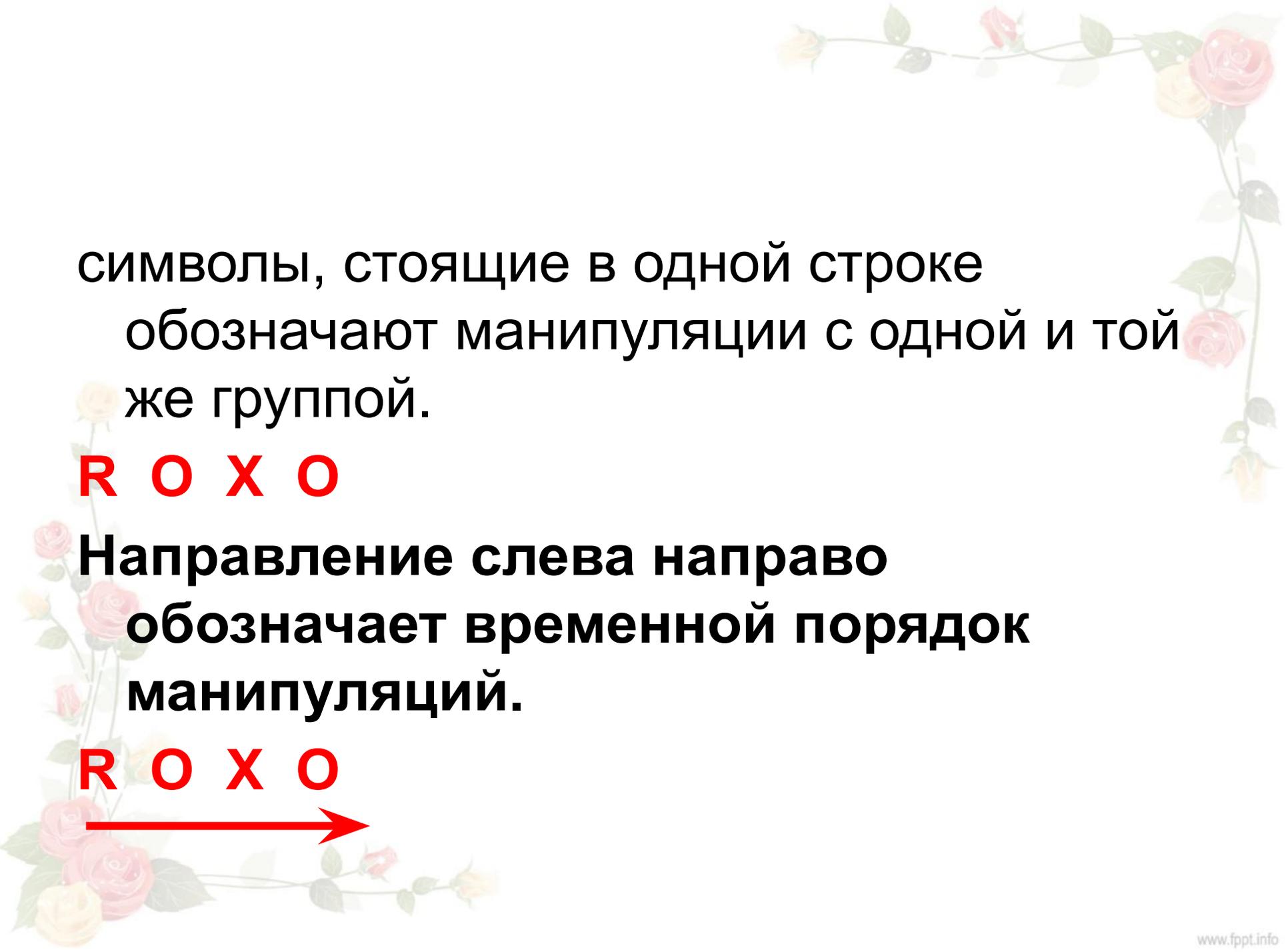
Истинноэкспериментальные

Квазиэкспериментальные

Доэкспериментальные



- 
- **X** (**eX**perimental) — экспериментальная переменная или событие, влияние которой/которого подлежит измерению
 - — **O** (**O**bservation) — процедура/процесс наблюдения, измерения, или результаты измерения;
 - — **R** (**R**andom) — случайный отбор испытуемых в группу;



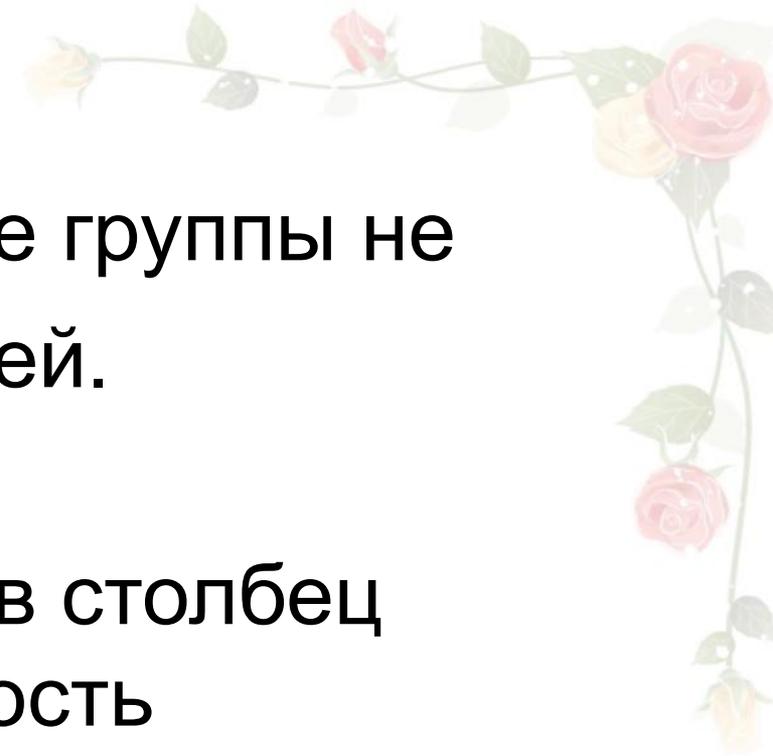
СИМВОЛЫ, СТОЯЩИЕ В ОДНОЙ СТРОКЕ
обозначают манипуляции с одной и той
же группой.

R O X O

Направление слева направо
обозначает временной порядок
манипуляций.

R O X O



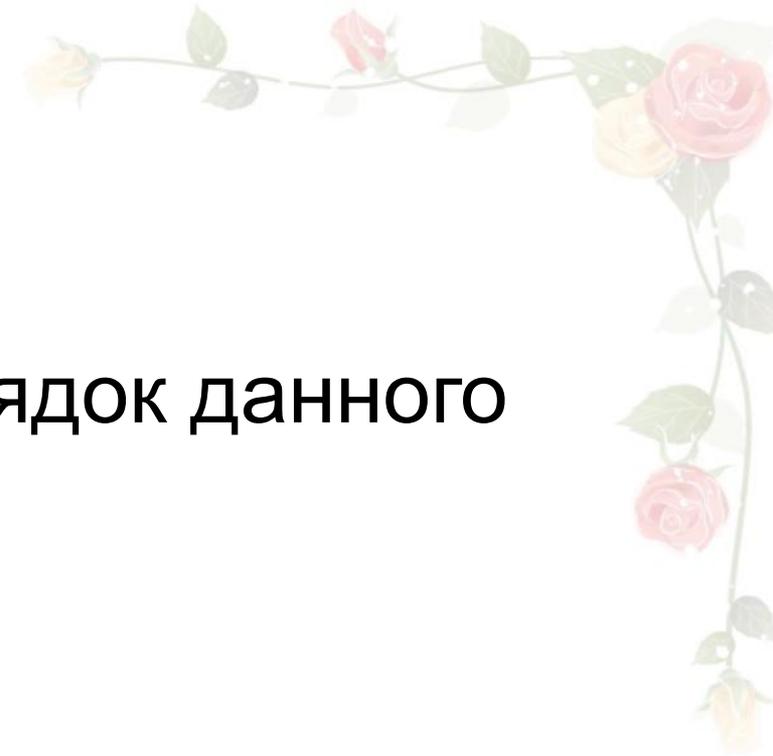


■■■■■■■■■■ сравнимые группы не
уравнены рандомизацией.

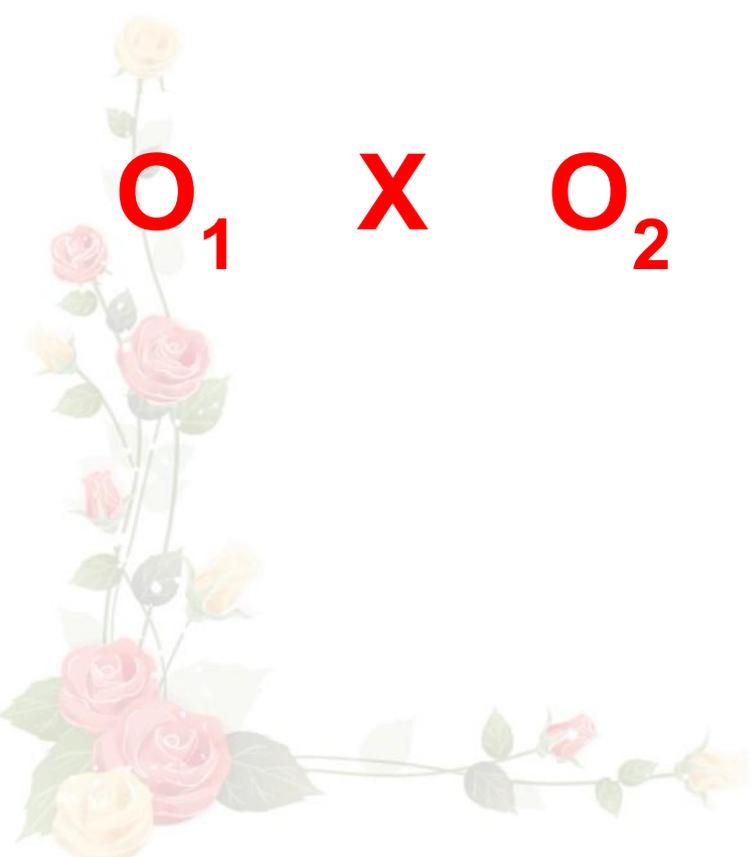
Расположение символов в столбец
означают одновременность
воздействий.

X O

O



Индексы обозначают порядок данного действия



O_1 X O_2

ДОЭКСПЕРИМЕНТАЛЬНЫЕ ПЛАНЫ



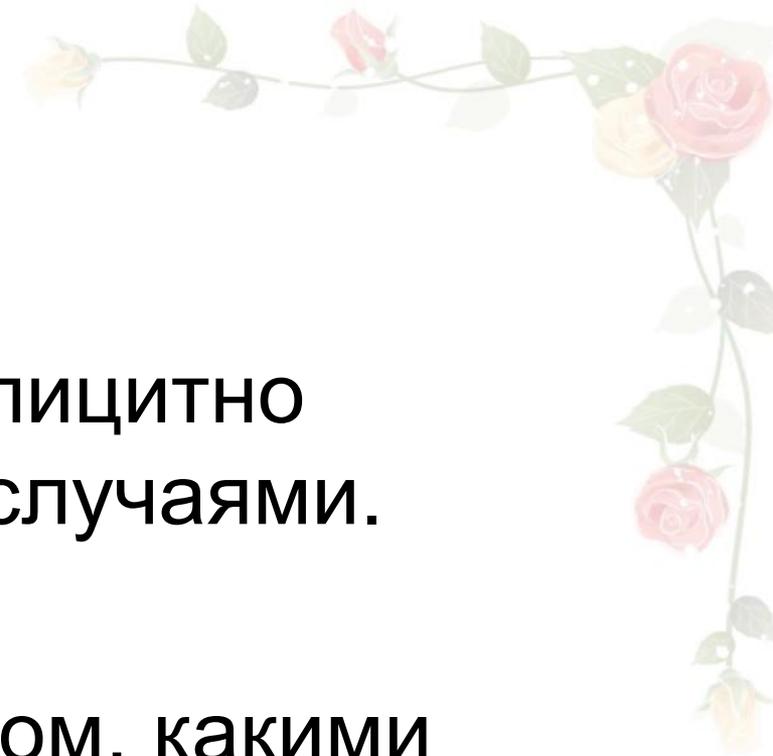
«Исследование одного случая»

X

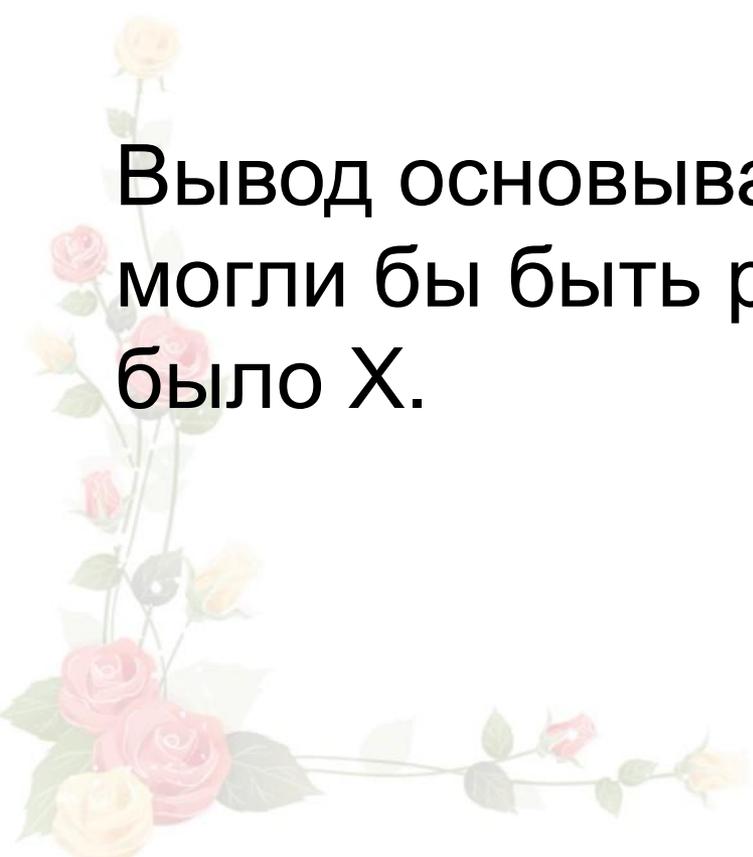
O

Нижняя точка отсчета

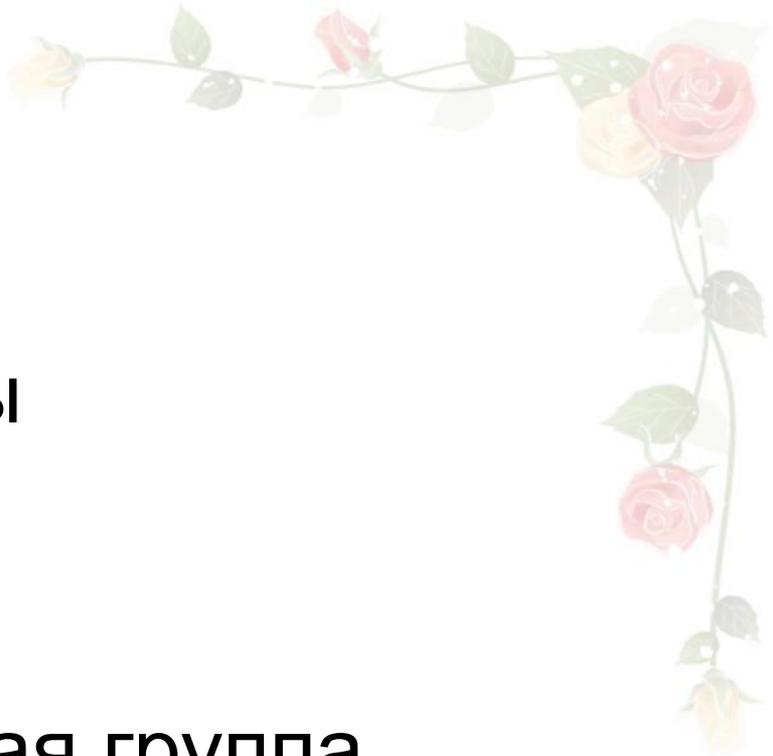




Полученные данные имплицитно
сравниваются с другими случаями.



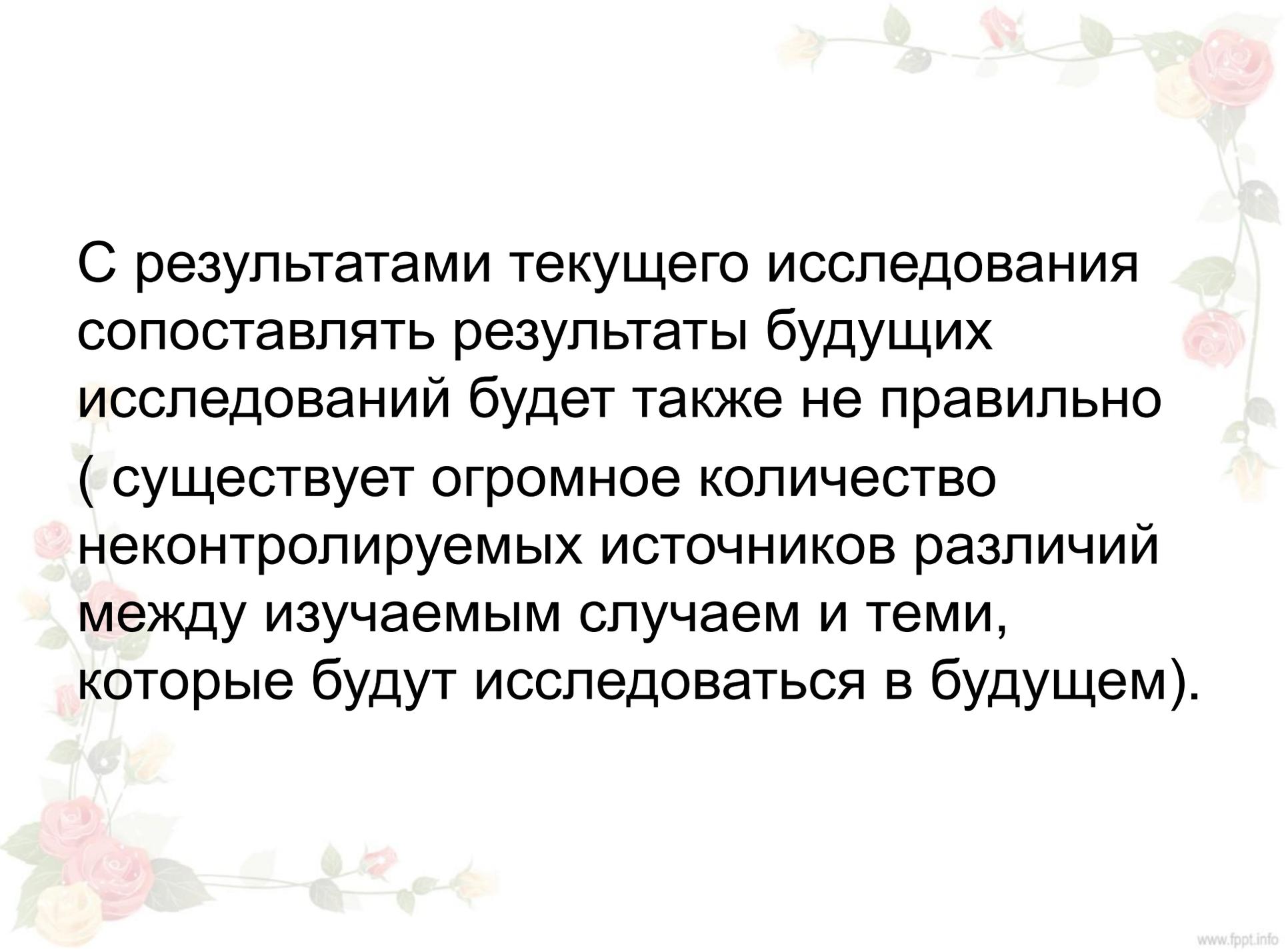
Вывод основывается на том, какими
могли бы быть результаты, если бы не
было X.



Стандартизованные тесты

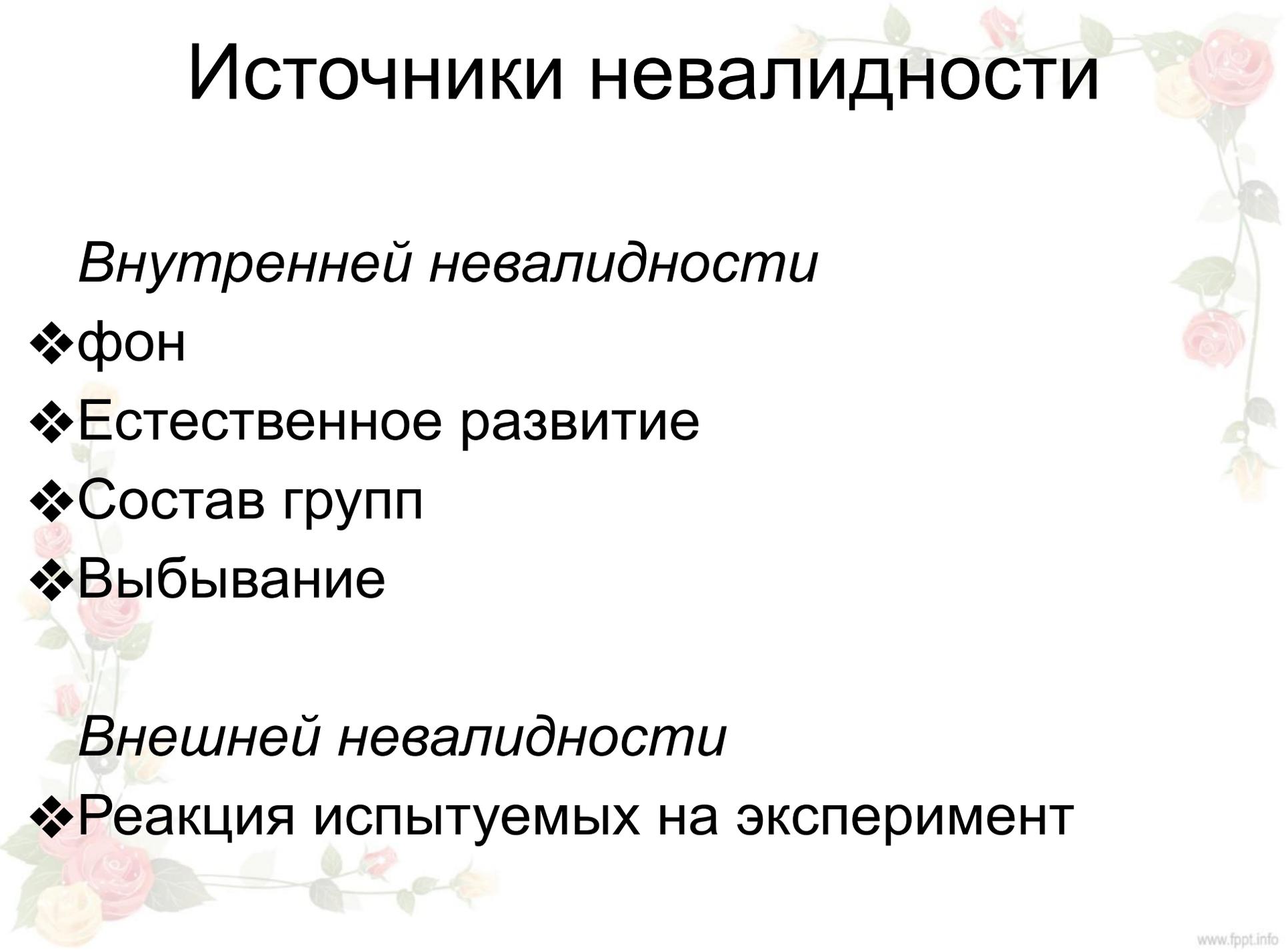


Стандартная - нормативная группа
бесполезна в качестве контрольной



С результатами текущего исследования сопоставлять результаты будущих исследований будет также не правильно (существует огромное количество неконтролируемых источников различий между изучаемым случаем и теми, которые будут исследоваться в будущем).

Источники невалидности



Внутренней невалидности

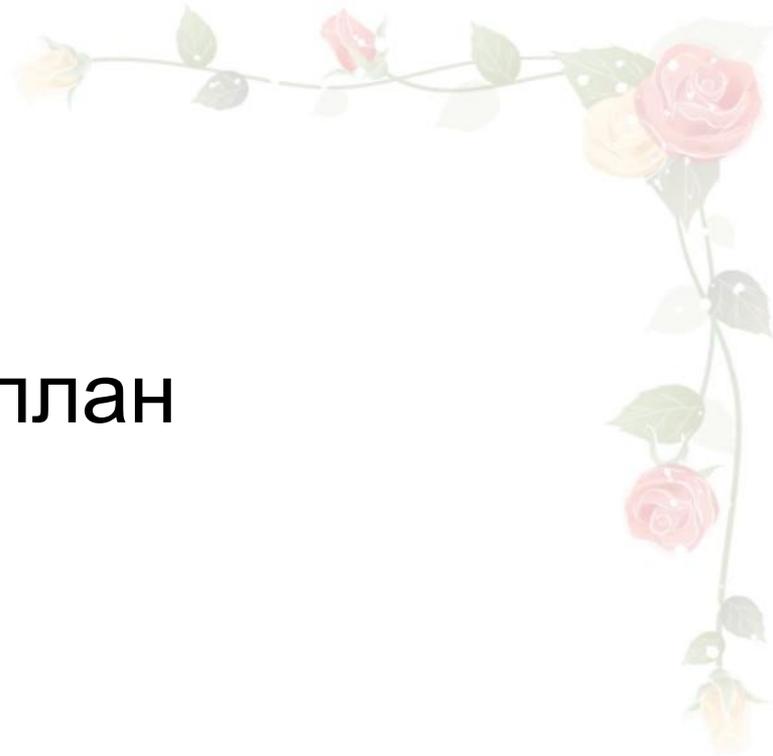
- ❖ фон
- ❖ Естественное развитие
- ❖ Состав групп
- ❖ Выбывание

Внешней невалидности

- ❖ Реакция испытуемых на эксперимент

Исследование на одной группе с
предварительным и итоговым
тестированием

O_1 X O_2



Лучше, чем предыдущий план



контролируется



состав групп



выбывание

Источники невалидности



Внутренней невалидности

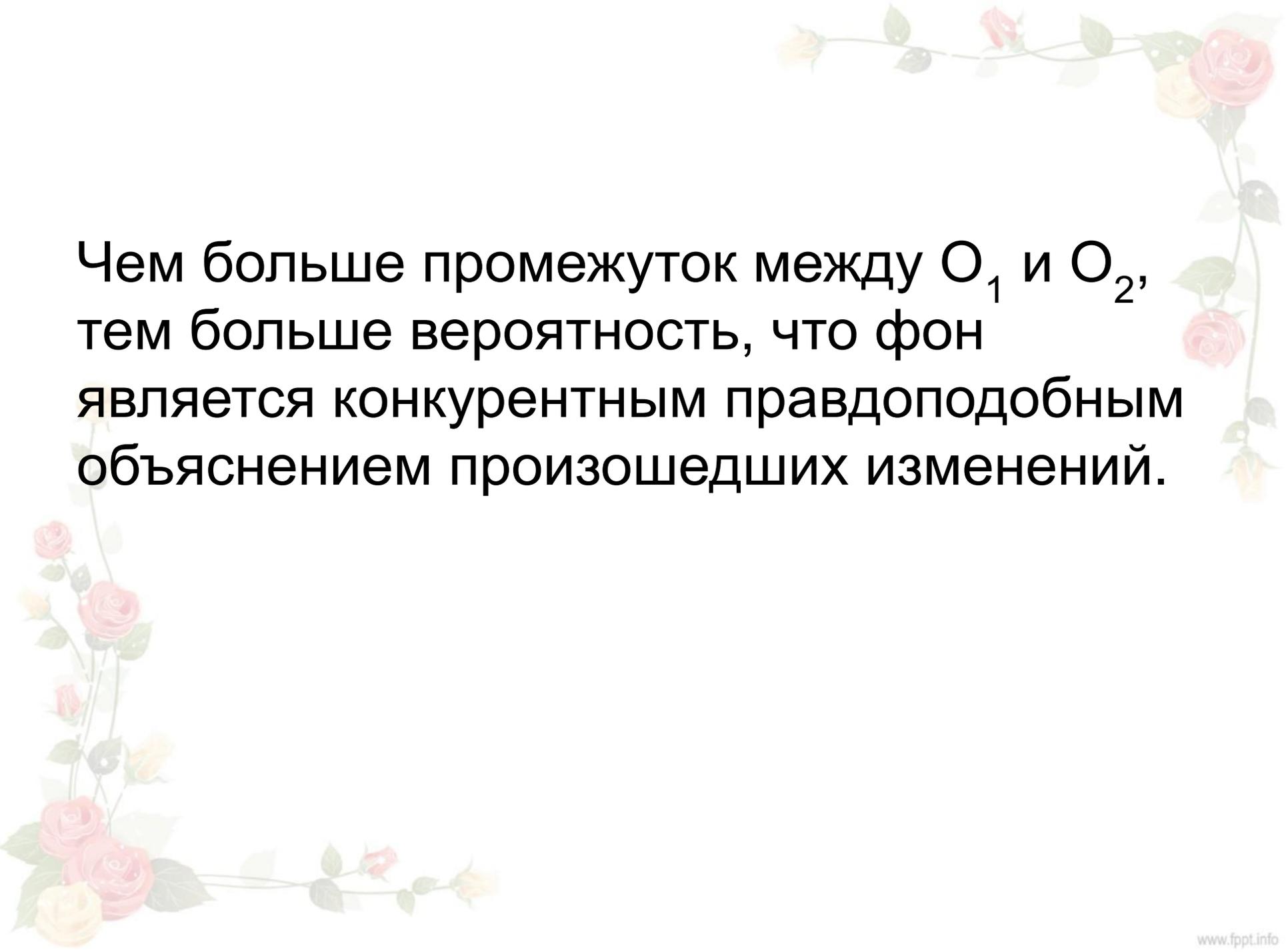
- ❖ фон
- ❖ Естественное развитие
- ❖ Эффект тестирования
- ❖ Инструментальная погрешность
- ❖ Состав групп
- ❖ Выбывание
- ❖ *Взаимодействие состава групп с естественным развитием*

фон

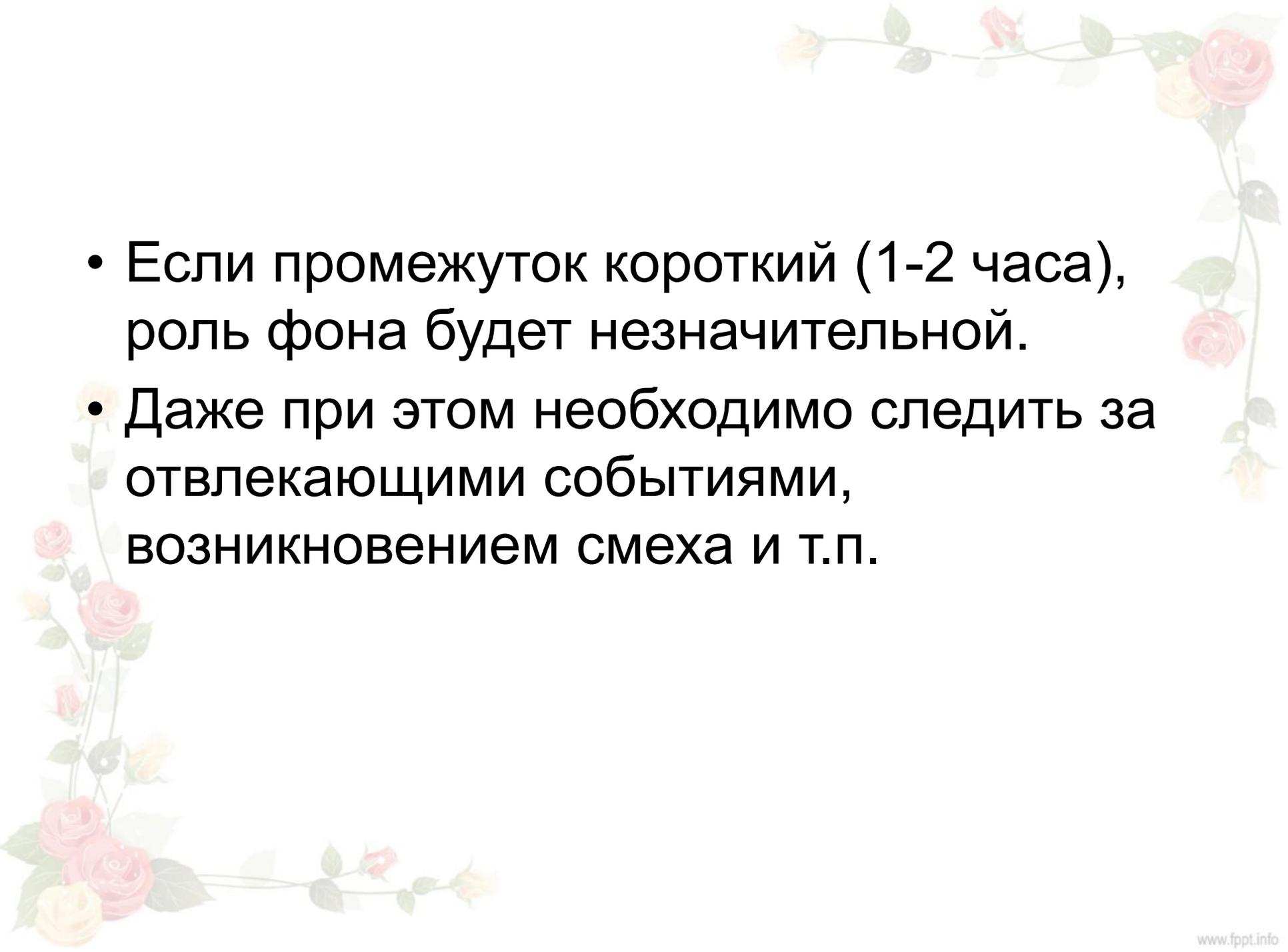
- Между O_1 (предварительное тестирование) и O_2 (итоговое тестирование) помимо экспериментального воздействия происходят другие события.
- если тестирования происходят в разные дни причиной различий могут послужить другие события, происшедшие в промежутке между тестированиями

Чтобы стать правдоподобной конкурентной гипотезой, такое событие должно произойти с большинством испытуемых (совместные занятия, известия)





Чем больше промежуток между O_1 и O_2 ,
тем больше вероятность, что фон
является конкурентным правдоподобным
объяснением произошедших изменений.

- 
- Если промежуток короткий (1-2 часа), роль фона будет незначительной.
 - Даже при этом необходимо следить за отвлекающими событиями, возникновением смеха и т.п.

«Экспериментальная изоляция»



К фону можно отнести время года,
приближение сессии (можно отнести и к
естественному развитию) и т.д.

Влияние этих событий может быть ошибочно принято за действие x .

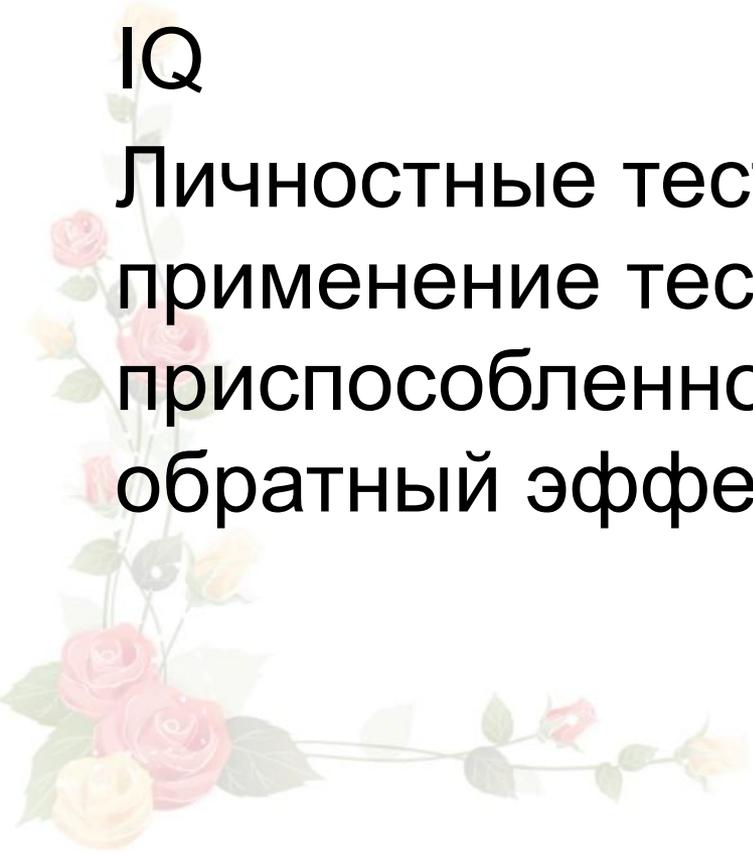
Естественное развитие

- Полученные различия могут оказаться следствием усталости, скуки, голода ..., а не действия х.
- «Спонтанная ремиссия»

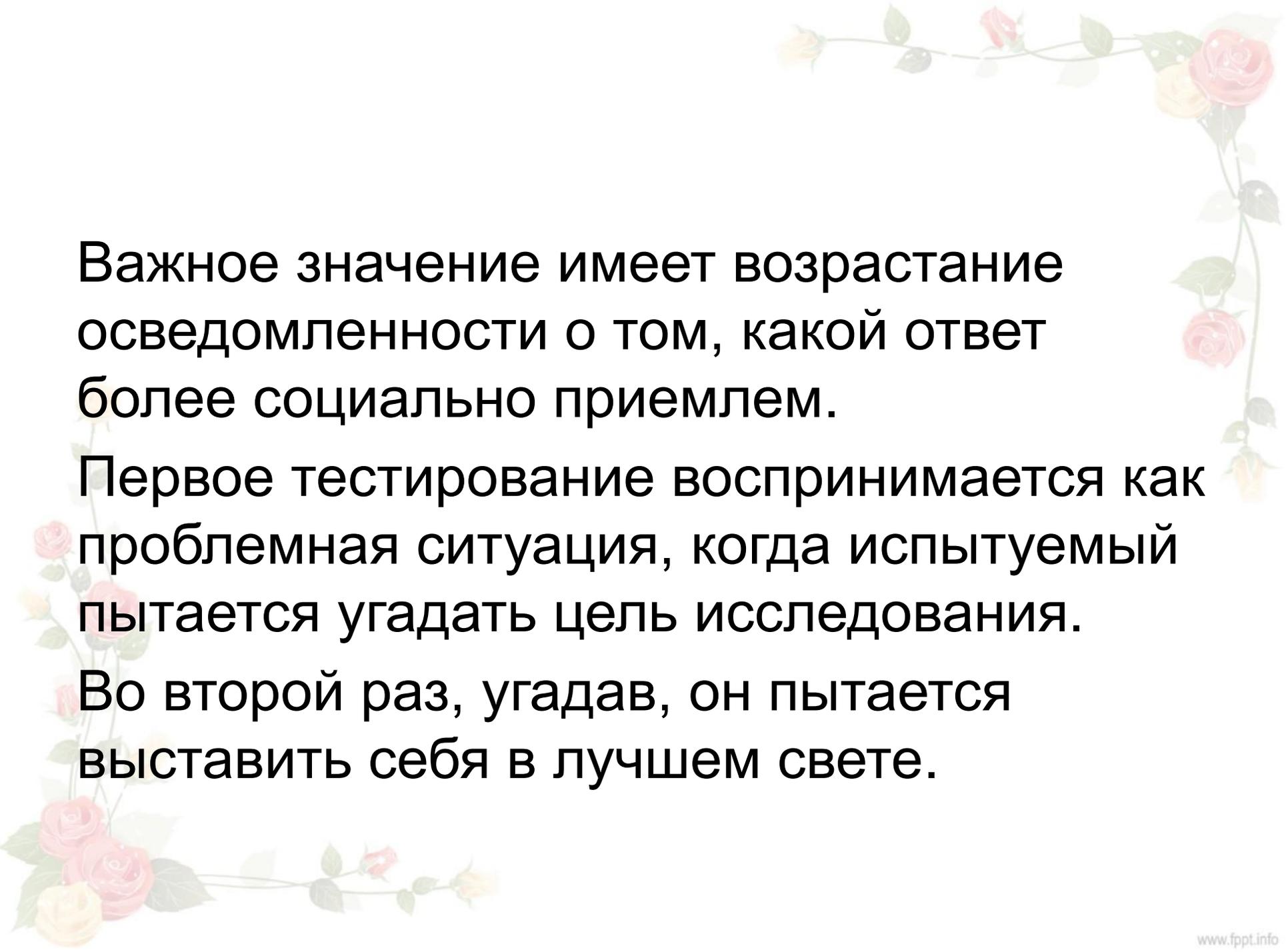


Эффекты тестирования

IQ



Личностные тесты (повторное применение тестов показывает лучшую приспособленность, но может быть и обратный эффект)



Важное значение имеет возрастание осведомленности о том, какой ответ более социально приемлем.

Первое тестирование воспринимается как проблемная ситуация, когда испытуемый пытается угадать цель исследования.

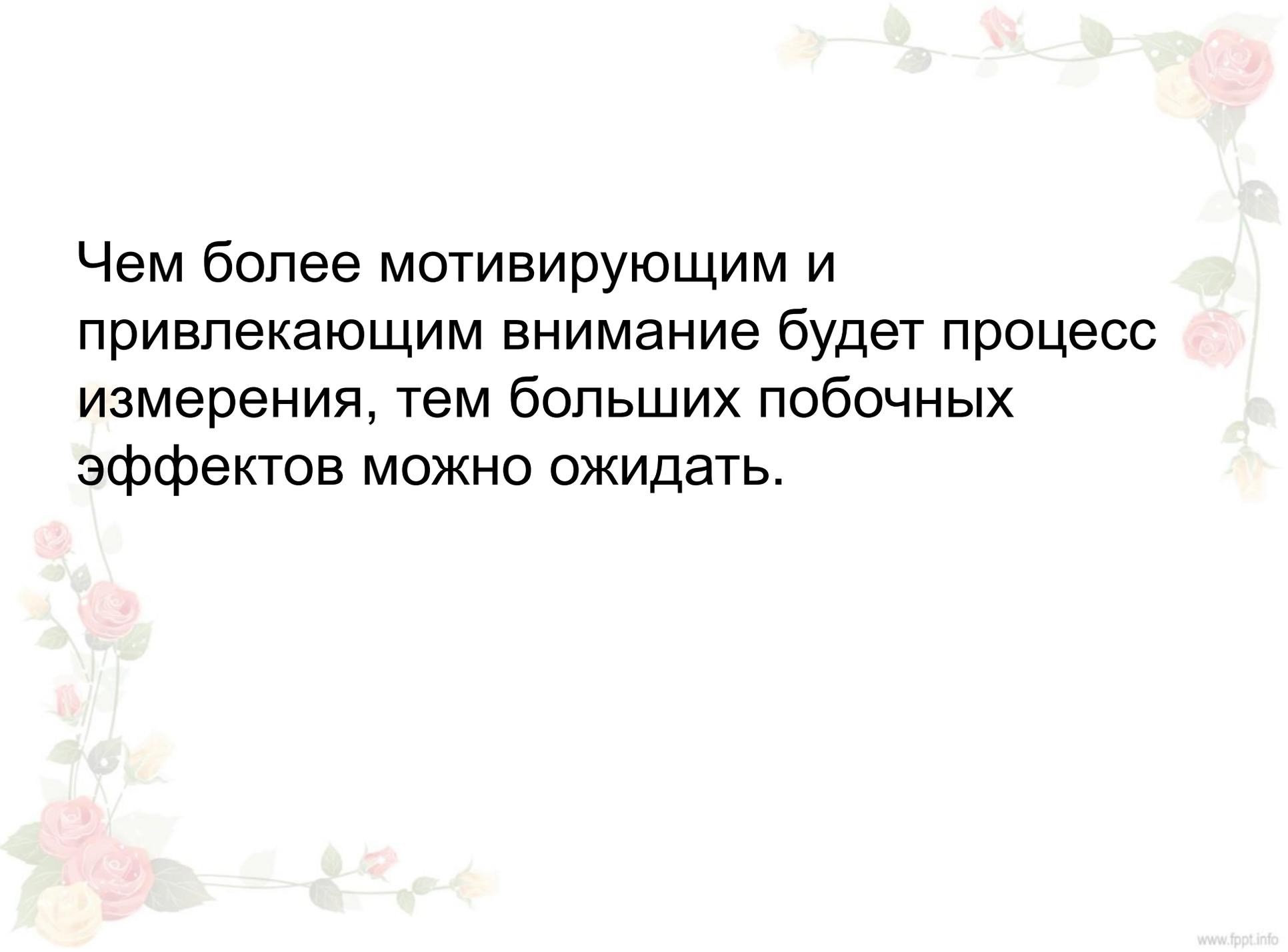
Во второй раз, угадав, он пытается выставить себя в лучшем свете.

Реактивность испытуемых

- Сам процесс измерения приводит к изменениям.
- Реактивность может возникать тогда, когда тестирование выступает как стимул к изменениям.

Появление наблюдателя уже меняет поведение.





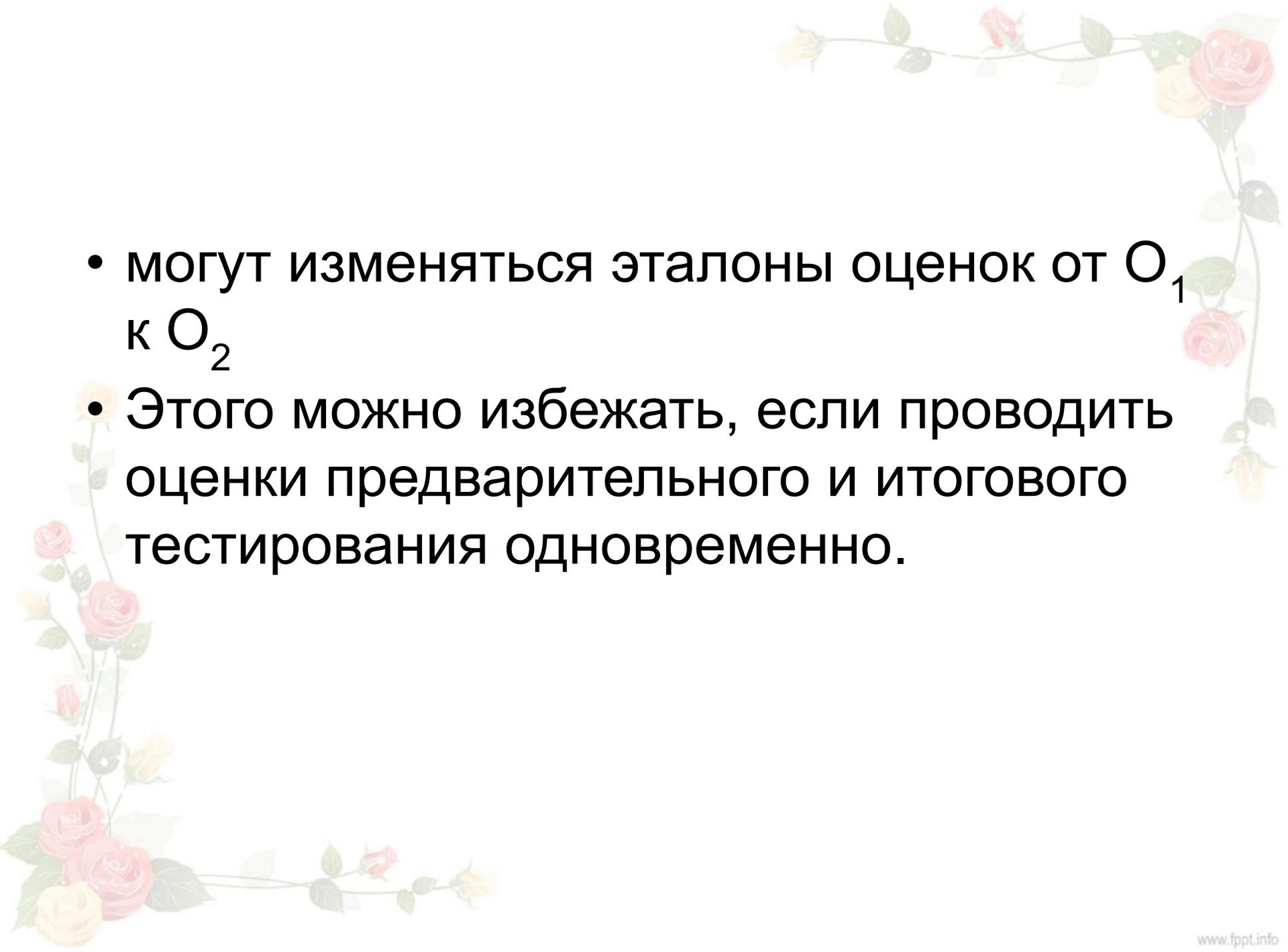
Чем более мотивирующим и привлекающим внимание будет процесс измерения, тем больших побочных эффектов можно ожидать.

Инструментальная погрешность

Усталость, научение экспериментатора, лучшее знакомство с порядком проведения исследования приведут к различиям между O_1 и O_2

Экспериментатор может стать более опытным или более пресыщенным.



- 
- могут изменяться эталоны оценок от O_1 к O_2
 - Этого можно избежать, если проводить оценки предварительного и итогового тестирования одновременно.

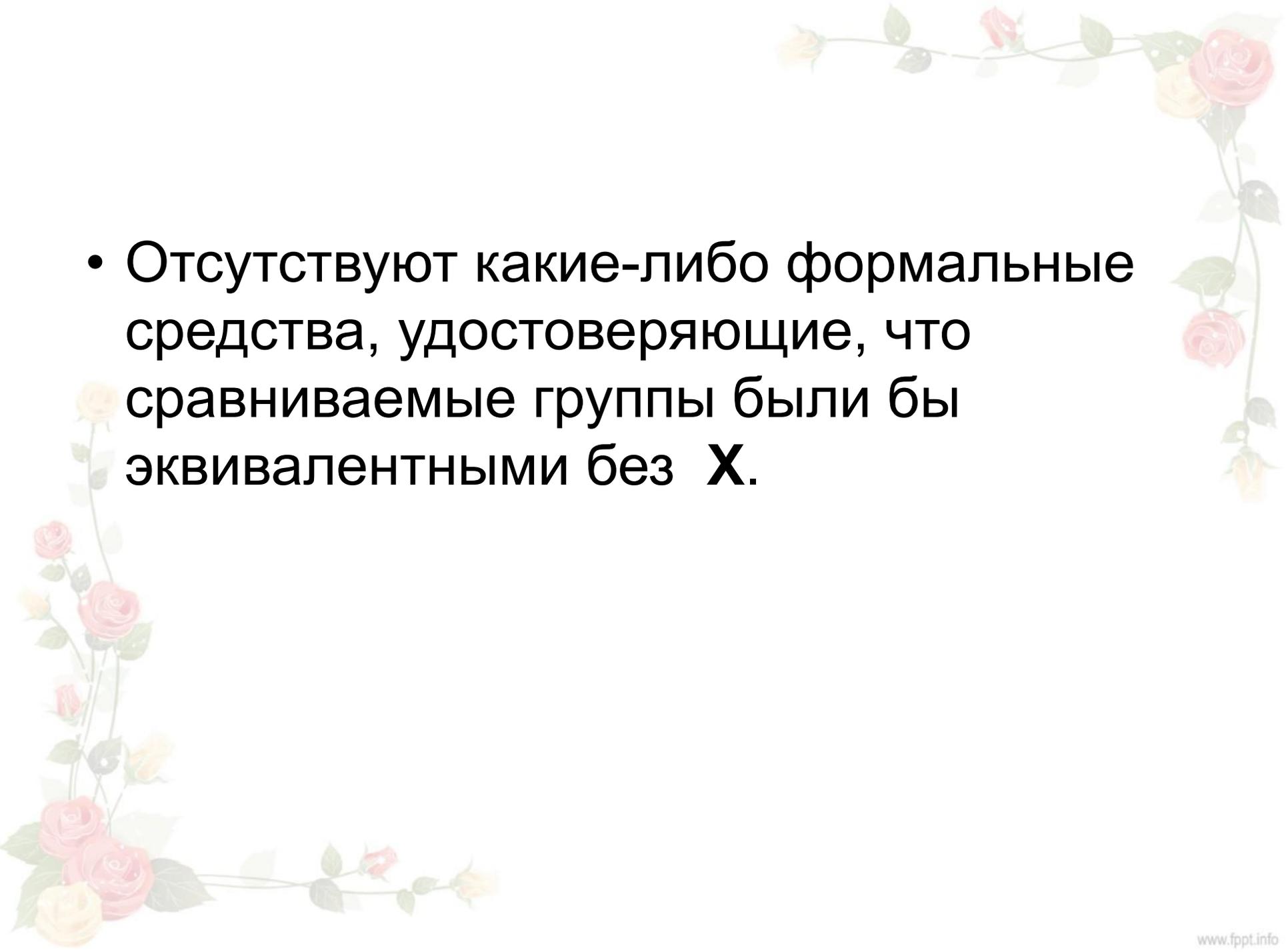
Внешней невалидности

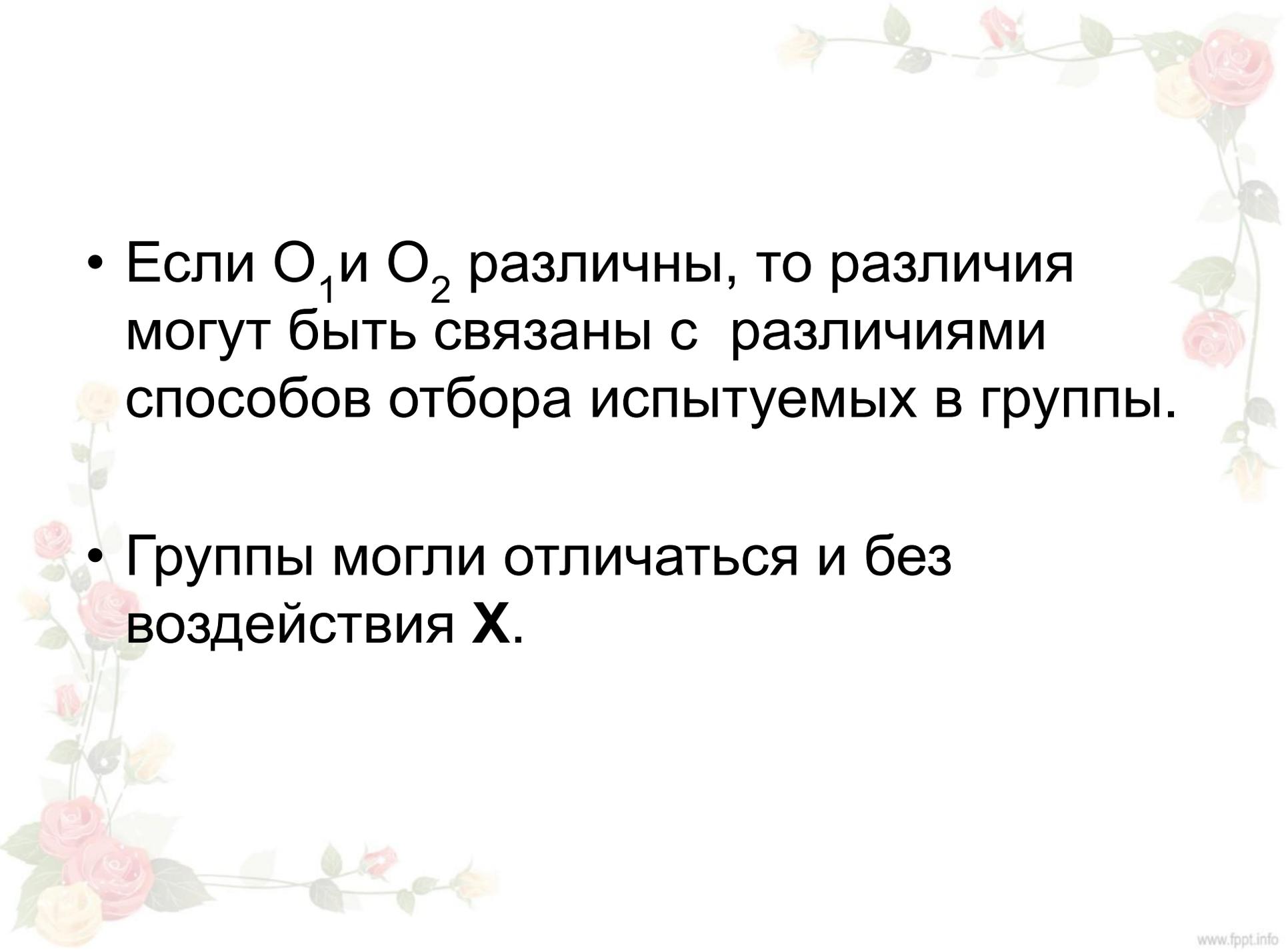
- ❖ взаимодействие тестирования и x
- ❖ Взаимодействие состава групп и x

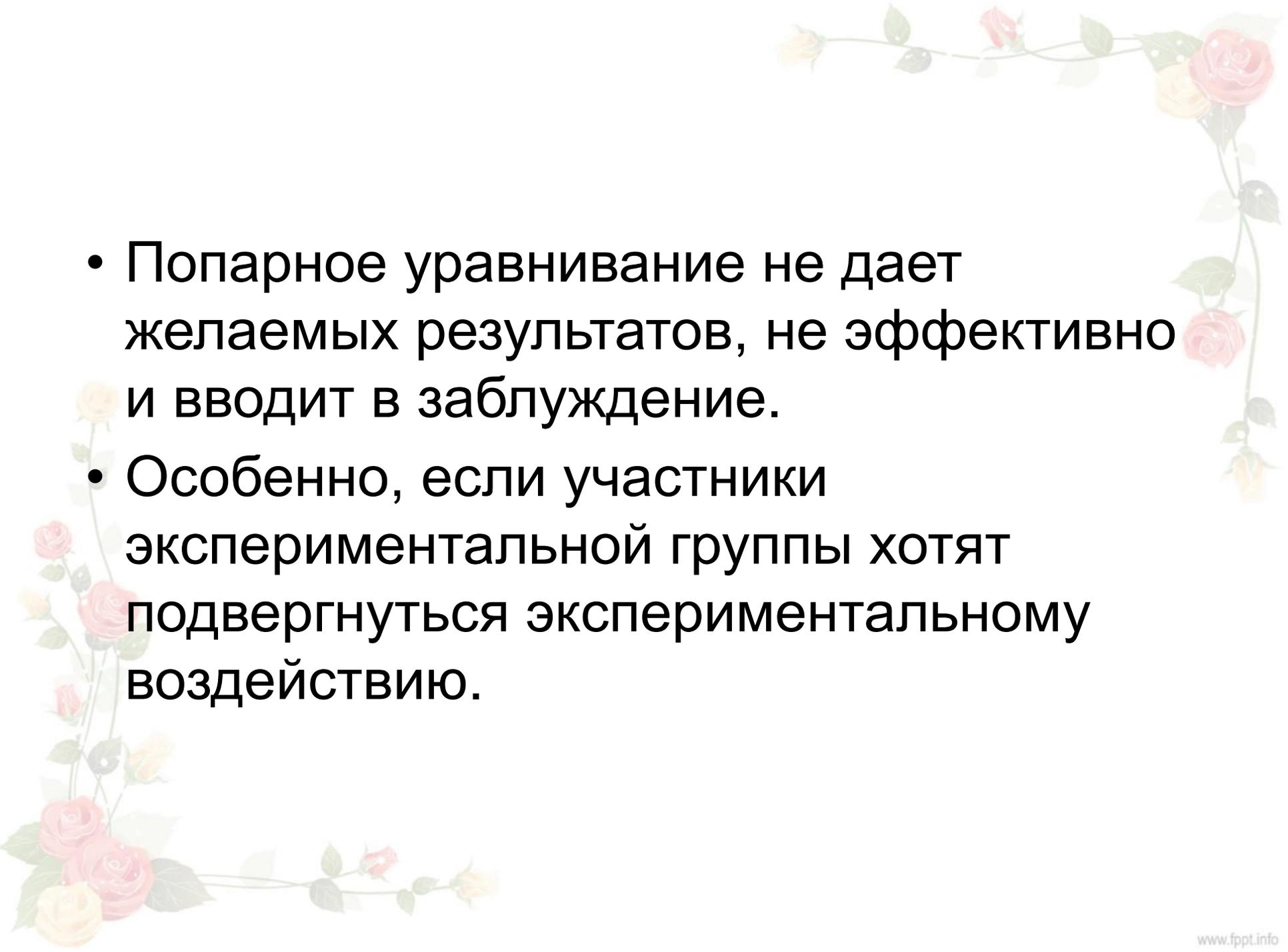
Сравнение статистических групп

X O_1

O_2

- 
- Отсутствуют какие-либо формальные средства, удостоверяющие, что сравниваемые группы были бы эквивалентными без **X**.

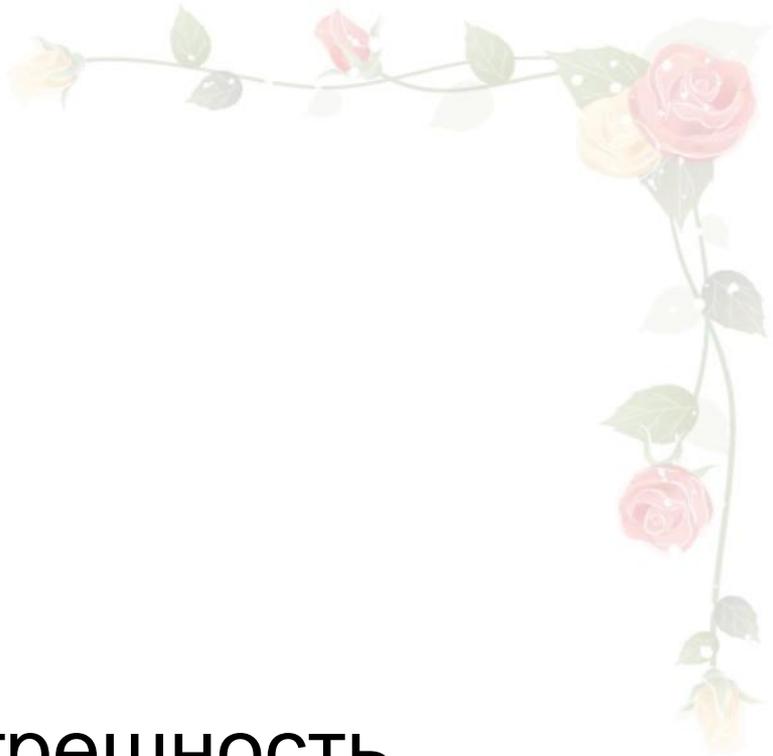
- 
- Если O_1 и O_2 различны, то различия могут быть связаны с различиями способов отбора испытуемых в группы.
 - Группы могли отличаться и без воздействия **X**.

- 
- Попарное уравнивание не дает желаемых результатов, не эффективно и вводит в заблуждение.
 - Особенно, если участники экспериментальной группы хотят подвергнуться экспериментальному воздействию.

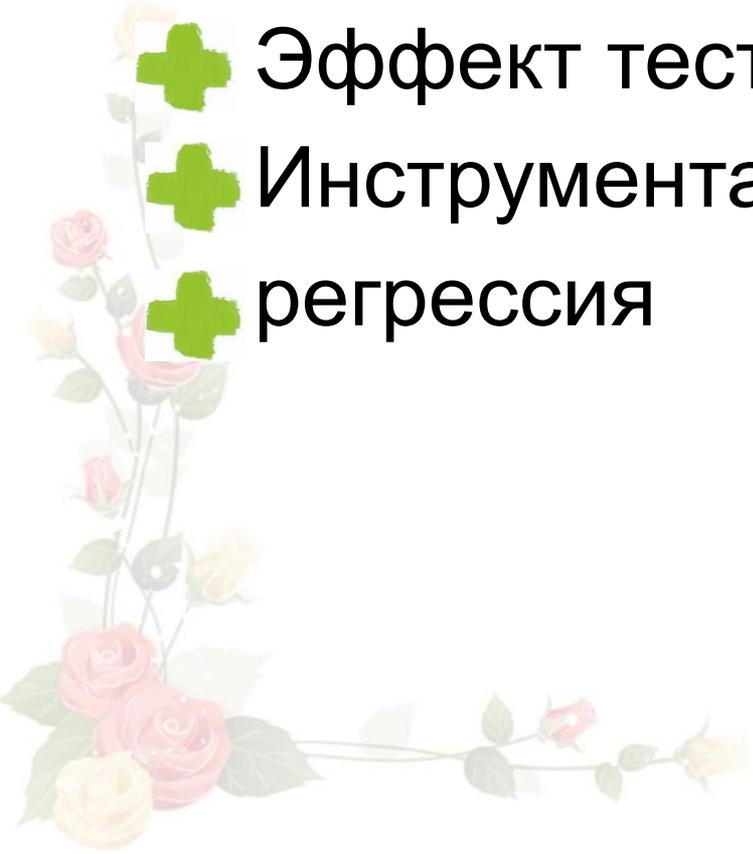
выбывание

Различия $O_1 - O_2$ в группах в результате различного выбывания испытуемых из сравниваемых групп.

Даже, если когда-то группы были эквиваленты, к моменту проведения сравнения они могут отличаться не по причине изменений, произошедших с членами групп, а по причине селективного выбывания испытуемых из одной группы.

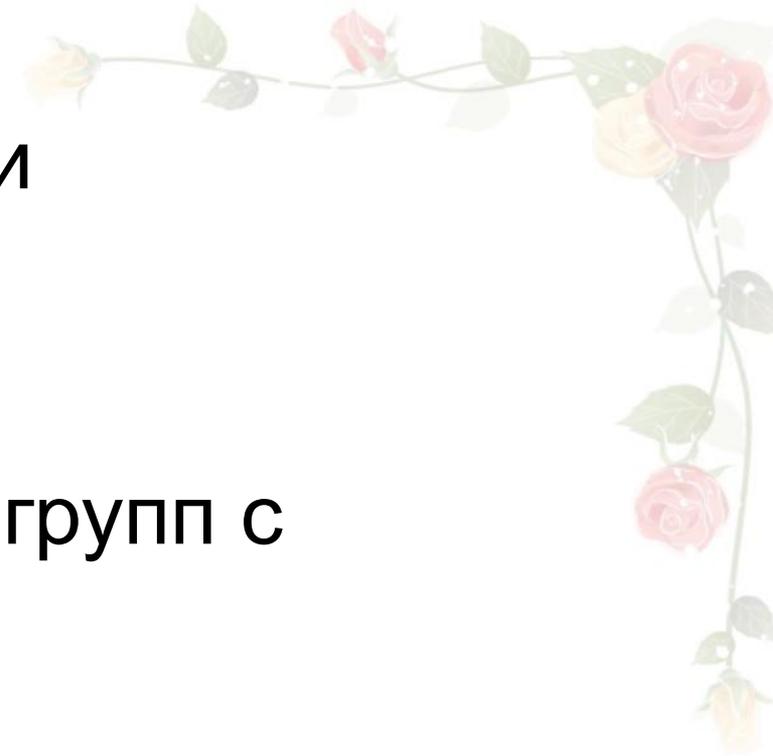


контролируется

-  Фон
 -  Эффект тестирования
 -  Инструментальная погрешность
 -  регрессия
- 

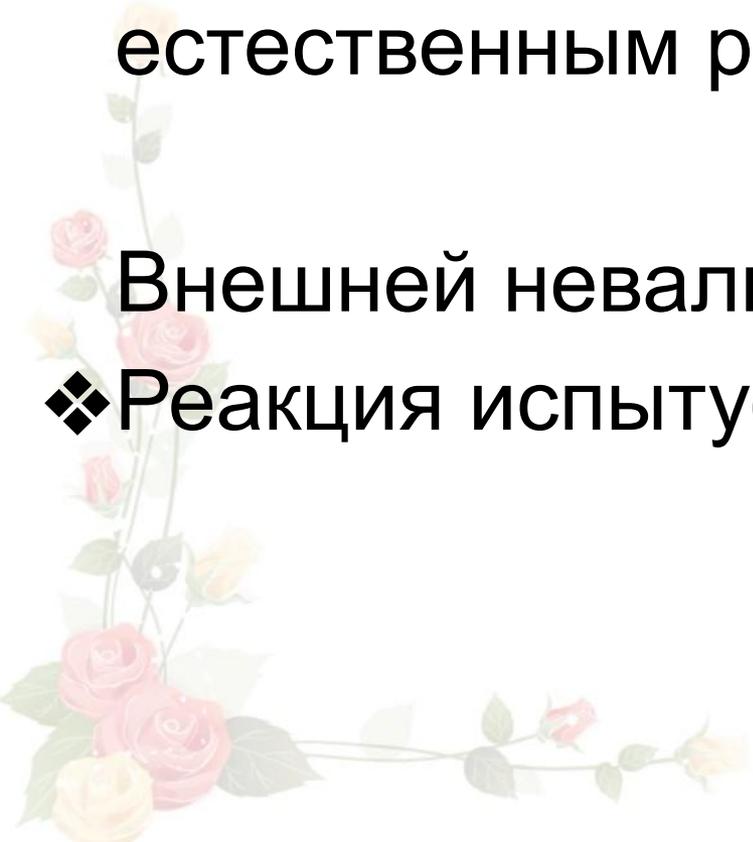
Источники невалидности





Внутренней невалидности

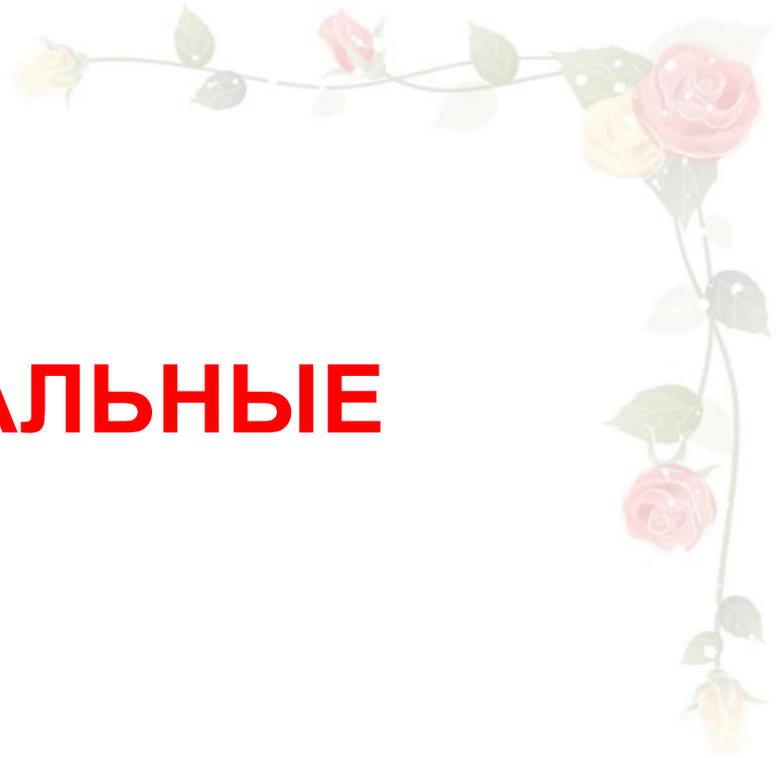
- ❖ Состав групп
- ❖ Выбывание
- ❖ Взаимодействие состава групп с естественным развитием



Внешней невалидности

- ❖ Реакция испытуемых на эксперимент

КВАЗИЭКСПЕРИМЕНТАЛЬНЫЕ ПЛАНЫ



План с предварительными и итоговым тестированием (неэквивалентные группы)

O_1 X O_2



O_1 O_2

План с контрольными выборками для предварительного и итогового тестирования

R O (X)

R X O

R O

R O

Планы временных серий

- $O_1 O_2 O_3 O_4 X O_5 O_6 O_7 O_8$

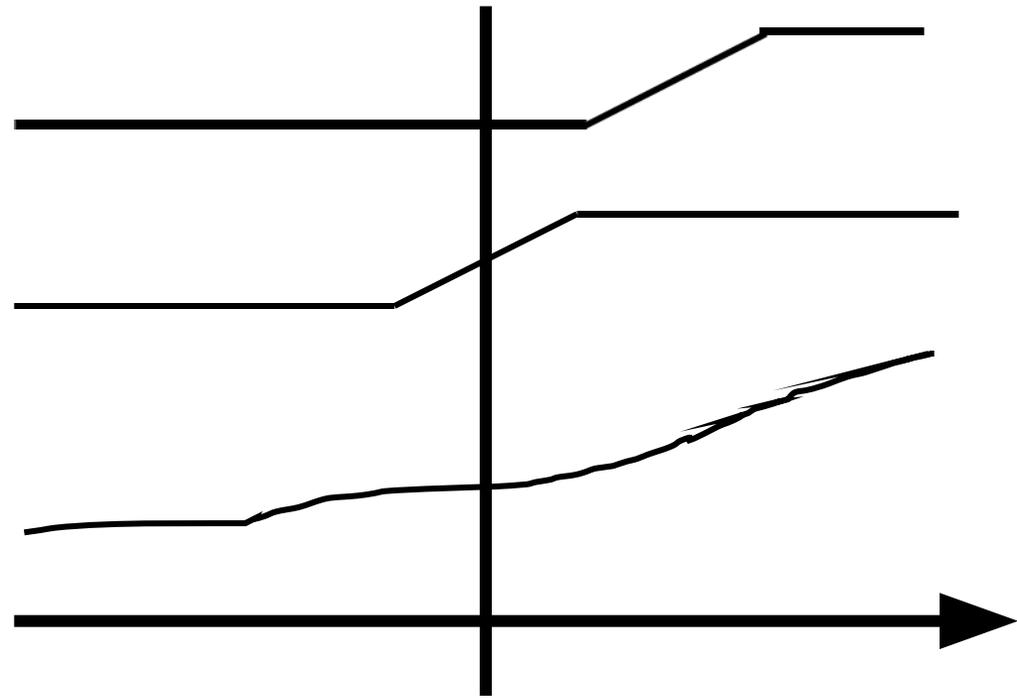
Многократные замеры зависимой переменной как до, так и после воздействия

- До воздействия пассивное состояние независимой переменной,
- X – активное воздействие.
- Изменение зависимой переменной – нарушение временного тренда.



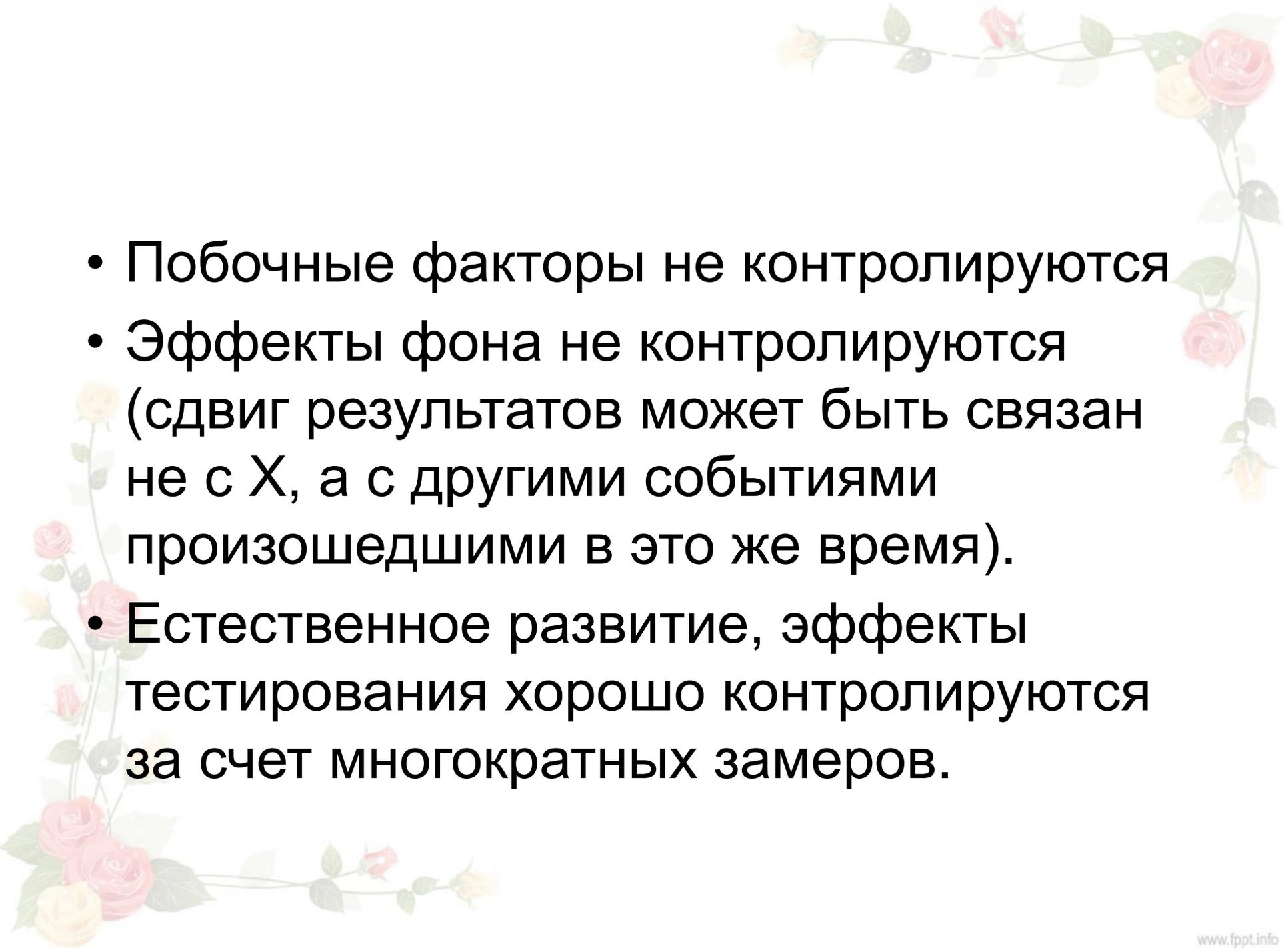
Время выступает аналогом независимой переменной при изучении временного тренда.

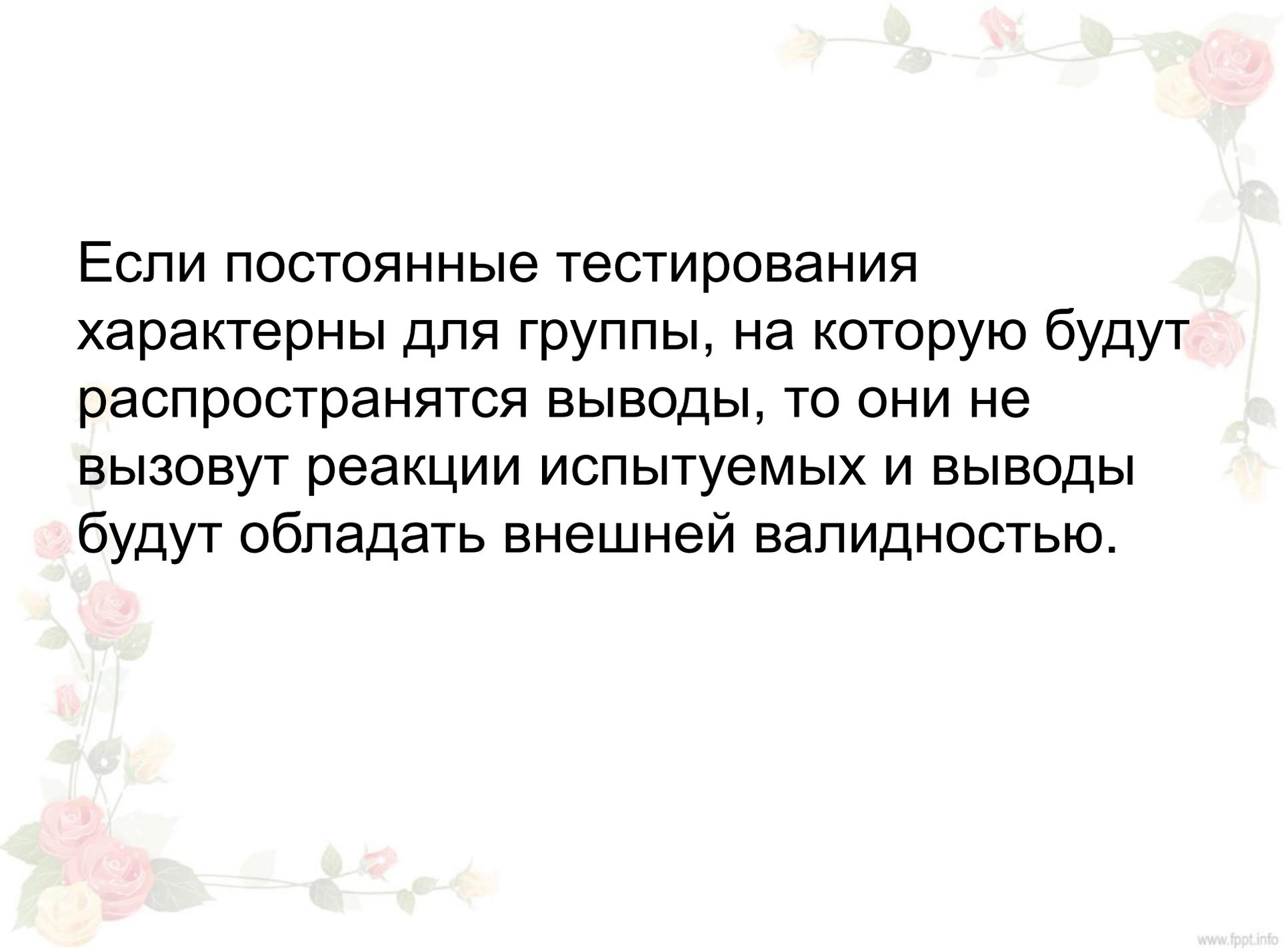
Отсутствует контроль над уровнями независимой переменной с использованием разных групп.



O_1 O_2 O_3 O_4 **X** O_5 O_6 O_7 O_8

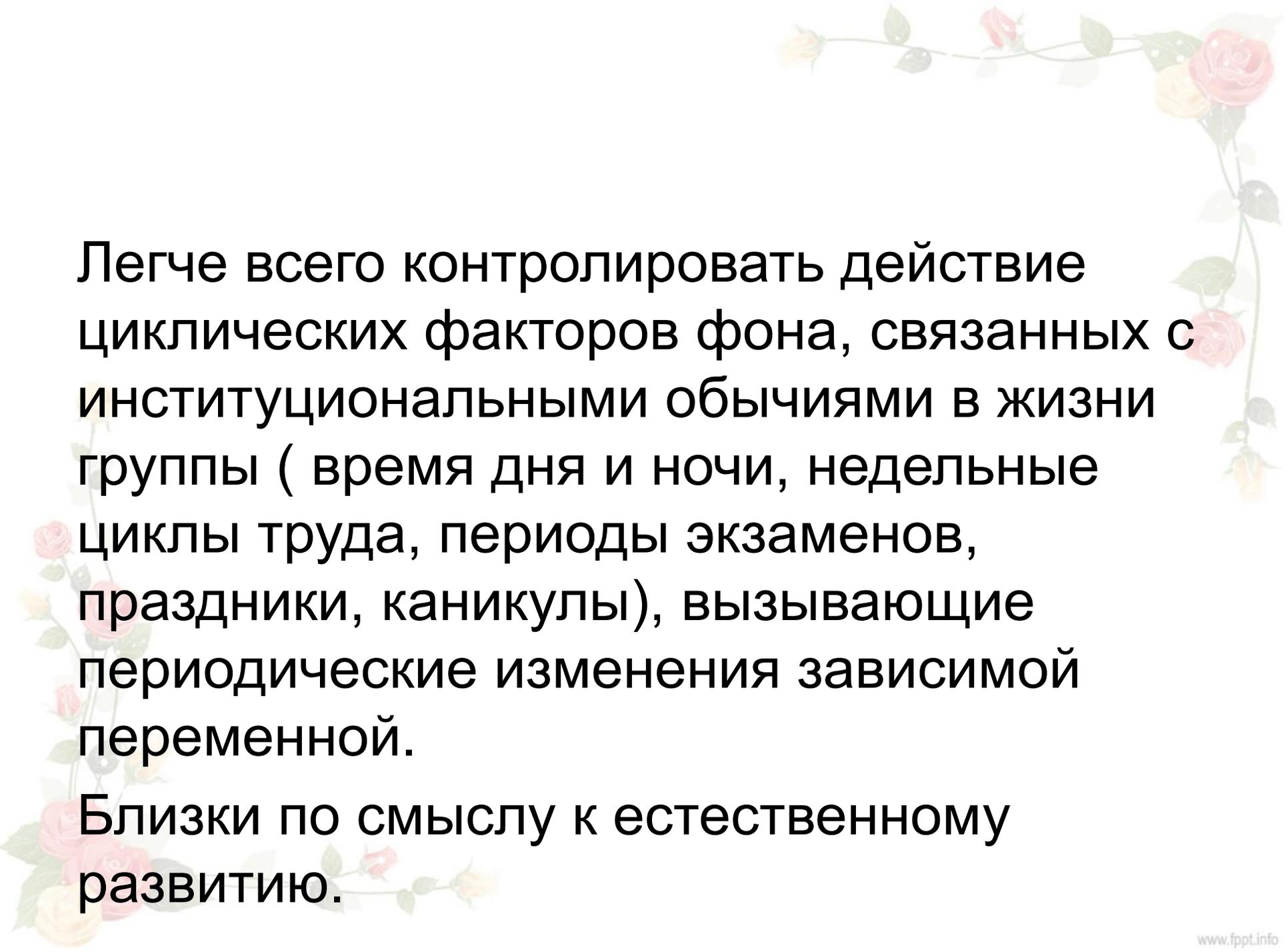
Необходимо выяснить, через какой интервал времени должен наступать эффект от X....

- 
- Побочные факторы не контролируются
 - Эффекты фона не контролируются (сдвиг результатов может быть связан не с X , а с другими событиями произошедшими в это же время).
 - Естественное развитие, эффекты тестирования хорошо контролируются за счет многократных замеров.



Если постоянные тестирования характерны для группы, на которую будут распространяться выводы, то они не вызовут реакции испытуемых и выводы будут обладать внешней валидностью.

- Вероятность фоновых воздействий зависит от экспериментальной изоляции.



Легче всего контролировать действие циклических факторов фона, связанных с институциональными обычаями в жизни группы (время дня и ночи, недельные циклы труда, периоды экзаменов, праздники, каникулы), вызывающие периодические изменения зависимой переменной.

Близки по смыслу к естественному развитию.

Фактор инструментальной погрешности

- Различия в измерительных процедурах могут быть восприняты как X изменения.
- Если экспериментатор знает о гипотезе исследования

Серии эквивалентных воздействий

$M_a X_1 O$ $M_b X_0 O$ $M_c X_1 O$ $M_d X_0 O$ и т.д.

МНОЖЕСТВЕННЫЕ СЕРИИ ЗАМЕРОВ

0 0 0 X 0 0 0



0 0 0 0 0 0

СЕРИИ ВРЕМЕННЫХ ВЫБОРОК

$X_1 O$ $X_0 O$ $X_1 O$ $X_0 O$ и т.д.

Предварительное и итоговое тестирование на различных выборках

R O (X)

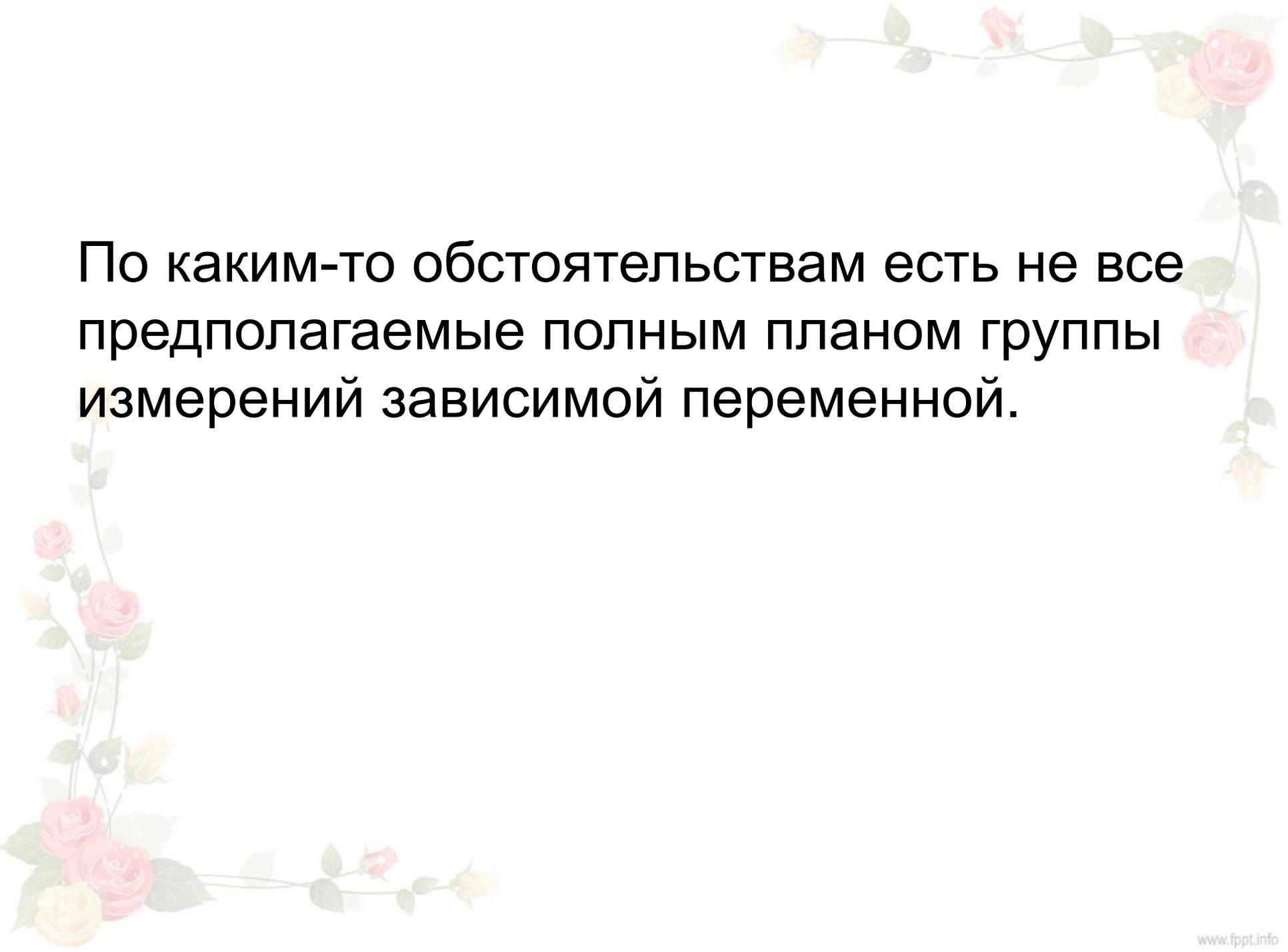
R X O

Лоскутные планы (неполные)

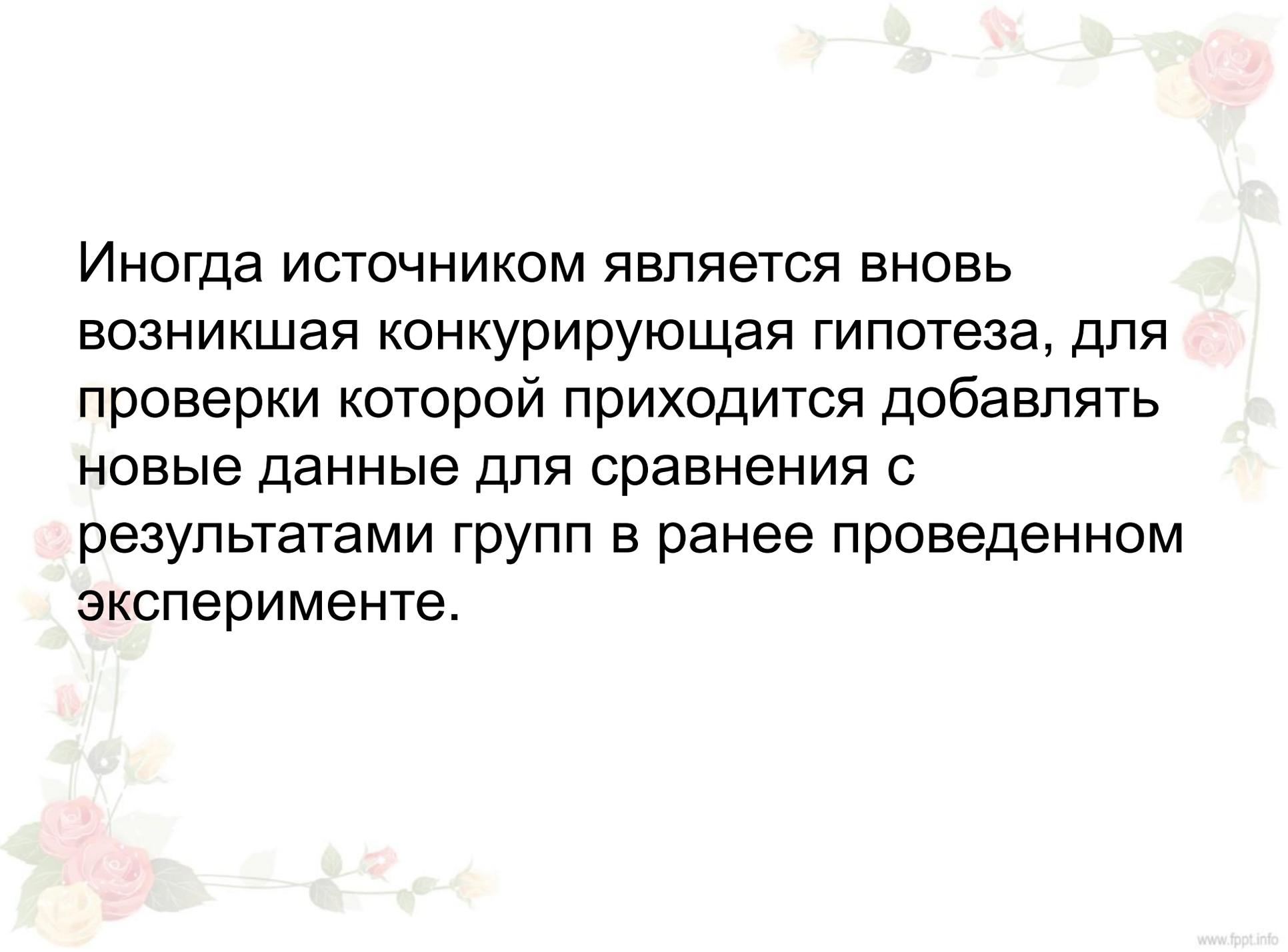


Когда отсутствуют некоторые группы для сравнения

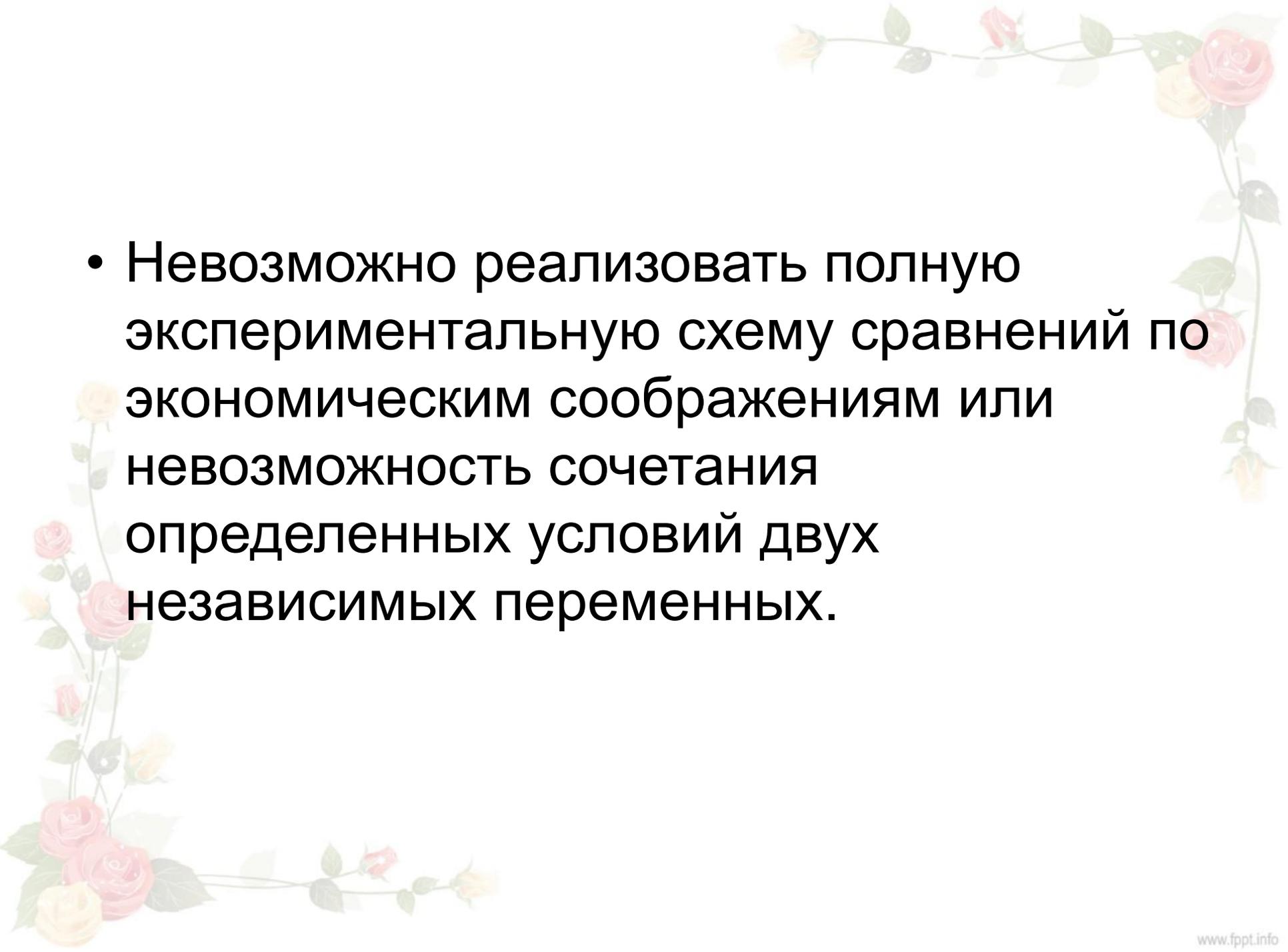


The slide features decorative rose vines in the corners. In the top right, a vine with green leaves and several roses in shades of pink, red, and yellow extends from the edge towards the center. In the bottom left, a similar vine with more roses and buds curves upwards from the corner. The background is plain white.

По каким-то обстоятельствам есть не все предполагаемые полным планом группы измерений зависимой переменной.



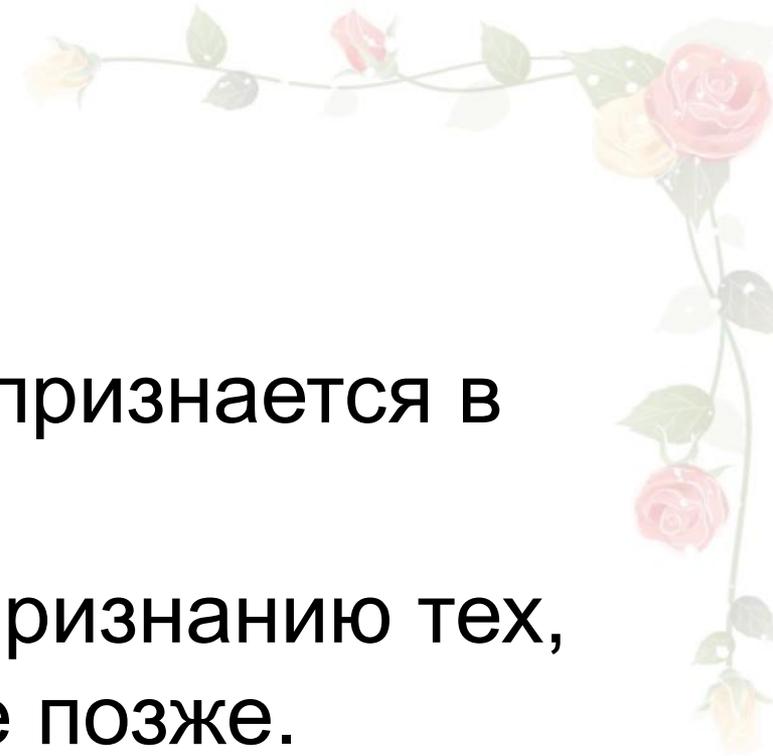
Иногда источником является вновь возникшая конкурирующая гипотеза, для проверки которой приходится добавлять новые данные для сравнения с результатами групп в ранее проведенном эксперименте.

- 
- Невозможно реализовать полную экспериментальную схему сравнений по экономическим соображениям или невозможность сочетания определенных условий двух независимых переменных.

Эффект Матфея

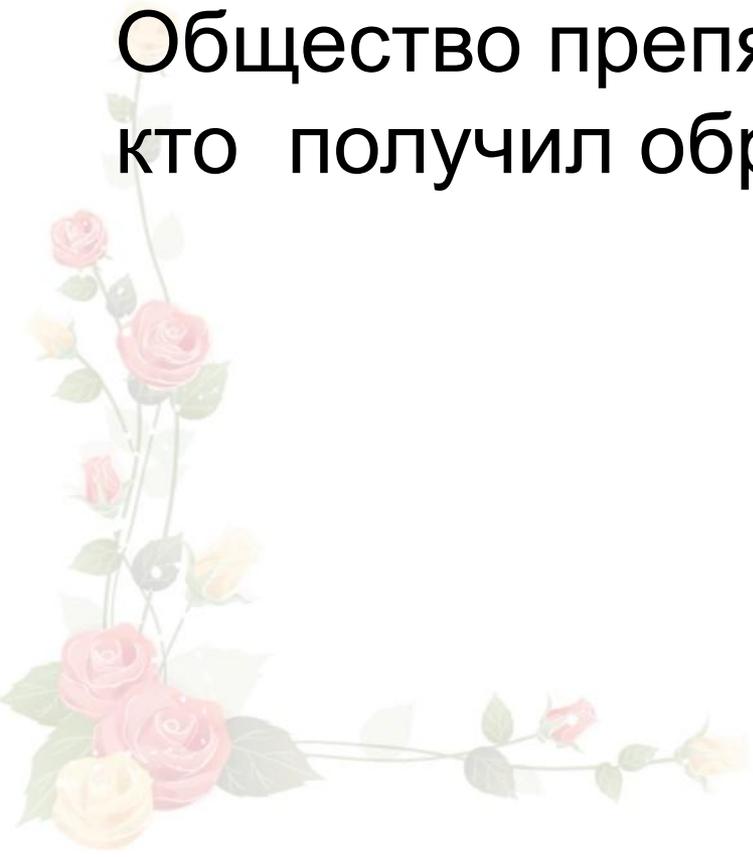
«имеющему да прибавится, а у
неимущего да отымится»

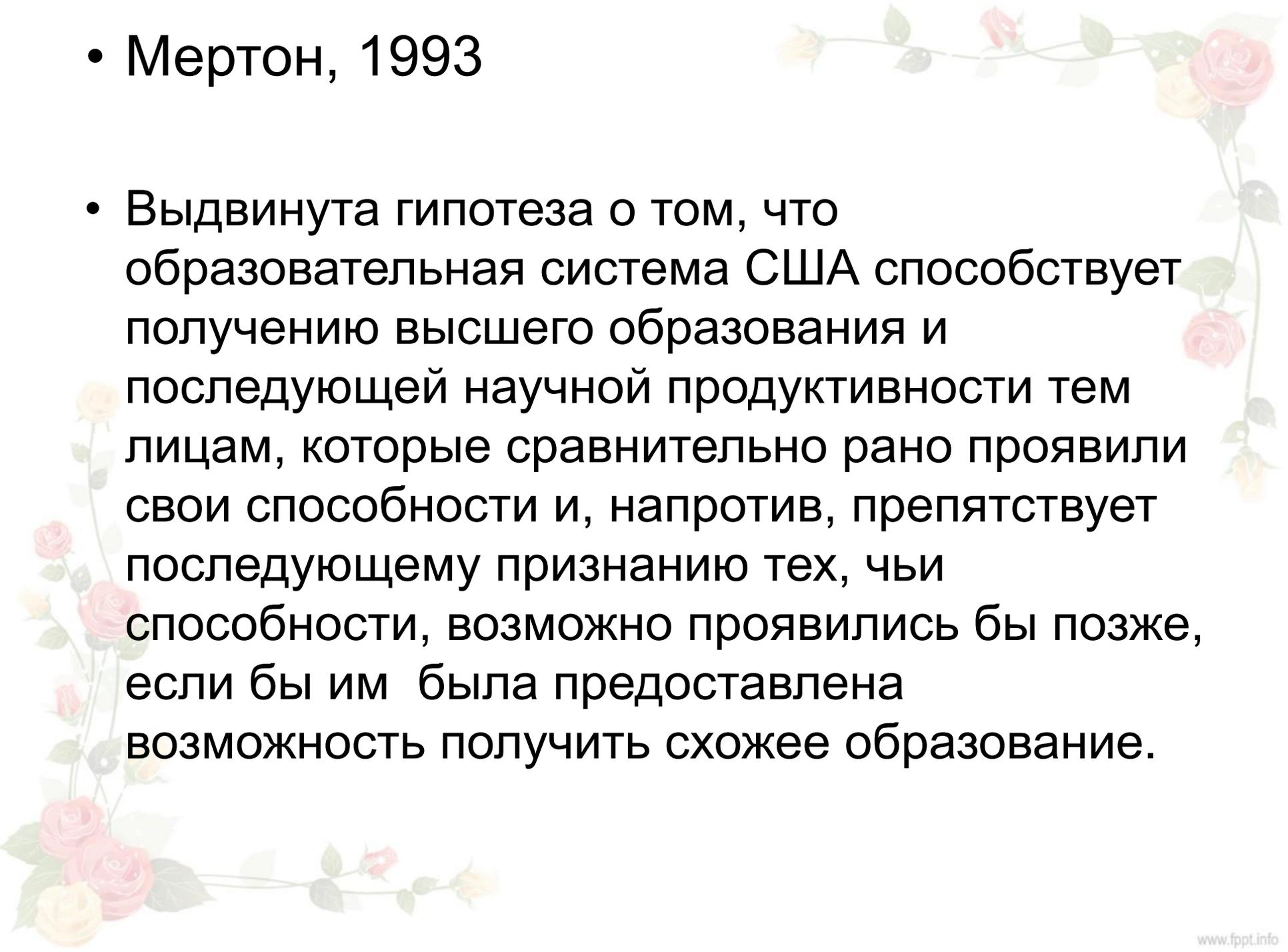




Успех ранних дарований признается в
большей степени.

Общество препятствует признанию тех,
кто получил образование позже.

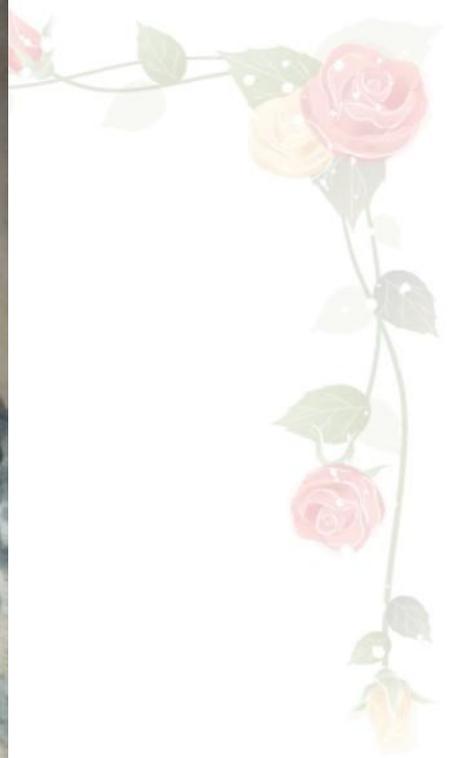


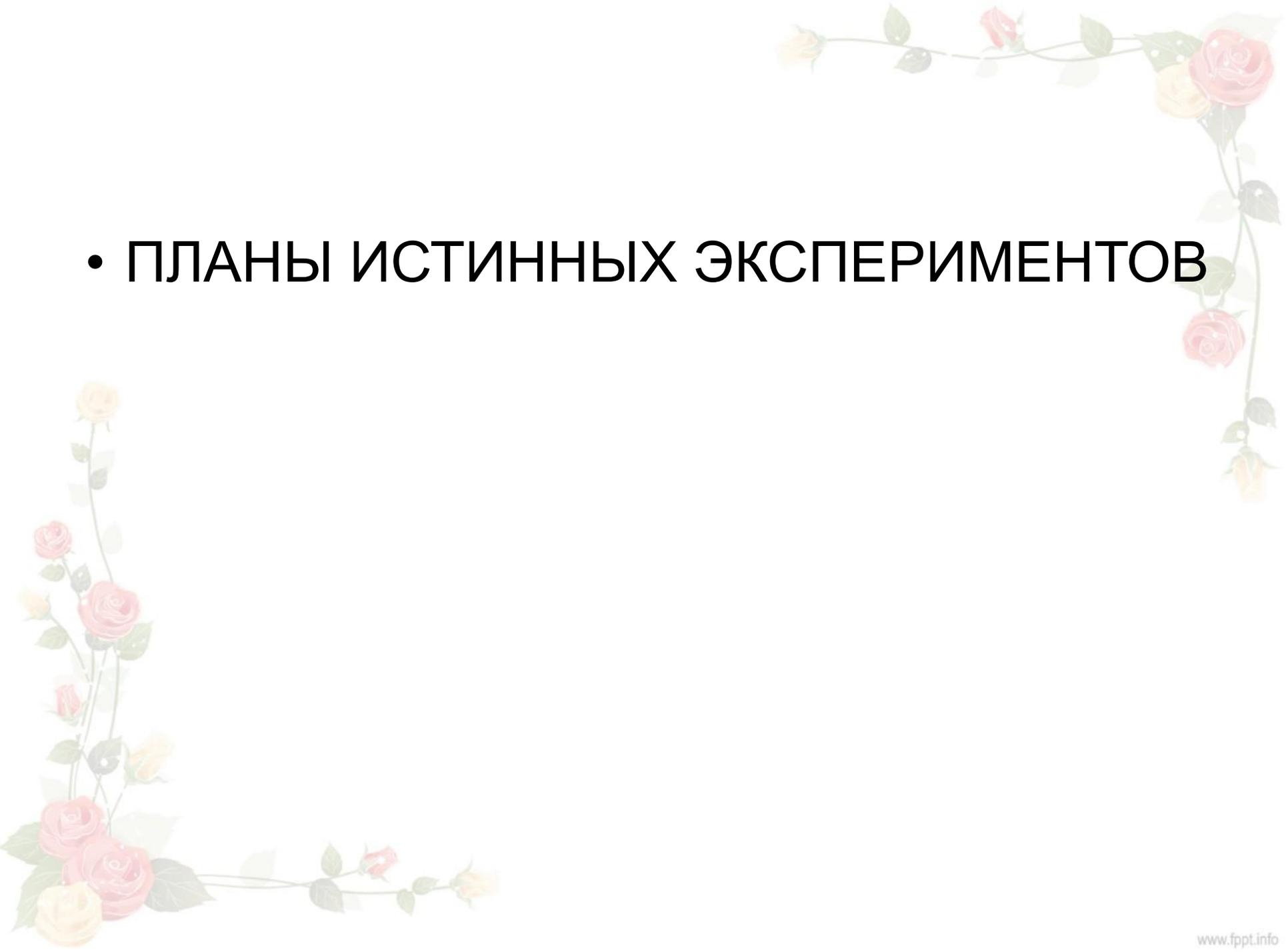


- Мертон, 1993

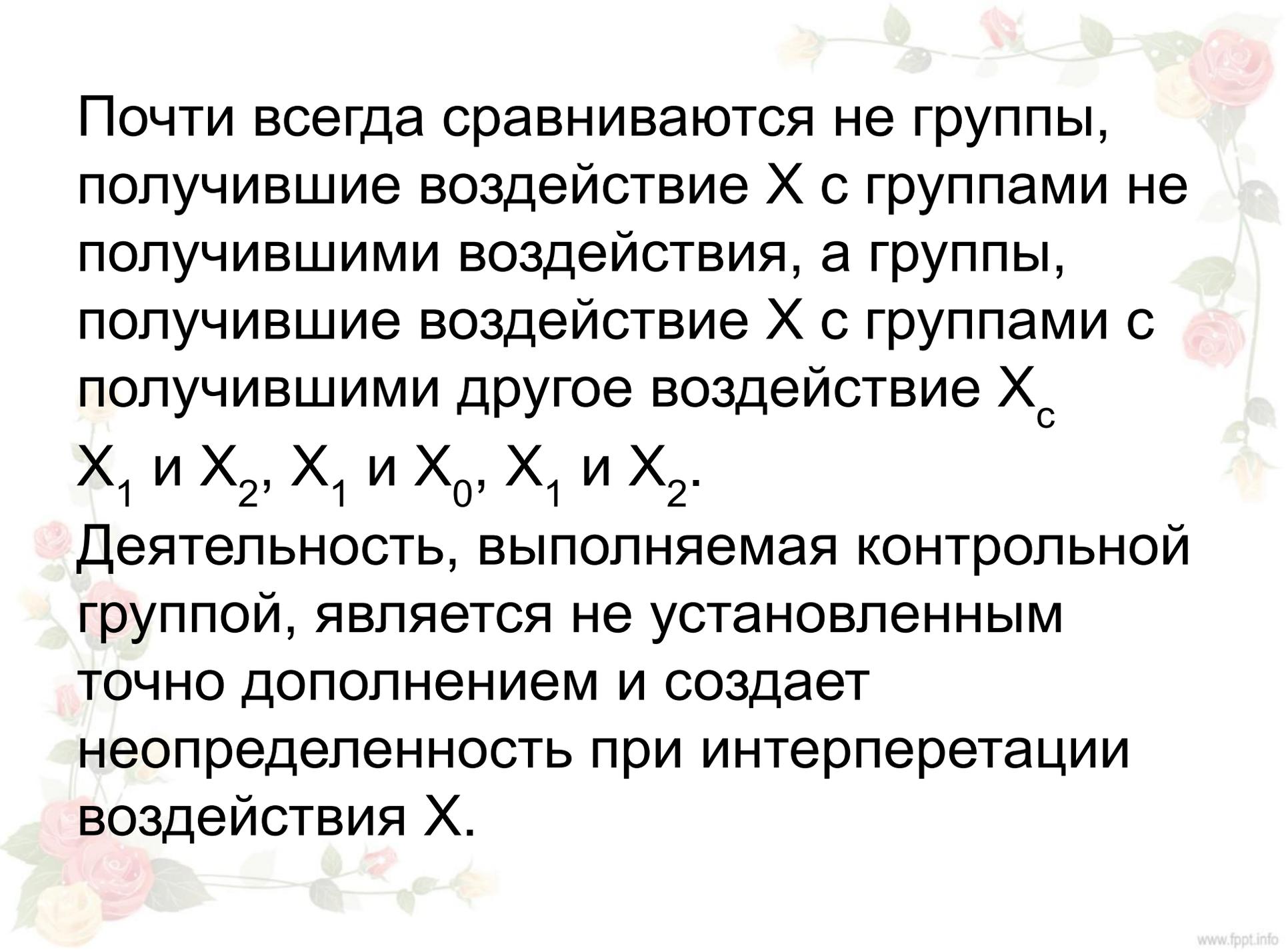
- Выдвинута гипотеза о том, что образовательная система США способствует получению высшего образования и последующей научной продуктивности тем лицам, которые сравнительно рано проявили свои способности и, напротив, препятствует последующему признанию тех, чьи способности, возможно проявились бы позже, если бы им была предоставлена возможность получить схожее образование.

	рано	поздно
успех	Рано/успех	поздно/успех 
неуспех	Рано/неуспех	Поздно/неуспех





- ПЛАНЫ ИСТИННЫХ ЭКСПЕРИМЕНТОВ



Почти всегда сравниваются не группы, получившие воздействие X с группами не получившими воздействия, а группы, получившие воздействие X с группами с получившими другое воздействие X_c

X_1 и X_2 , X_1 и X_0 , X_1 и X_2 .

Деятельность, выполняемая контрольной группой, является не установленным точно дополнением и создает неопределенность при интерперетации воздействия X .

**ПЛАН С ПРЕДВАРИТЕЛЬНЫМ И
ИТОГОВЫМ ТЕСТИРОВАНИЕМ И
КОНТРОЛЬНОЙ
(ЭКВИВАЛЕНТНОЙ) ГРУППОЙ**

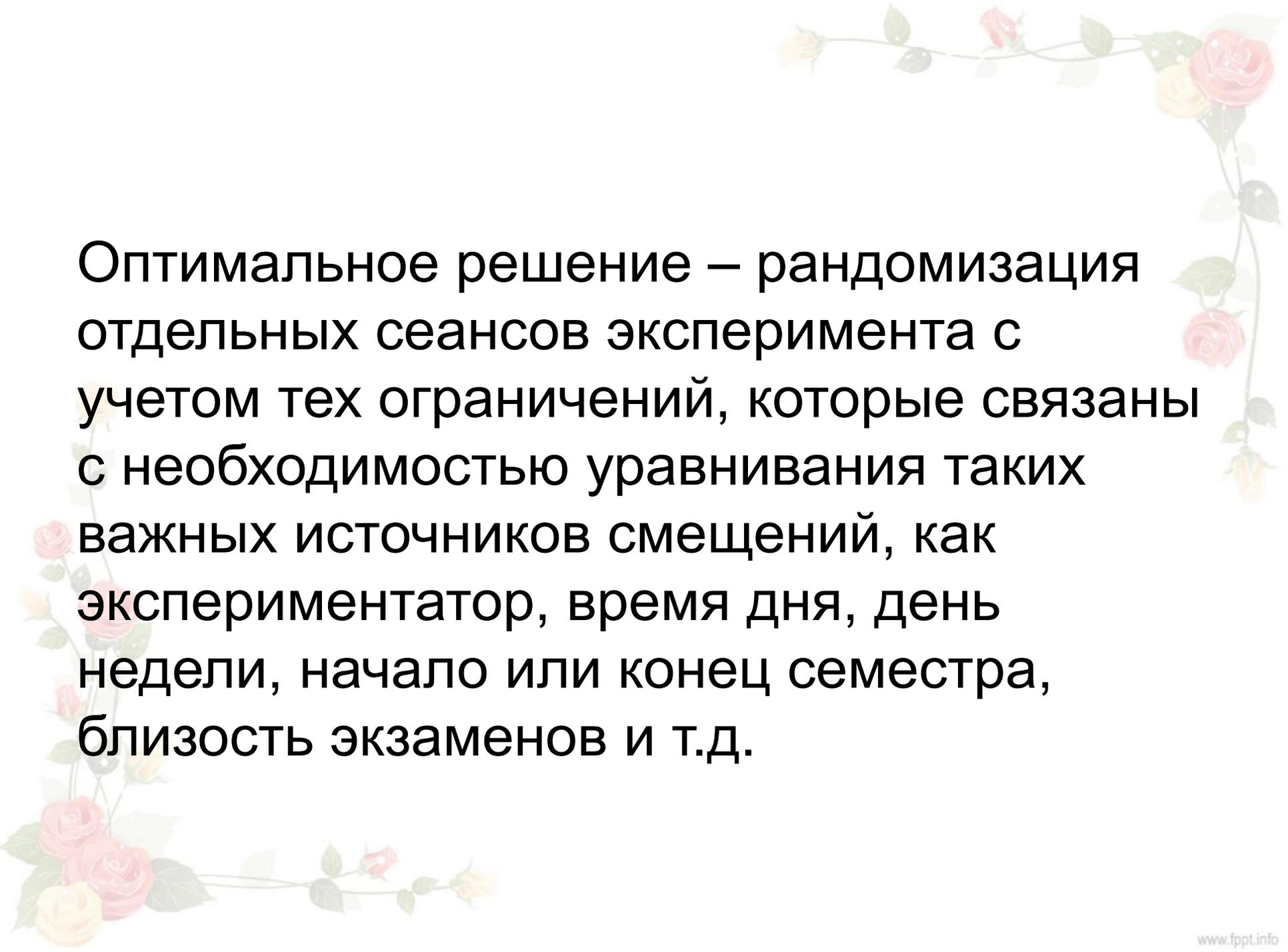
R O₁ X O₂
R O₃ O₄

- План рассчитан на одновременное исследование в экспериментальной и контрольной группах.

Если проводить одновременно необходимо использовать разных экспериментаторов.

Различия между экспериментаторами становятся одним из видов внутригрупповой истории, смешиваемой с Х.



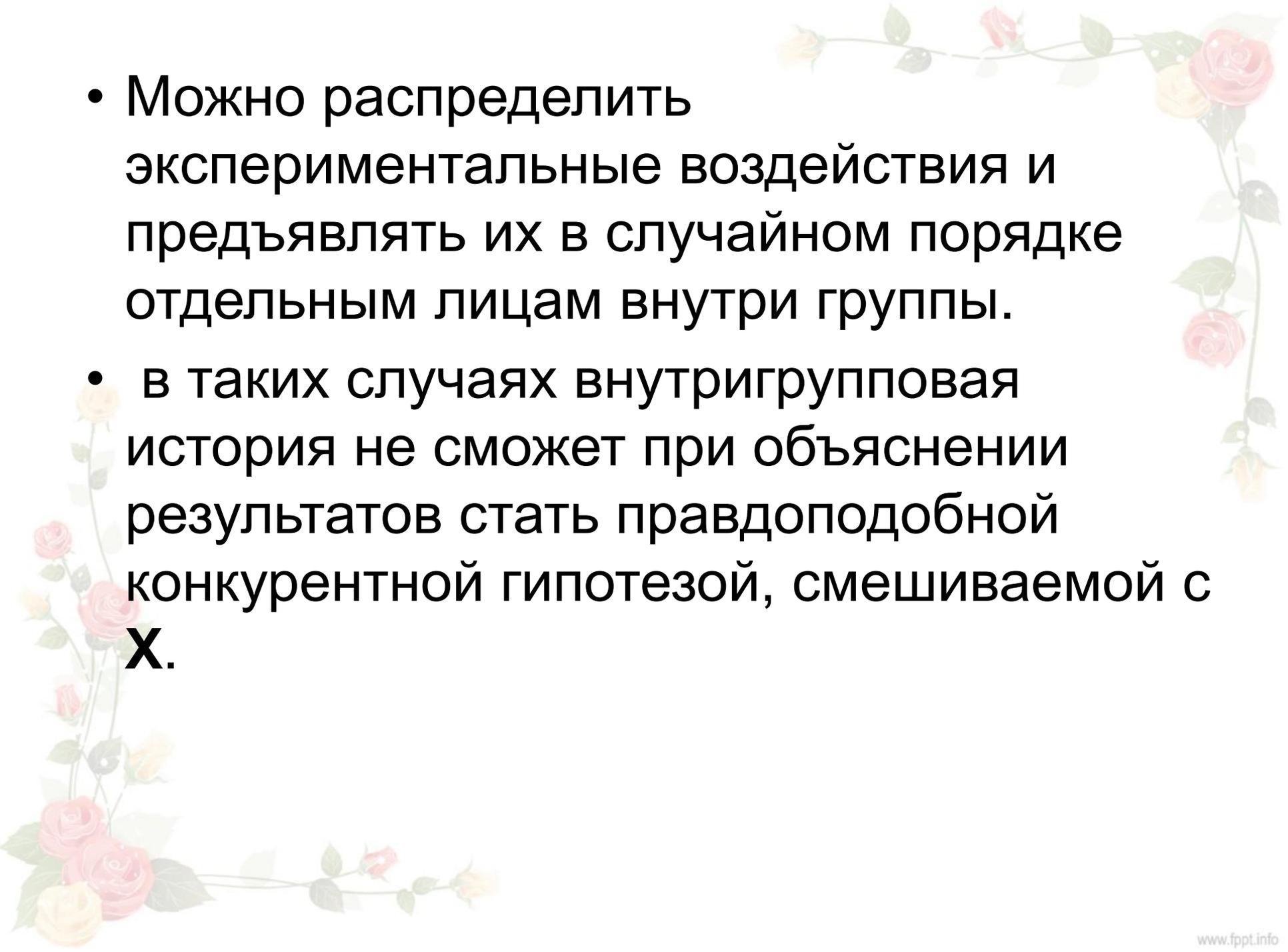


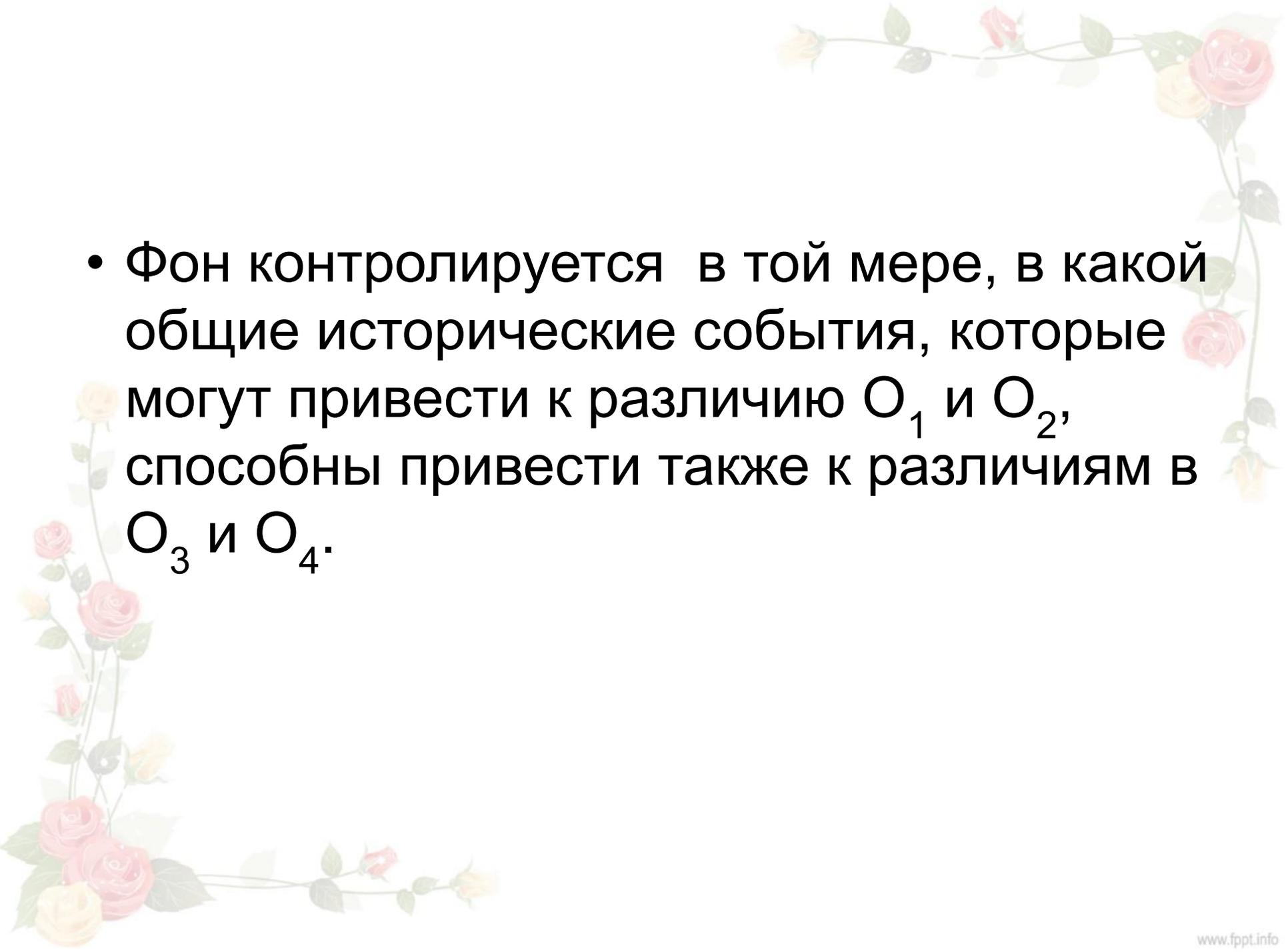
Оптимальное решение – рандомизация отдельных сеансов эксперимента с учетом тех ограничений, которые связаны с необходимостью уравнивания таких важных источников смещений, как экспериментатор, время дня, день недели, начало или конец семестра, близость экзаменов и т.д.



Групповые исследования

- Все испытуемые, принимающие участие в одном сеансе, подвергаются действию одних и тех же посторонних событий, что может оказаться отличной от X причиной сходства реакций.

- 
- Можно распределить экспериментальные воздействия и предъявлять их в случайном порядке отдельным лицам внутри группы.
 - в таких случаях внутригрупповая история не сможет при объяснении результатов стать правдоподобной конкурентной гипотезой, смешиваемой с **X**.

- 
- Фон контролируется в той мере, в какой общие исторические события, которые могут привести к различию O_1 и O_2 , способны привести также к различиям в O_3 и O_4 .

Не контролируется

Внутригрупповая история





- Если испытуемые экспериментальной группы подвергаются однократному воздействию и контрольная группа также исследуется один раз, то посторонние частные события, происходящие в каждой из этих групп, могут стать конкурентными объяснениями отличия $O_1 - O_2$ от $O_3 - O_4$.

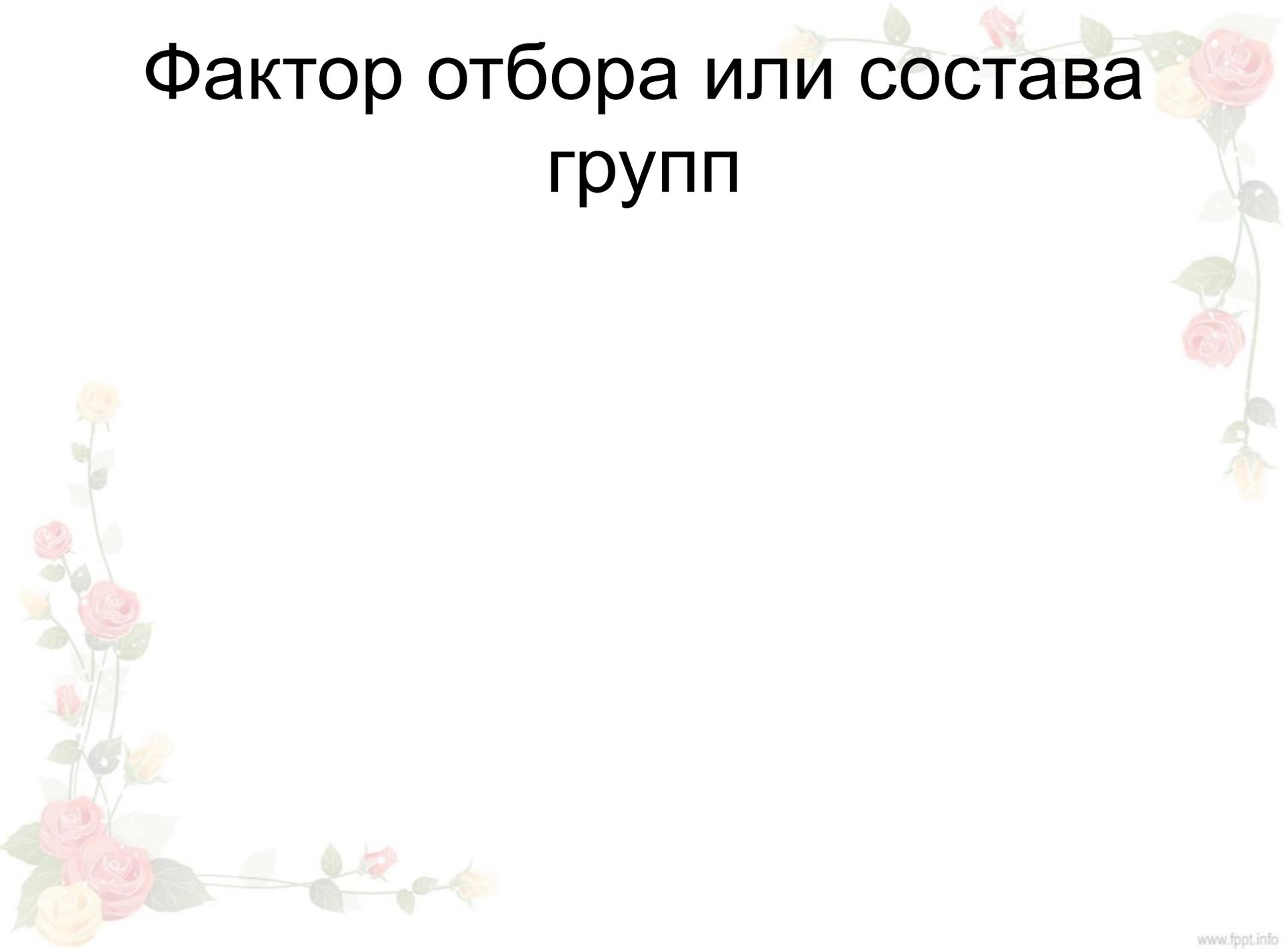
- *Вмешательство посторонних людей, звонок на мобильный телефон, шутка, предварительные замечания экспериментатора, отключение электричества.*

- Такие эксперименты нельзя считать истинными.

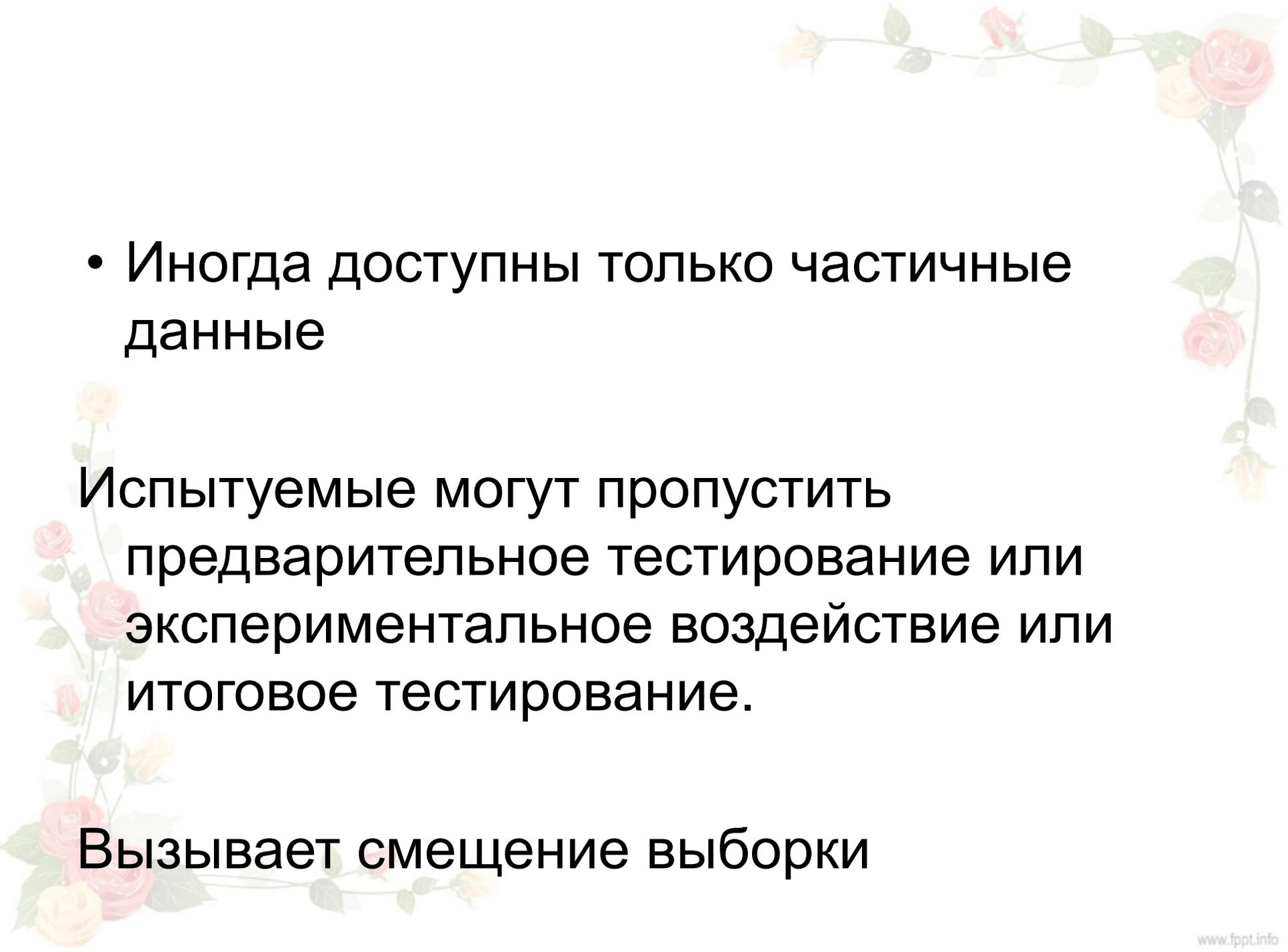
Тестирование индивидуально.

Даже при индивидуальной работе фон может оказаться не контролируемым, если вся экспериментальная группа проходит тестирование до контрольной.

Фактор отбора или состава групп

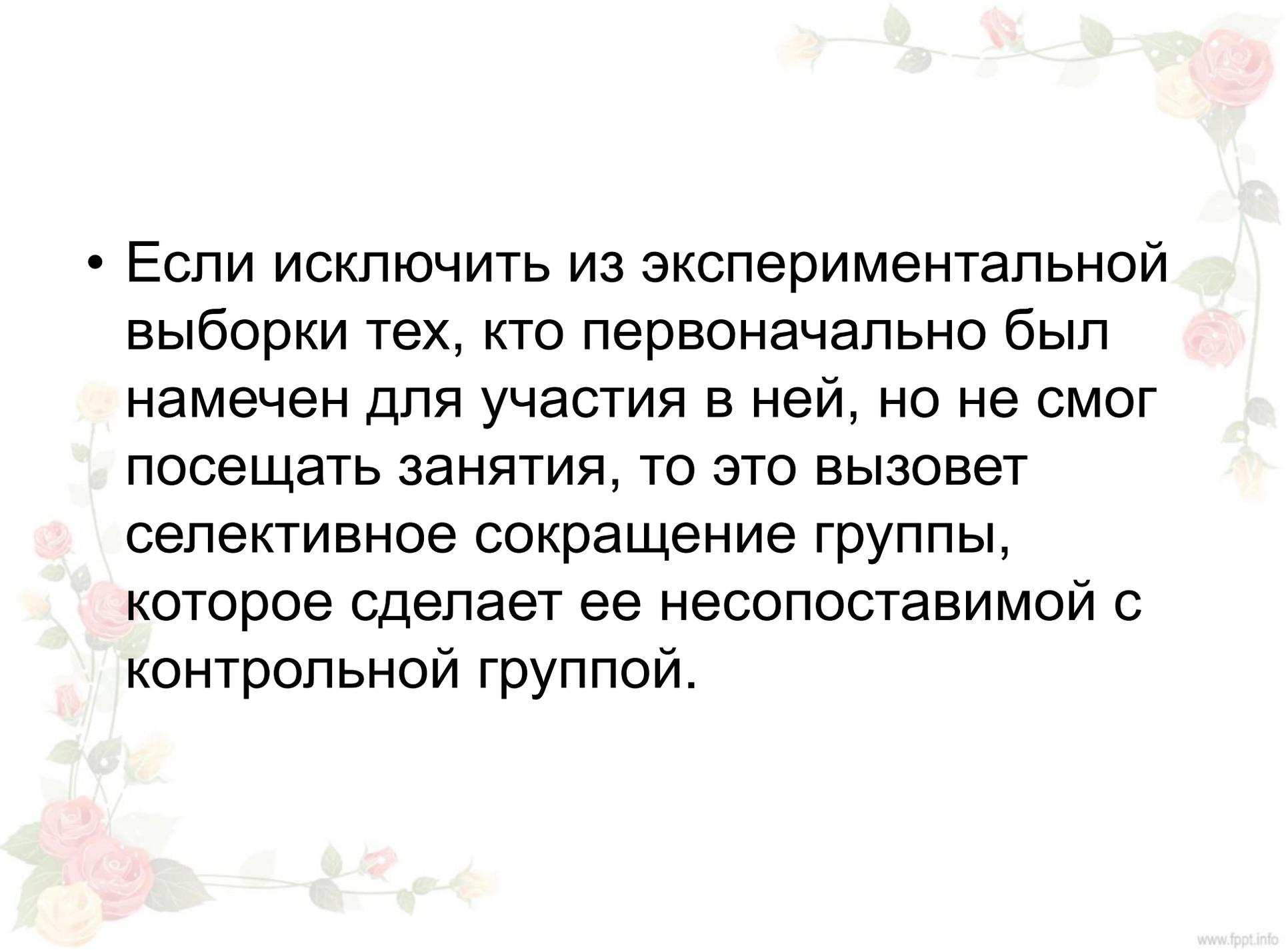


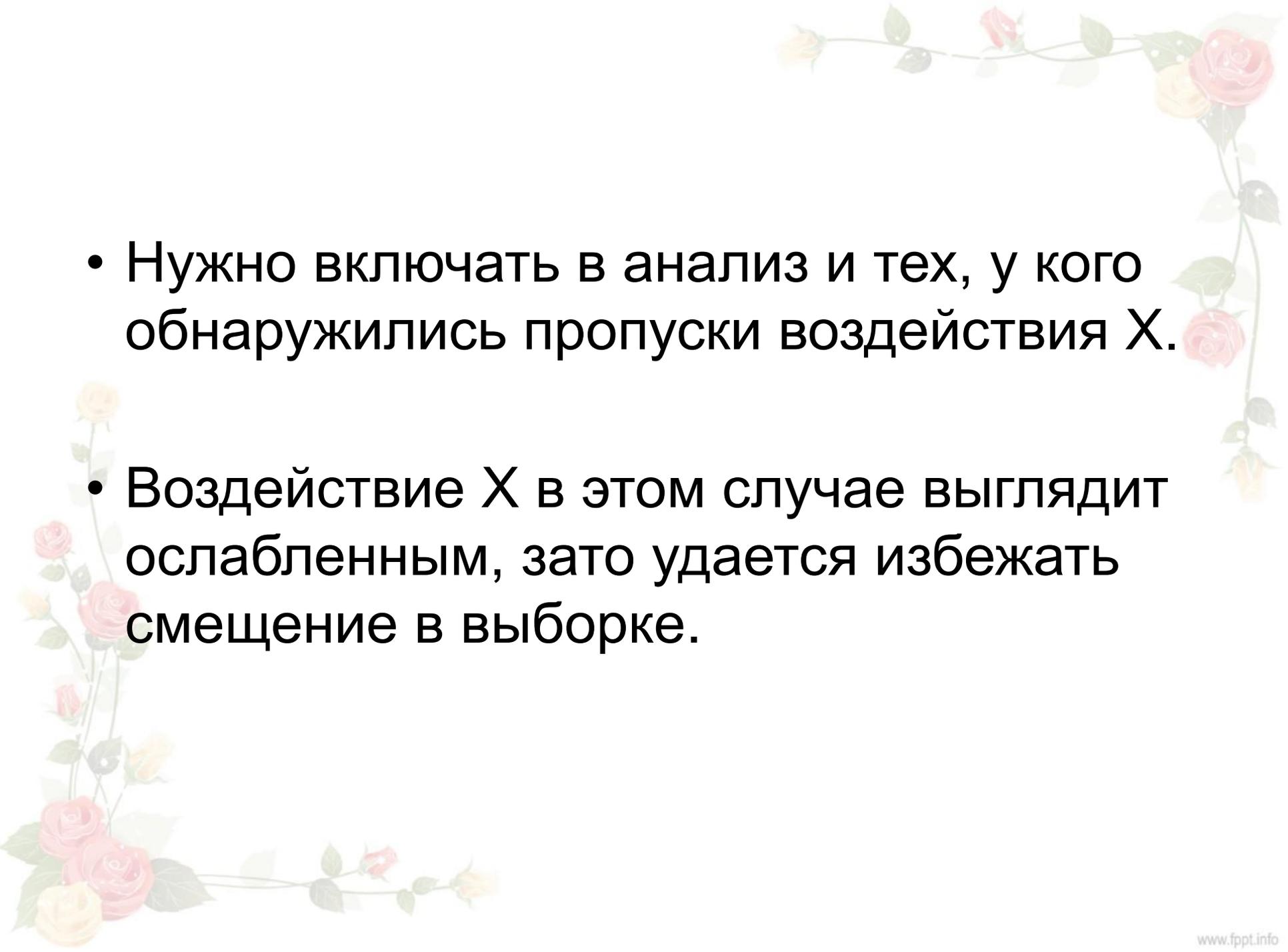
- Может ли выбывание объяснить различия между O1 и O2?

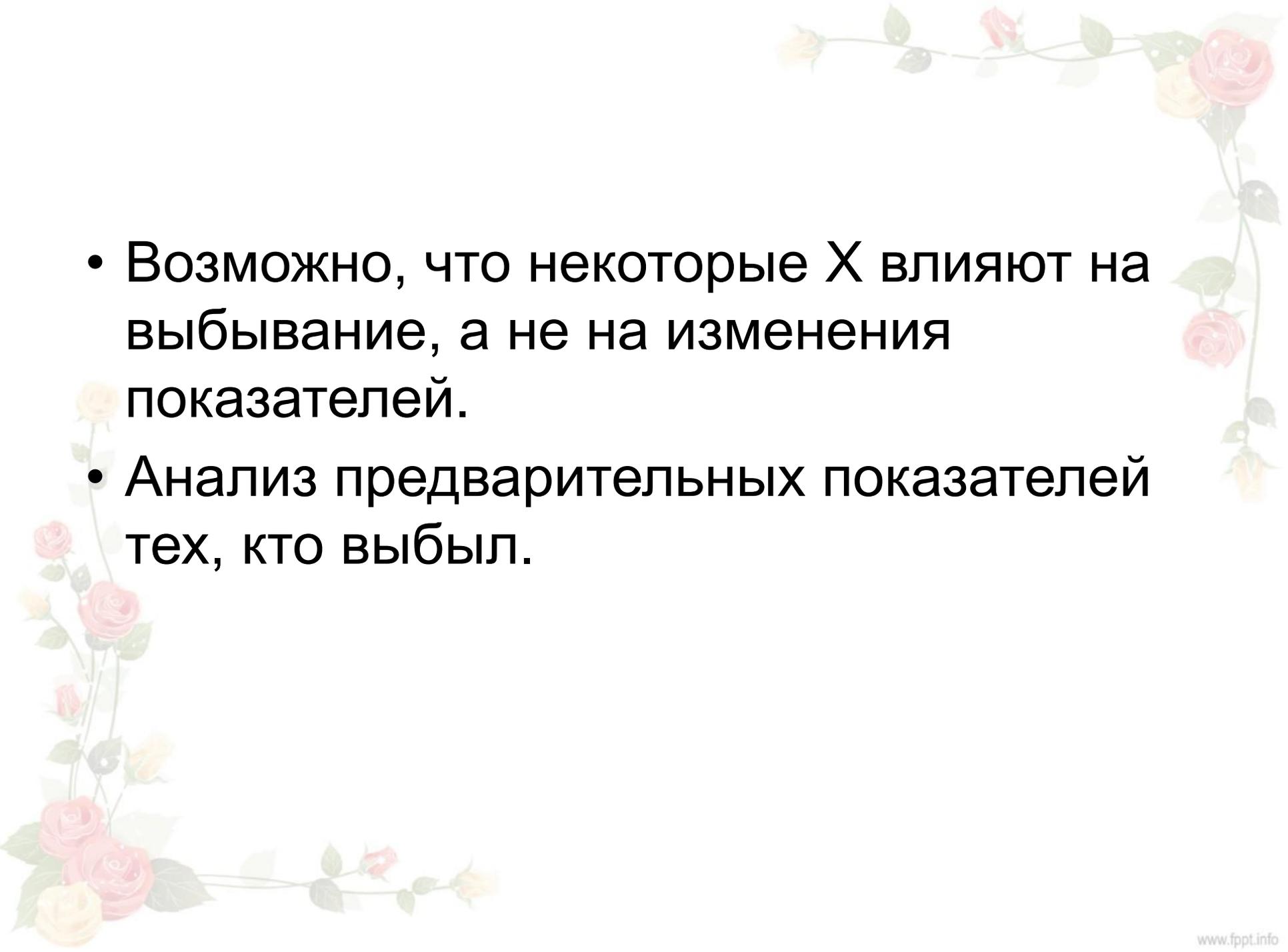
- 
- Иногда доступны только частичные данные

Испытуемые могут пропустить предварительное тестирование или экспериментальное воздействие или итоговое тестирование.

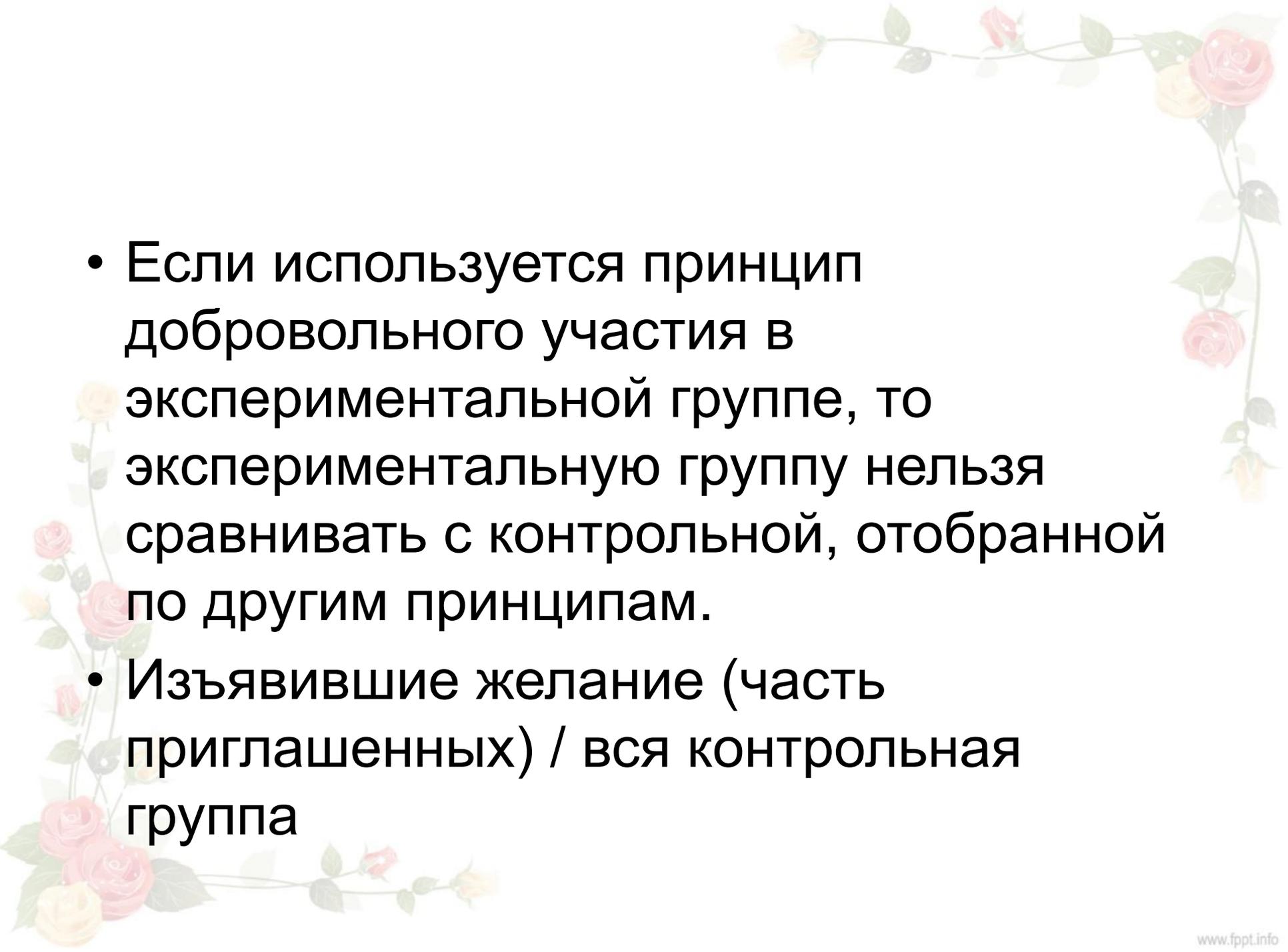
Вызывает смещение выборки

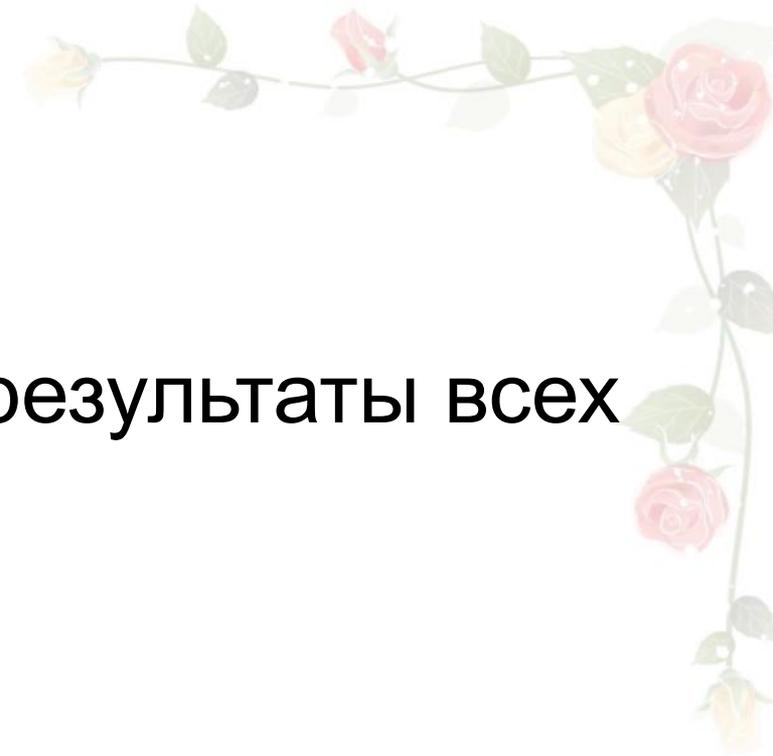
- 
- Если исключить из экспериментальной выборки тех, кто первоначально был намечен для участия в ней, но не смог посещать занятия, то это вызовет селективное сокращение группы, которое сделает ее несопоставимой с контрольной группой.

- 
- Нужно включать в анализ и тех, у кого обнаружались пропуски воздействия X .
 - Воздействие X в этом случае выглядит ослабленным, зато удастся избежать смещение в выборке.

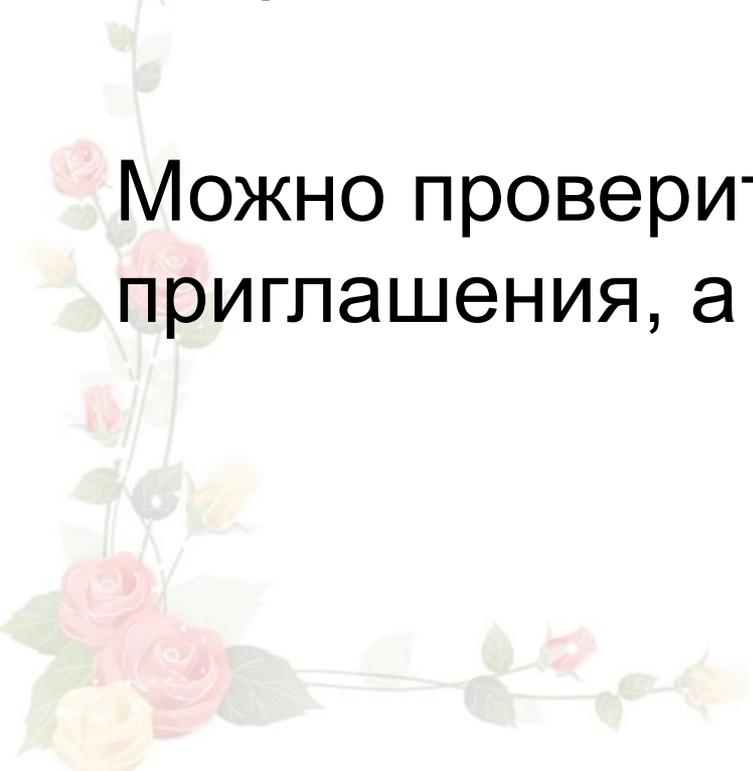
- 
- Возможно, что некоторые X влияют на выбывание, а не на изменения показателей.
 - Анализ предварительных показателей тех, кто выбыл.

- Различный характер выбывания в экспериментальной и контрольной группе.

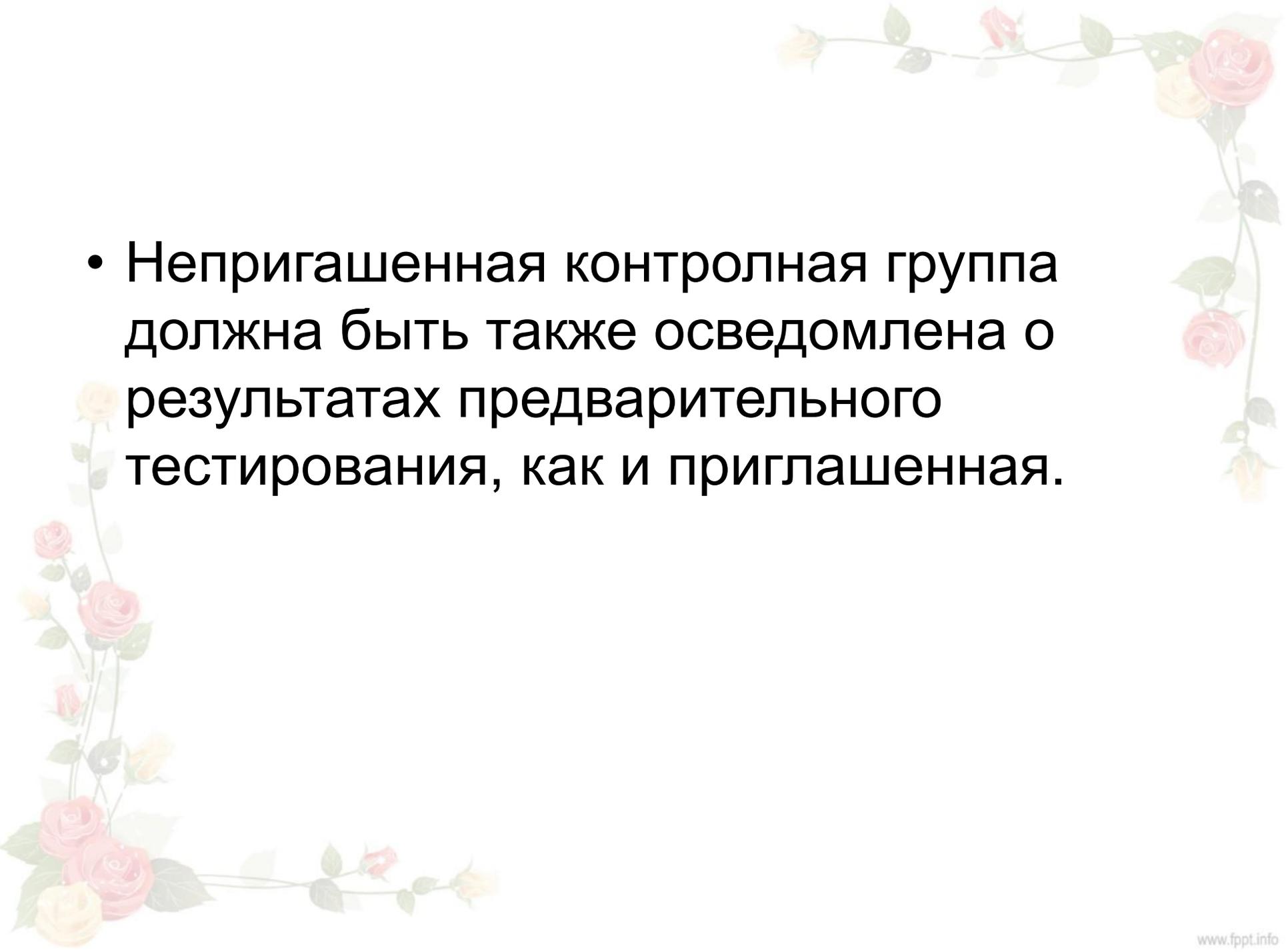
- 
- Если используется принцип добровольного участия в экспериментальной группе, то экспериментальную группу нельзя сравнивать с контрольной, отобранной по другим принципам.
 - Изъявившие желание (часть приглашенных) / вся контрольная группа



Необходимо сравнивать результаты всех приглашенных со всеми неприглашенными.



Можно проверить, влияет ли факт самого приглашения, а не X.

- 
- Непригашенная контрольная группа должна быть также осведомлена о результатах предварительного тестирования, как и приглашенная.

альтернатива

- Приглашать всех, кто нуждается в помощи и из них формировать контрольную и плацебо выборку.
- Манипулировать временем ожидания при формировании контрольной группы из тех, кто тоже хотел бы получить помощь. Это приводит к чрезмерному выбыванию из группы, тех для кого помощь отсрочена.

контролируется

- + Фон
- + Естественное развитие
- + Эффект тестирования
- + Инструментальная погрешность
- + Регрессия
- + Состав групп
- + Выбывание
- + Взаимодействие состава групп с естественным развитием

Источники внутренней невалидности не выявлены

Источники внешней невалидности

❖ взаимодействие тестирования и х

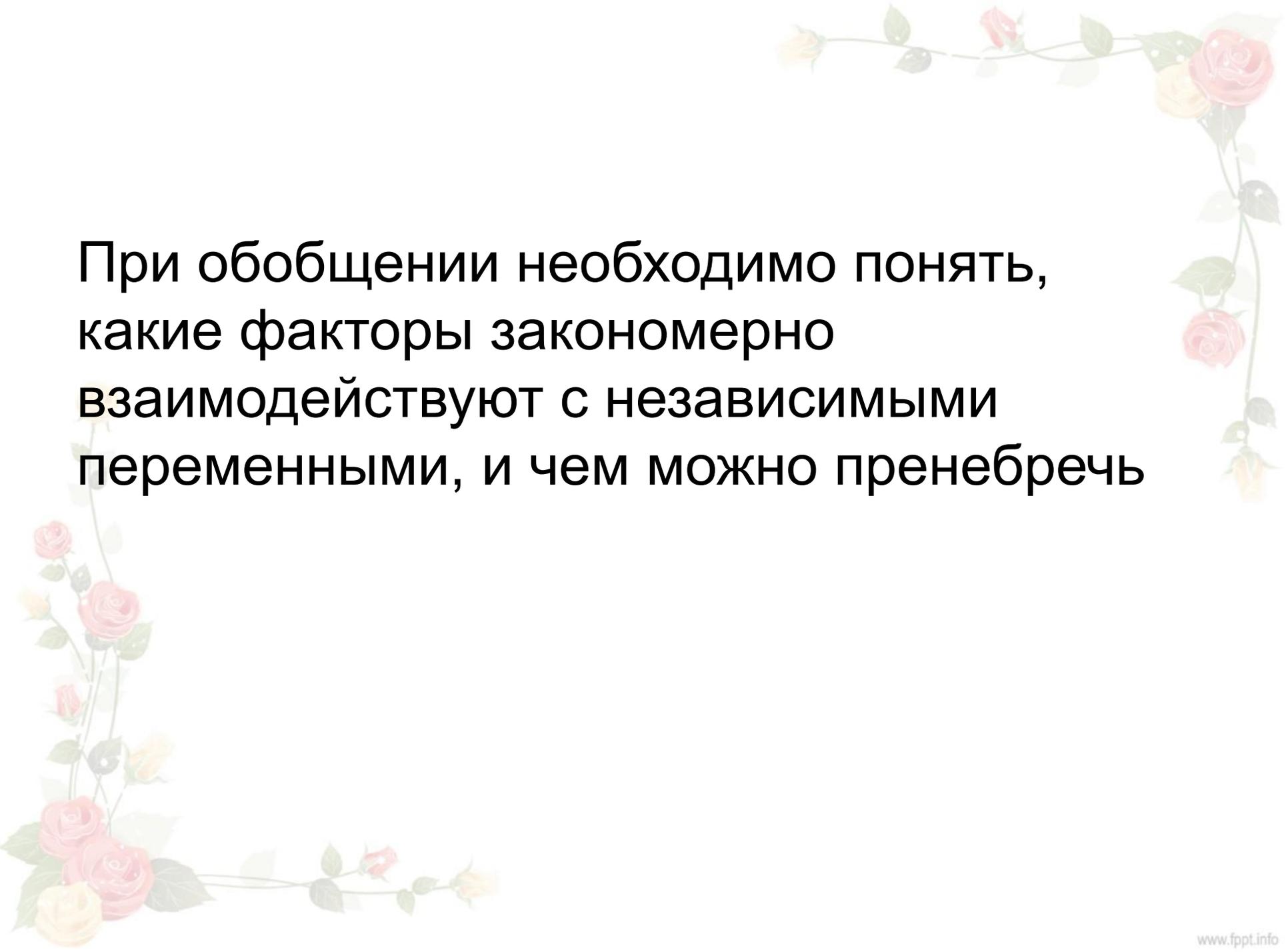


Внешняя валидность

Эффект продемонстрирован только для тех конкретных условий, которые оказались общими для экспериментальной и контрольной групп, то есть только для предварительно обследованных групп, включающих лиц определенного возраста, интеллекта, социально-экономического статуса, географического региона, взятых в определенный исторический момент.....



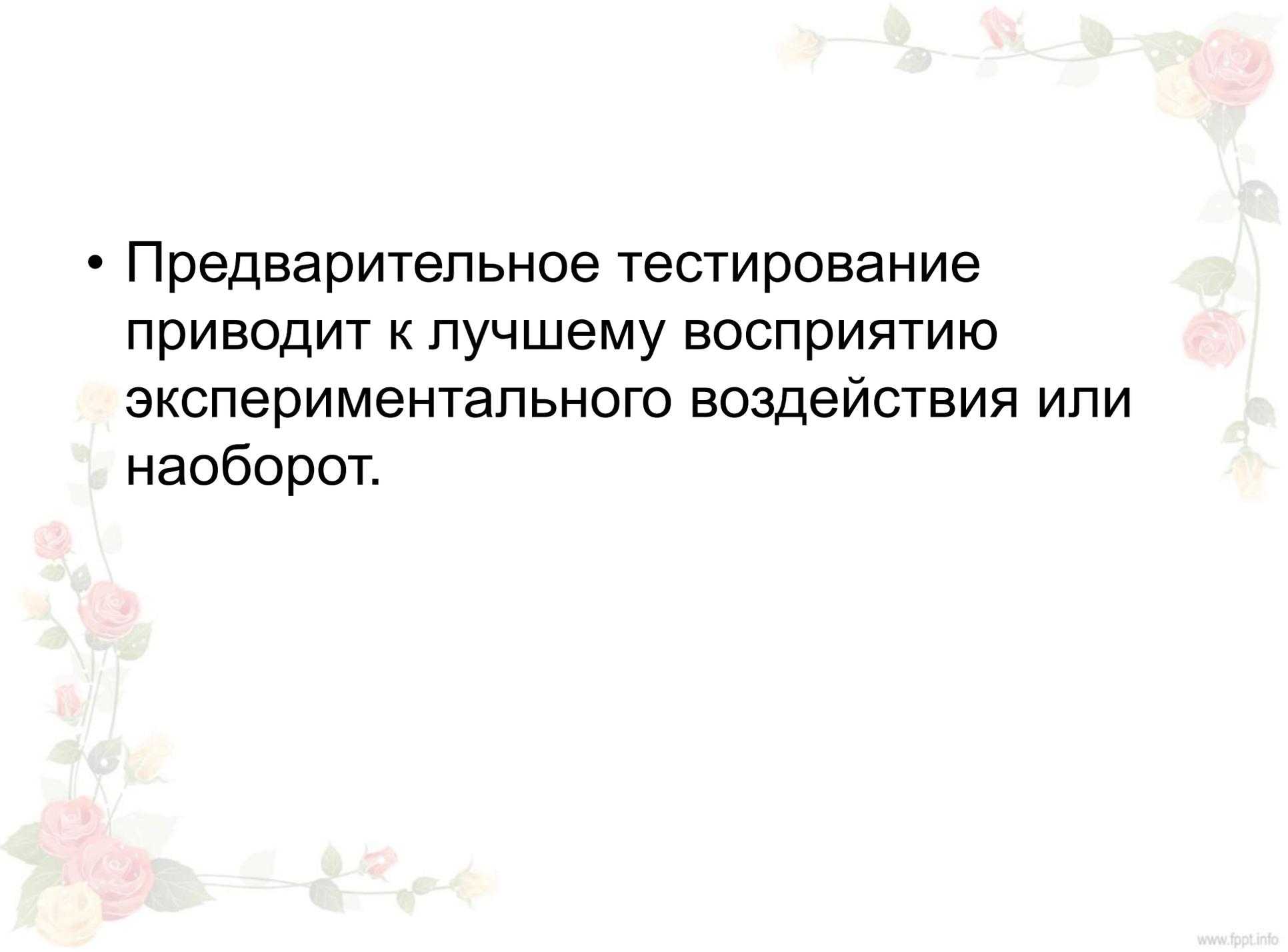
Deadly Sins

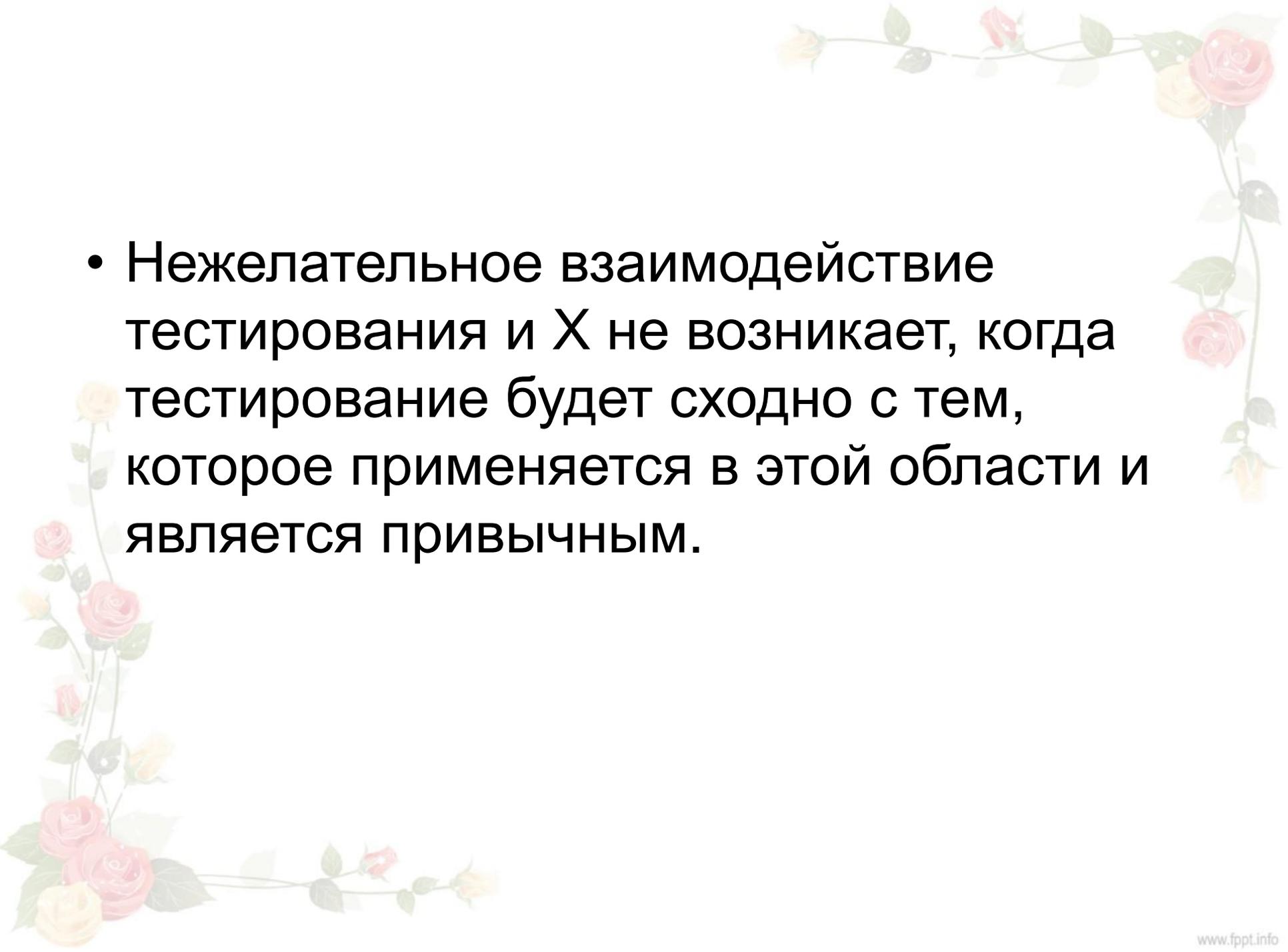


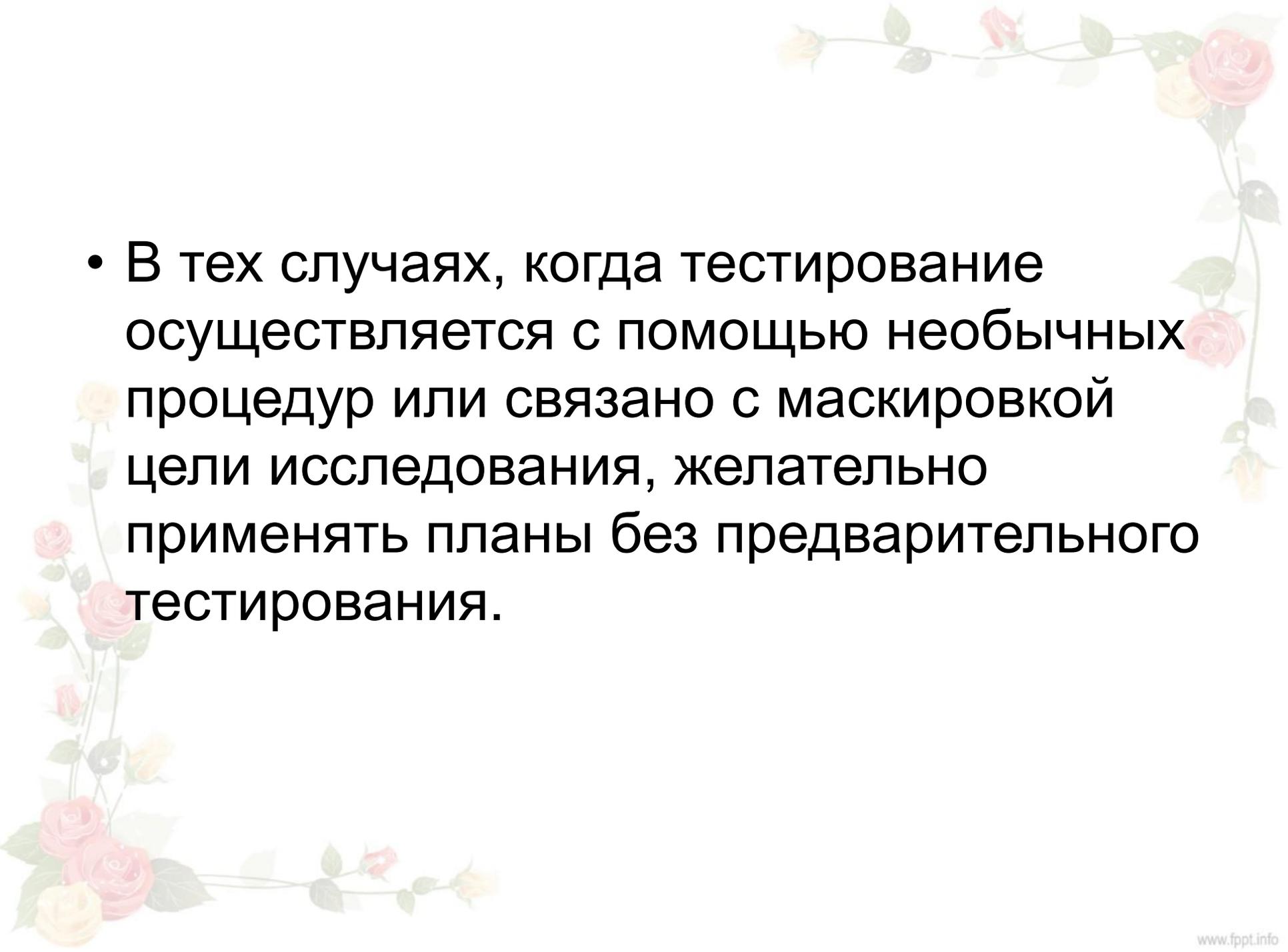
При обобщении необходимо понять, какие факторы закономерно взаимодействуют с независимыми переменными, и чем можно пренебречь

Источники внешней невалидности

❖ взаимодействие тестирования и х

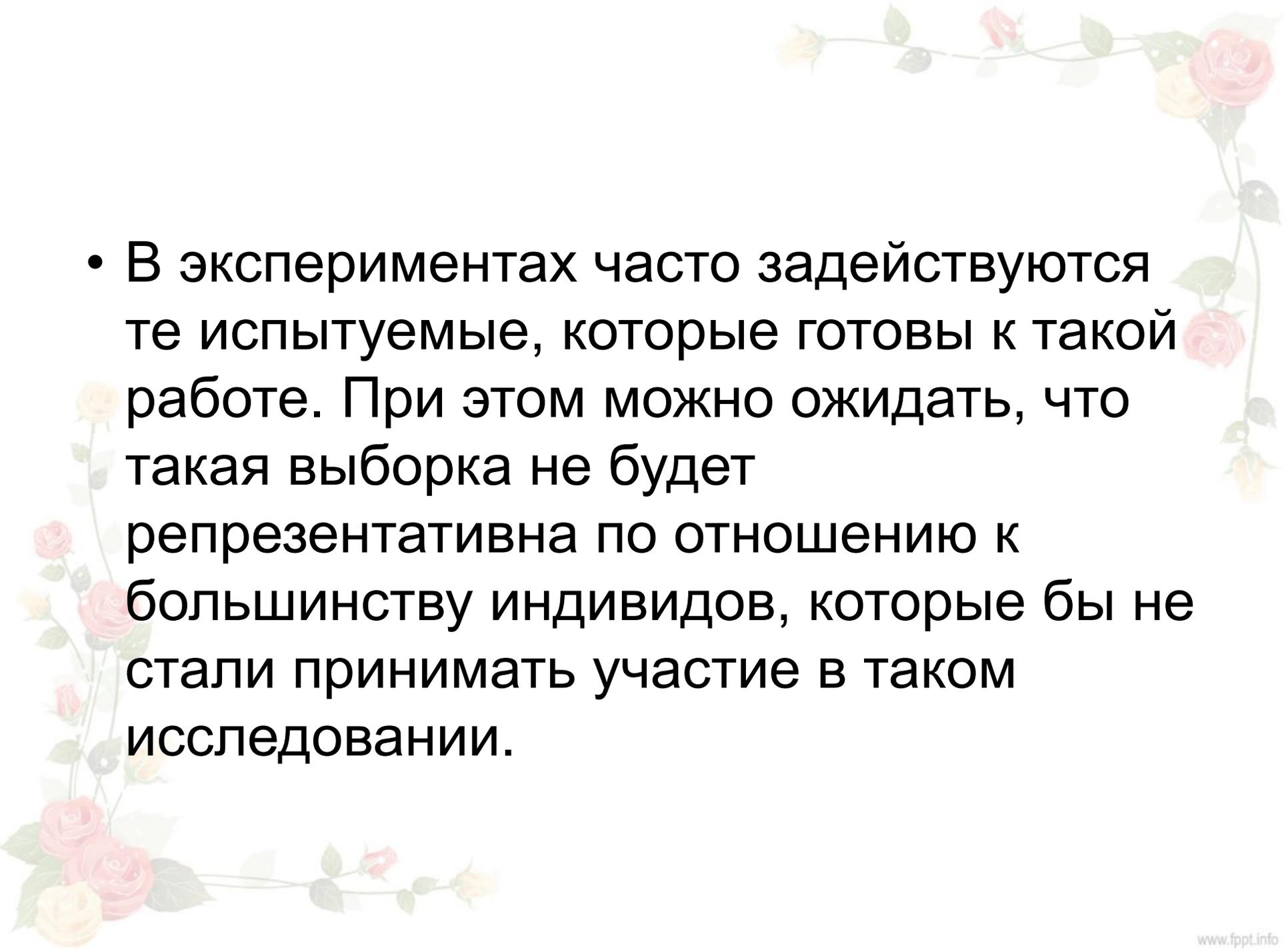
- 
- Предварительное тестирование приводит к лучшему восприятию экспериментального воздействия или наоборот.

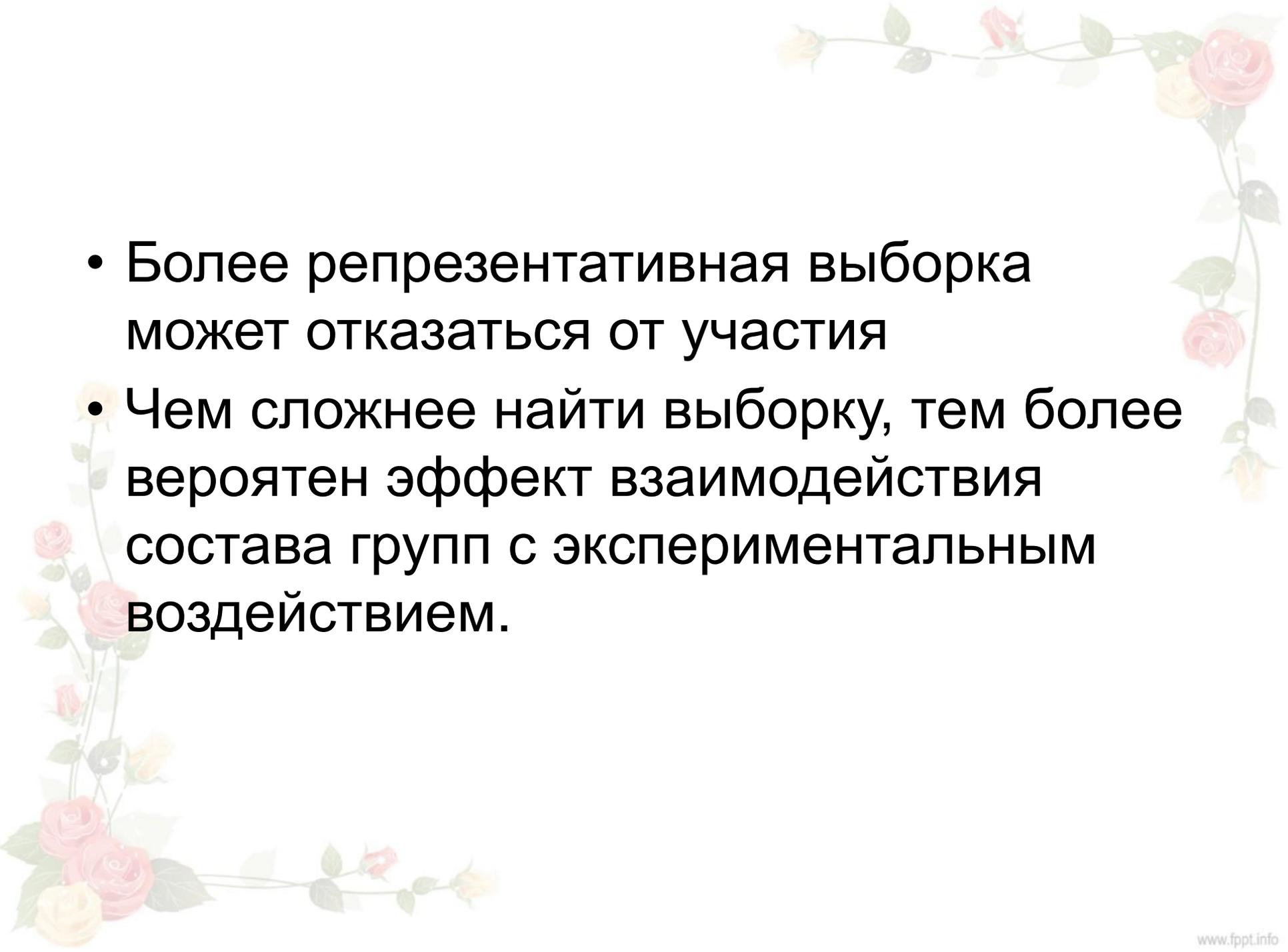
- 
- Нежелательное взаимодействие тестирования и X не возникает, когда тестирование будет сходно с тем, которое применяется в этой области и является привычным.

- 
- В тех случаях, когда тестирование осуществляется с помощью необычных процедур или связано с маскировкой цели исследования, желательно применять планы без предварительного тестирования.

Взаимодействие отбора и состава группы

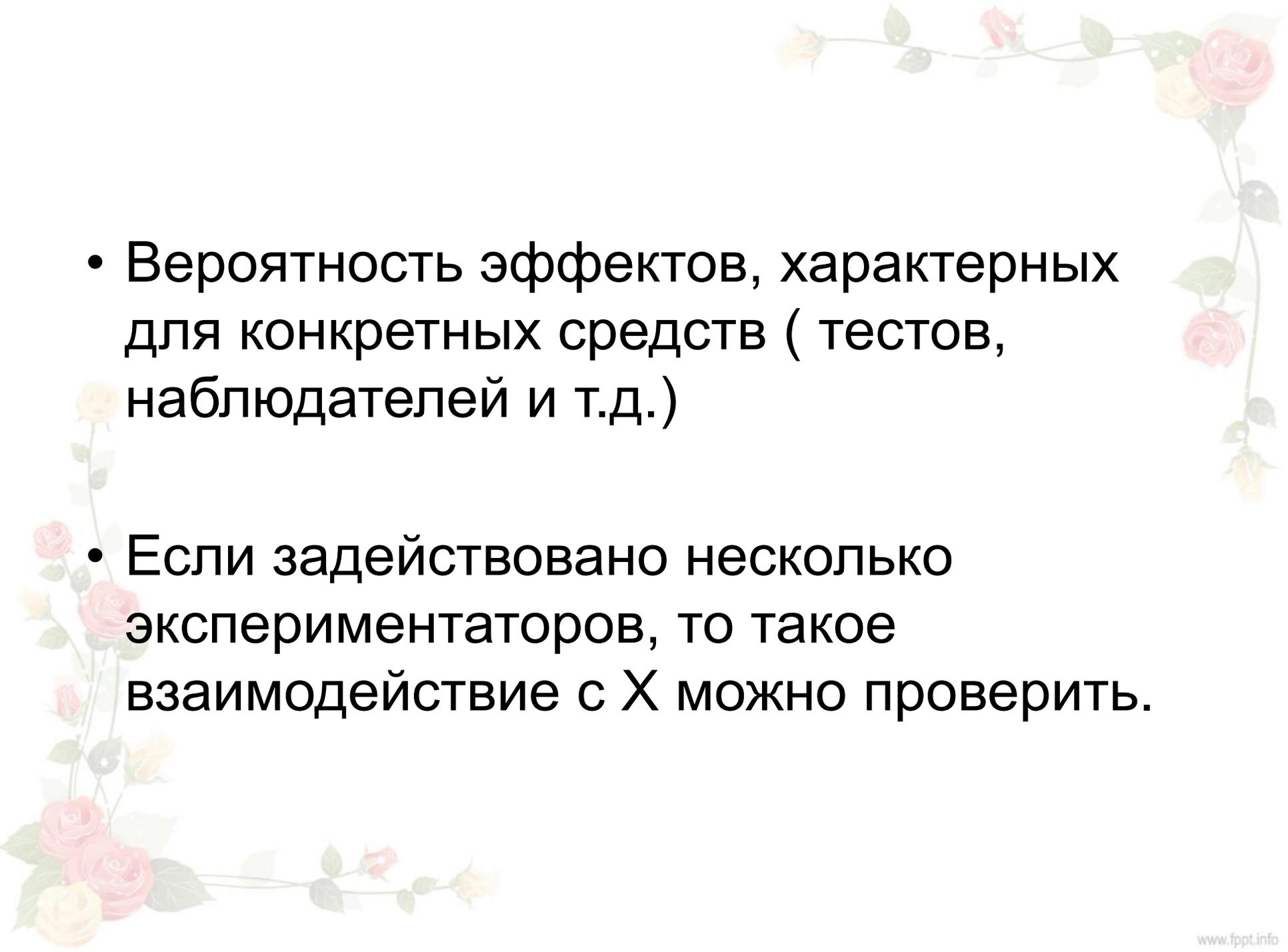
- Необходимо убедиться в отсутствии взаимодействия характеристик выборки с экспериментальным воздействием
- Пример: все школы отказываются от участия в исследовании, кроме одной (есть вероятность, что эта выборка достаточно сильно отличается от остальных выборок)

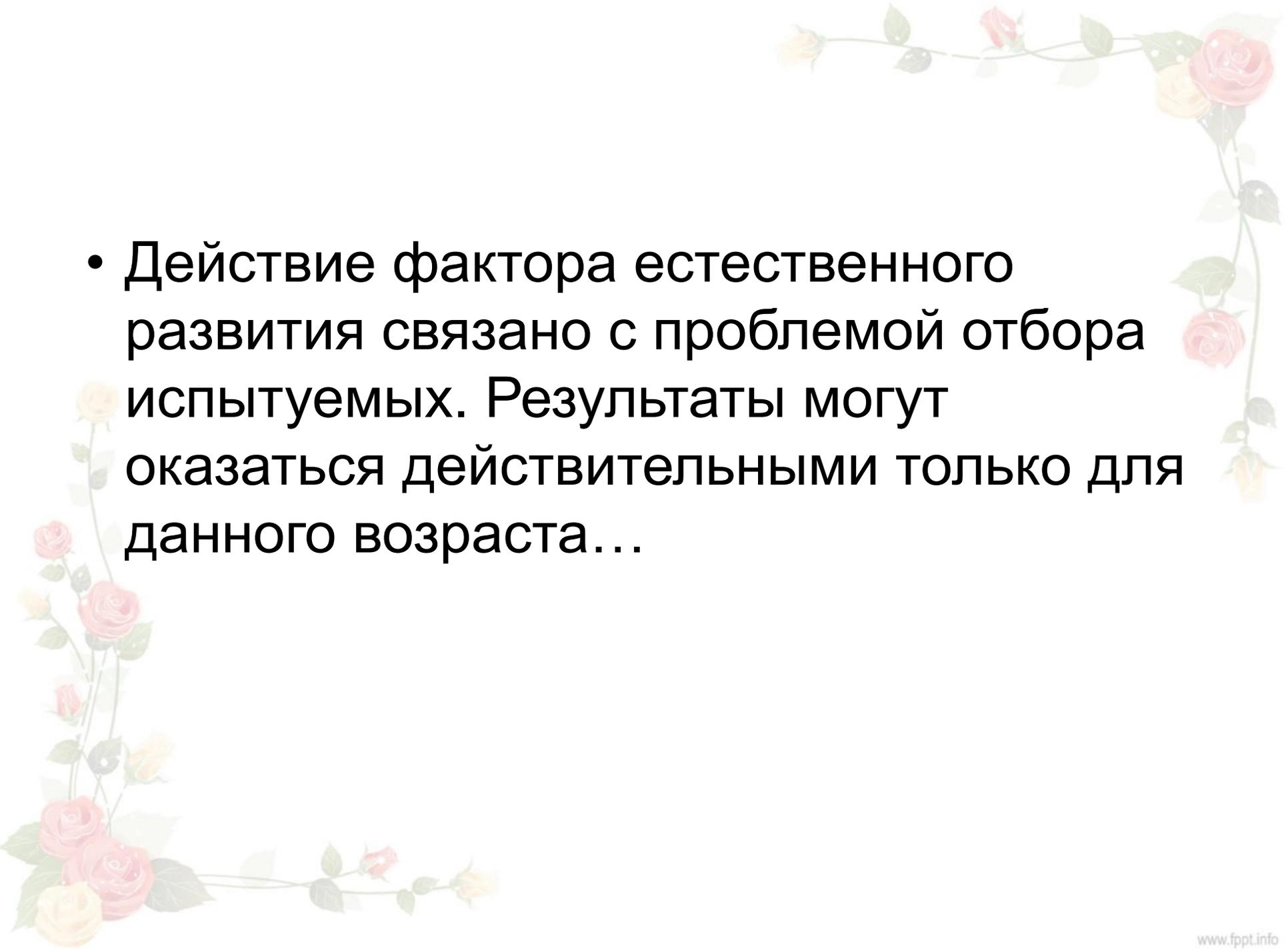
- 
- В экспериментах часто задействуются те испытуемые, которые готовы к такой работе. При этом можно ожидать, что такая выборка не будет репрезентативна по отношению к большинству индивидов, которые бы не стали принимать участие в таком исследовании.

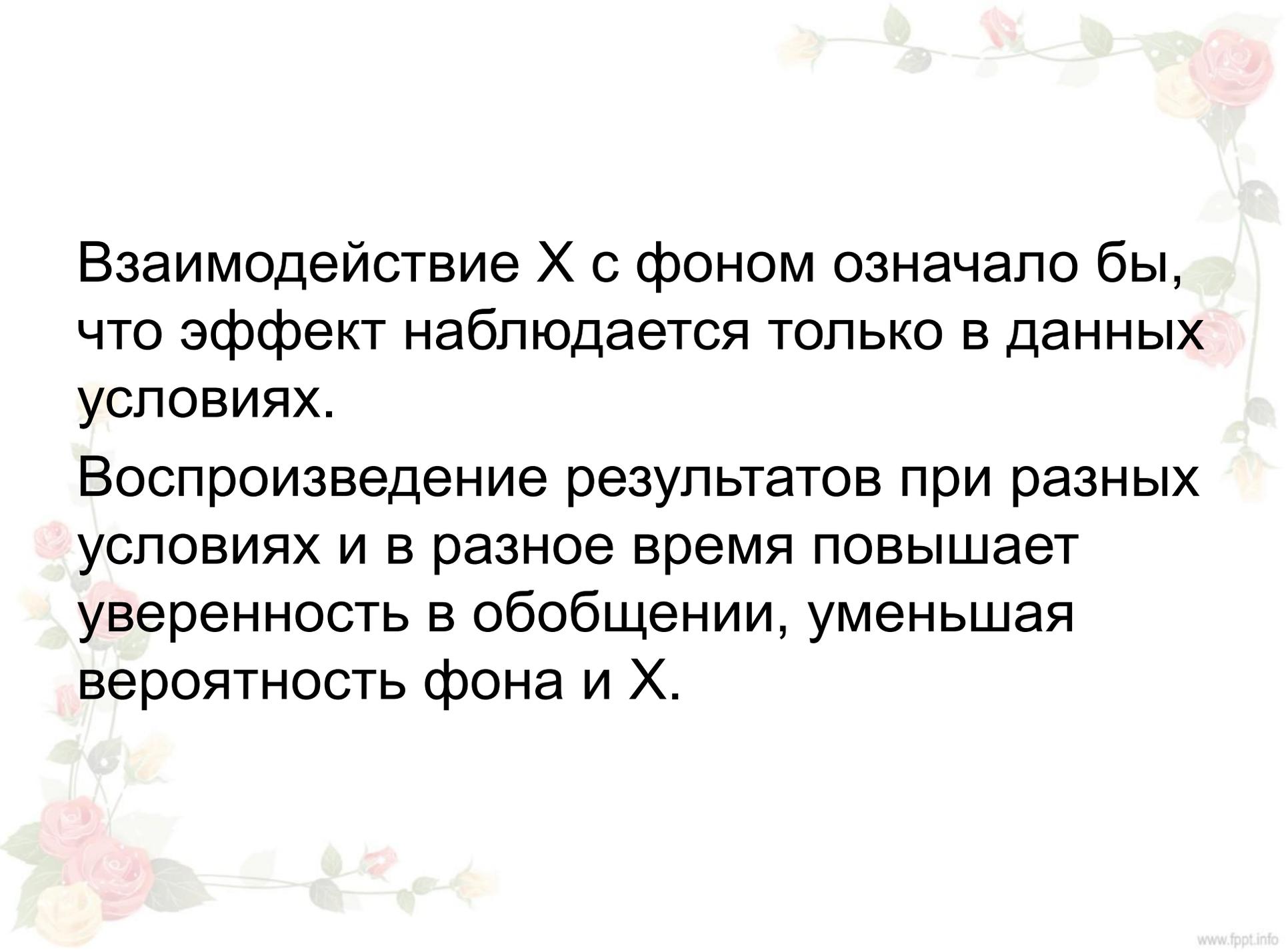
- 
- Более репрезентативная выборка может отказаться от участия
 - Чем сложнее найти выборку, тем более вероятен эффект взаимодействия состава групп с экспериментальным воздействием.

Угрозы внешней валидности

- Дифференцированное выбывание как следствие x .

- 
- Вероятность эффектов, характерных для конкретных средств (тестов, наблюдателей и т.д.)
 - Если задействовано несколько экспериментаторов, то такое взаимодействие с X можно проверить.

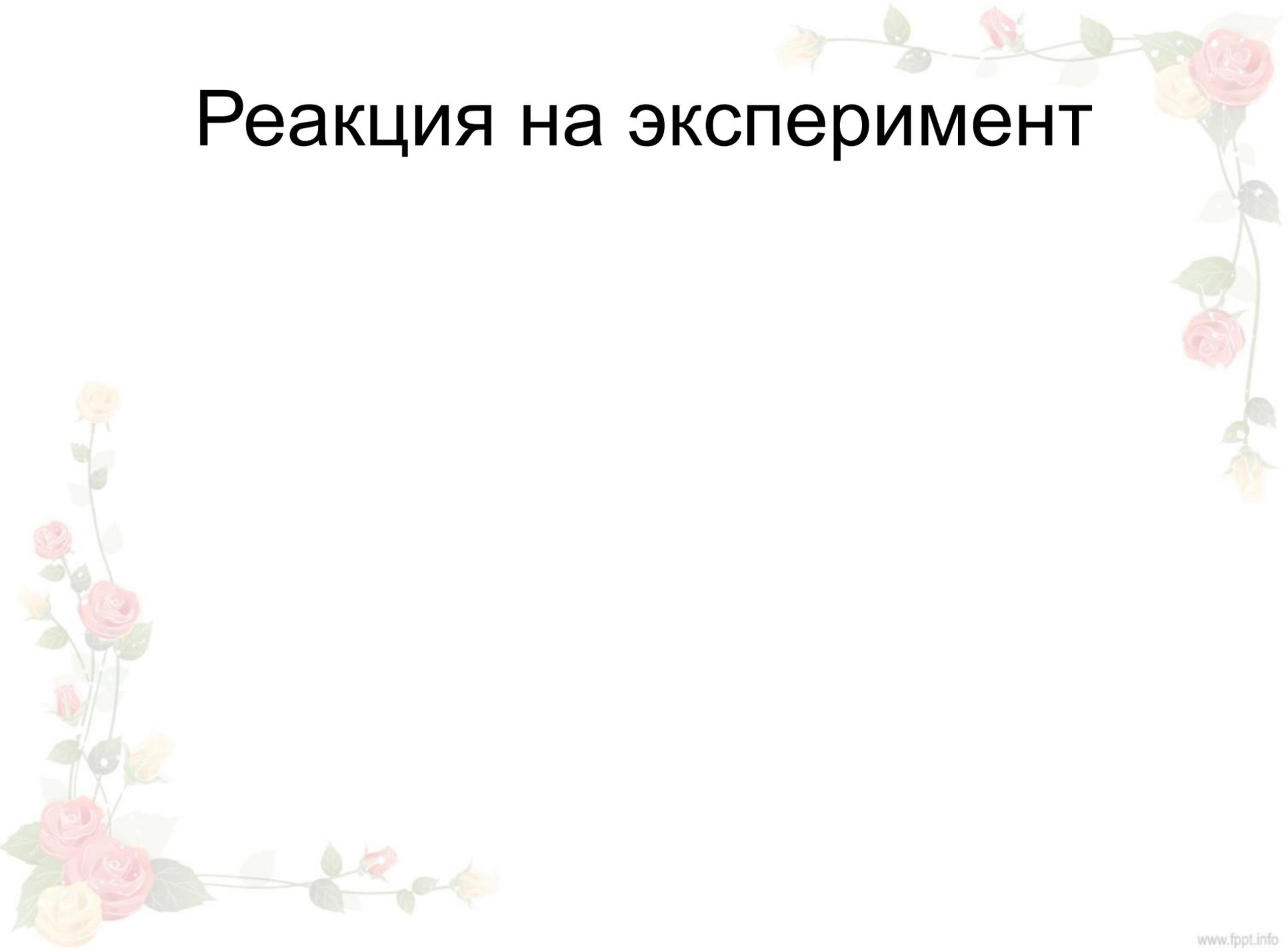
- 
- Действие фактора естественного развития связано с проблемой отбора испытуемых. Результаты могут оказаться действительными только для данного возраста...

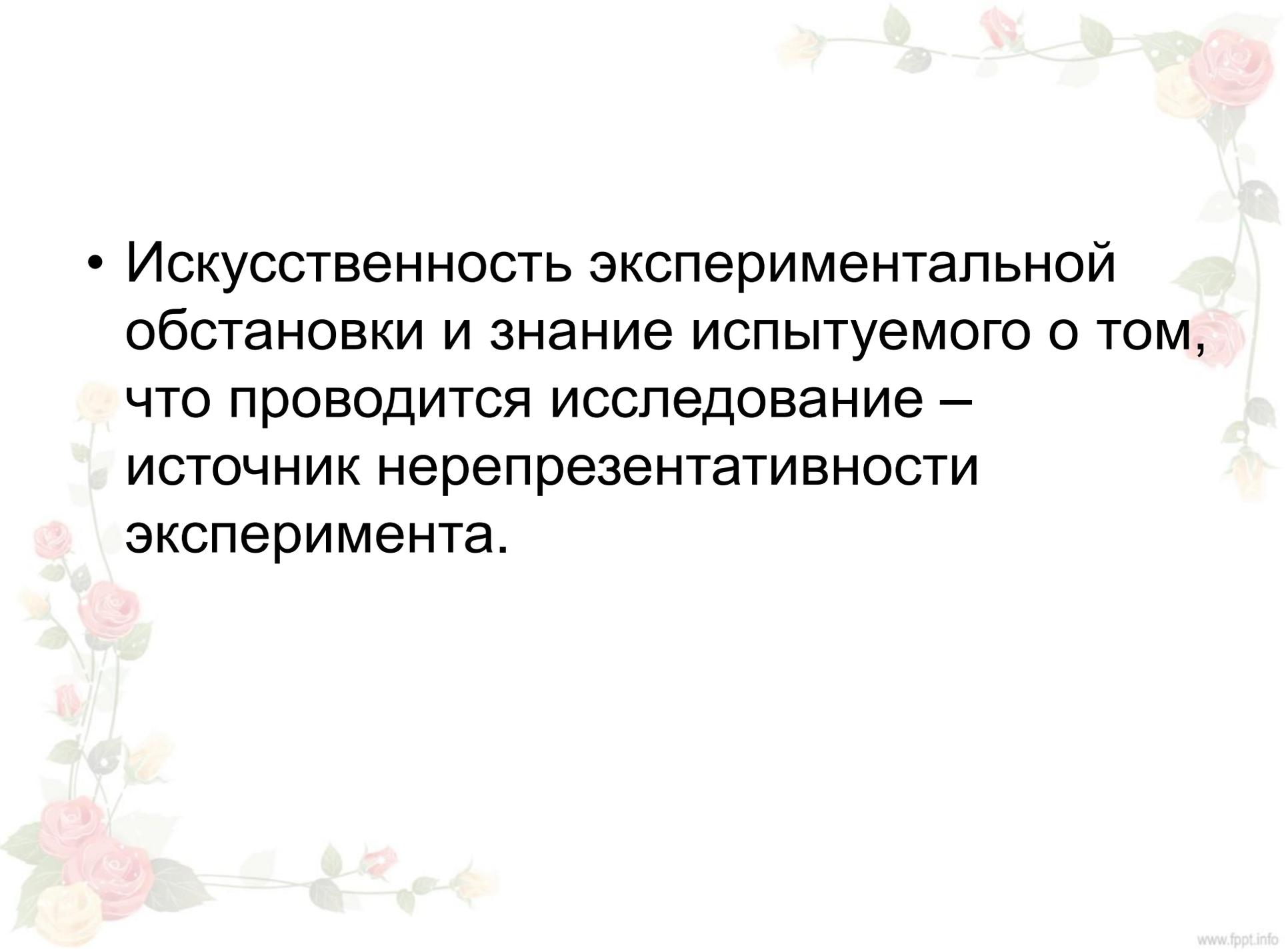


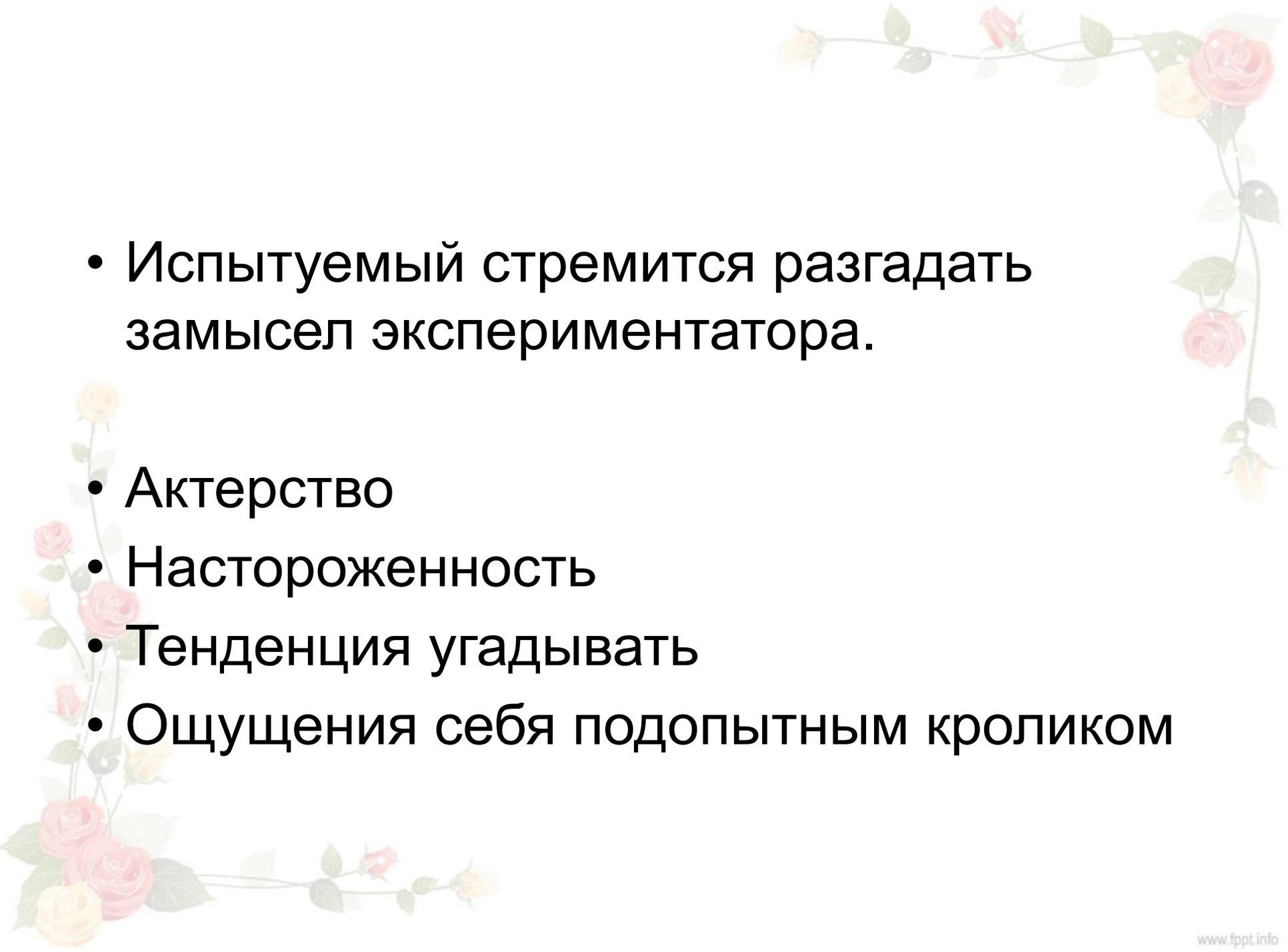
Взаимодействие X с фоном означало бы, что эффект наблюдается только в данных условиях.

Воспроизведение результатов при разных условиях и в разное время повышает уверенность в обобщении, уменьшая вероятность фона и X .

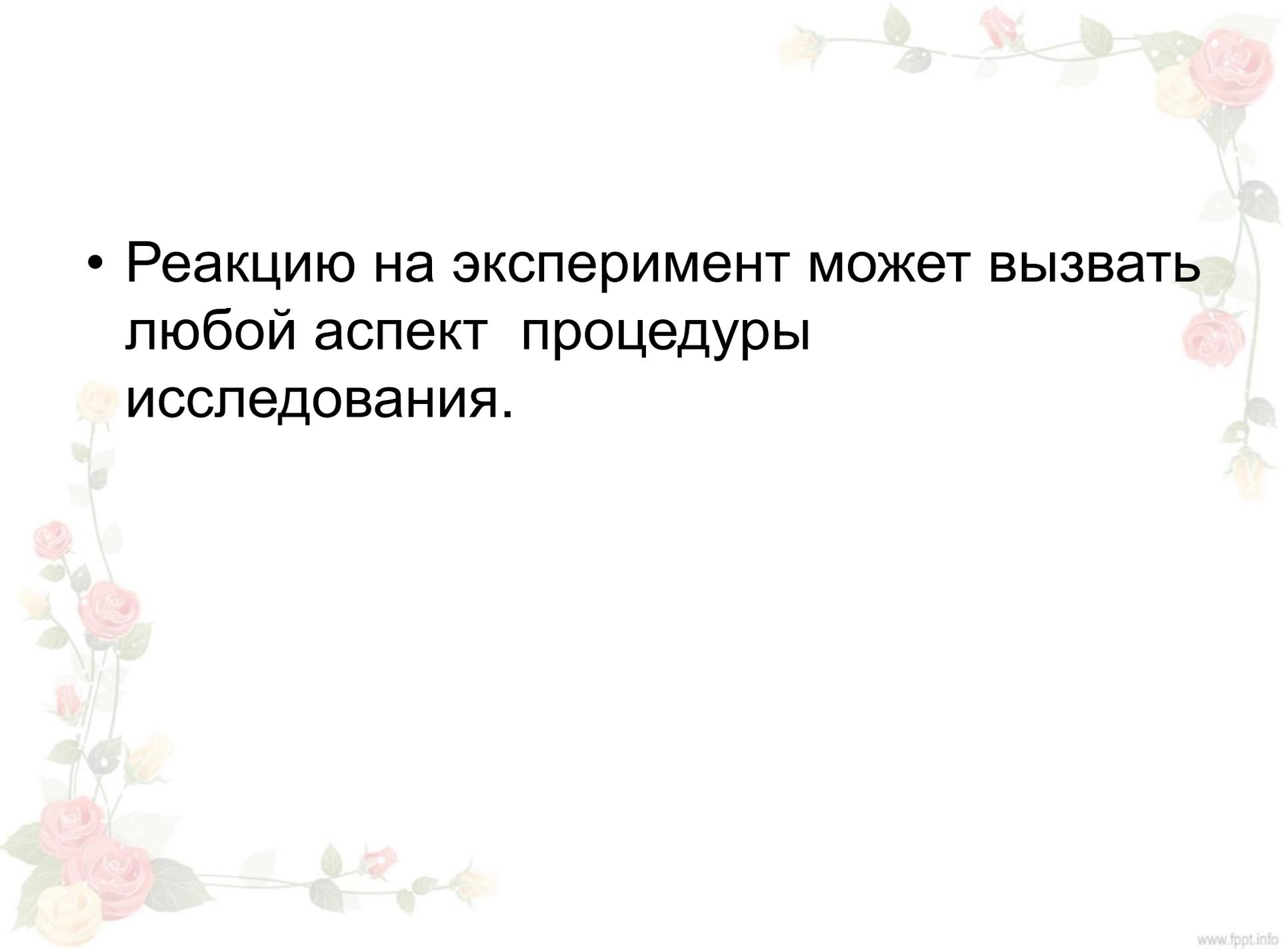
Реакция на эксперимент

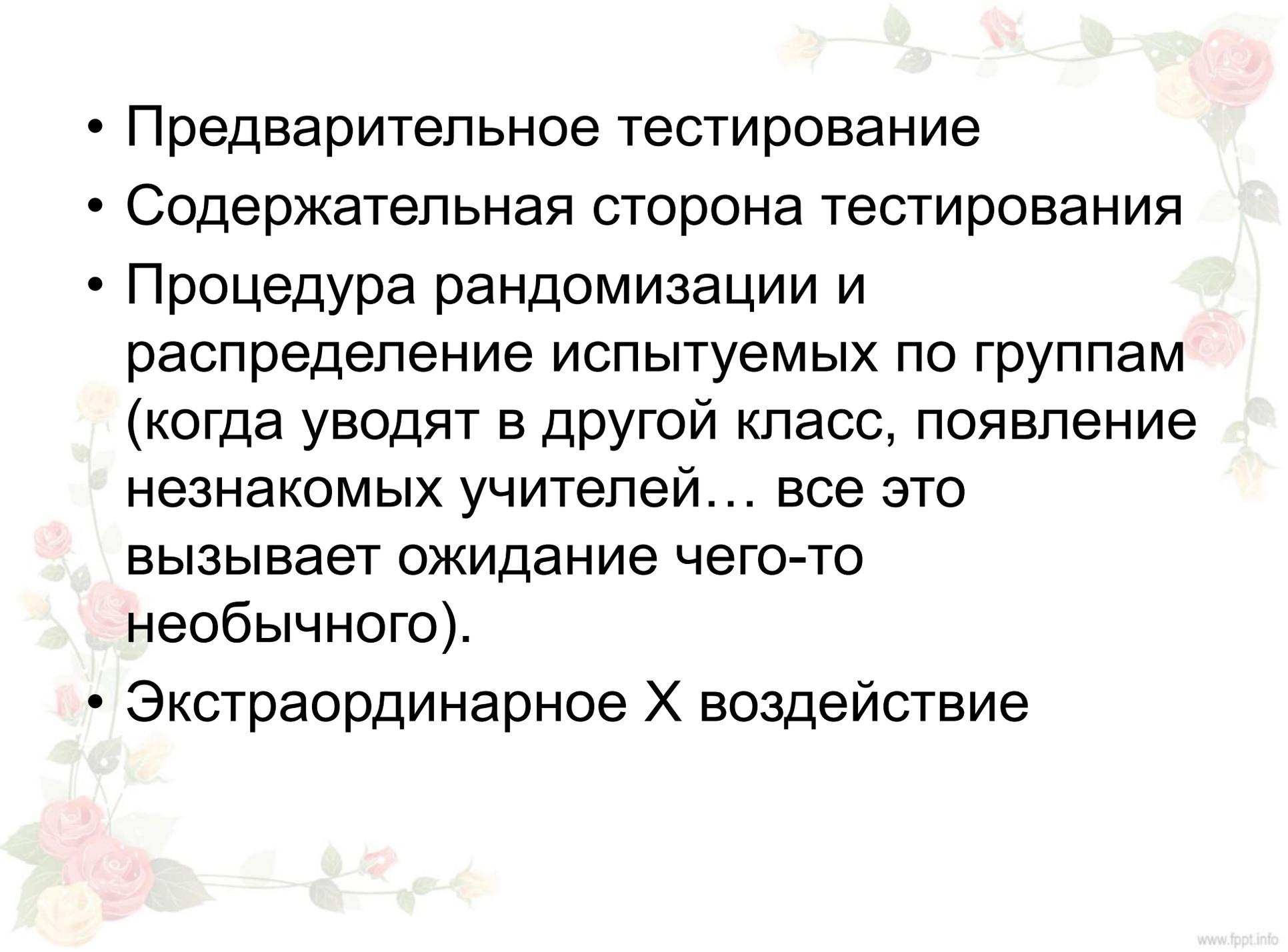


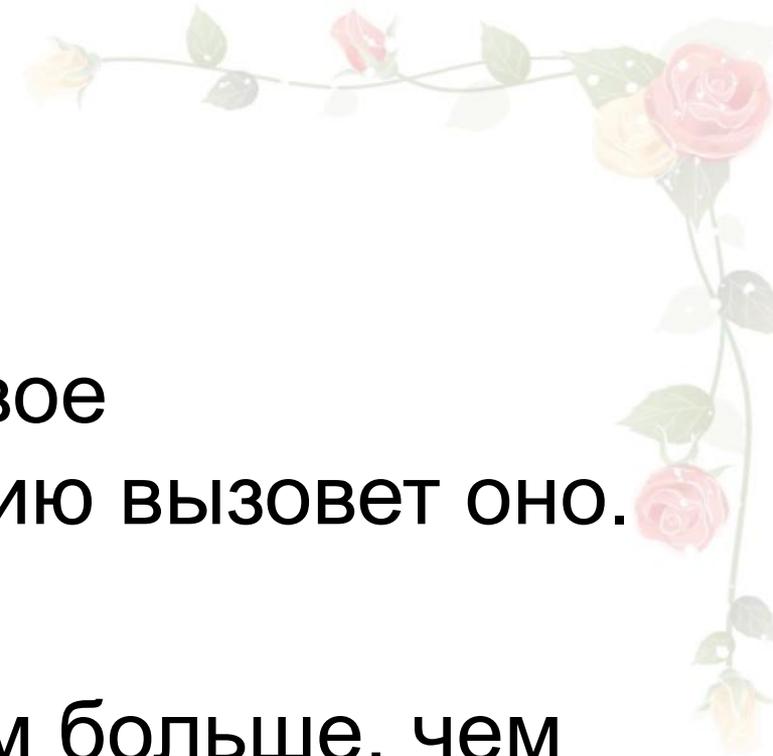
- 
- Искусственность экспериментальной обстановки и знание испытуемого о том, что проводится исследование – источник нерепрезентативности эксперимента.

- 
- Испытуемый стремится разгадать замысел экспериментатора.
 - Актерство
 - Настороженность
 - Тенденция угадывать
 - Ощущения себя подопытным кроликом

- Все это не типично для нормальной жизни и примешиваясь к X затрудняет обобщение результатов.

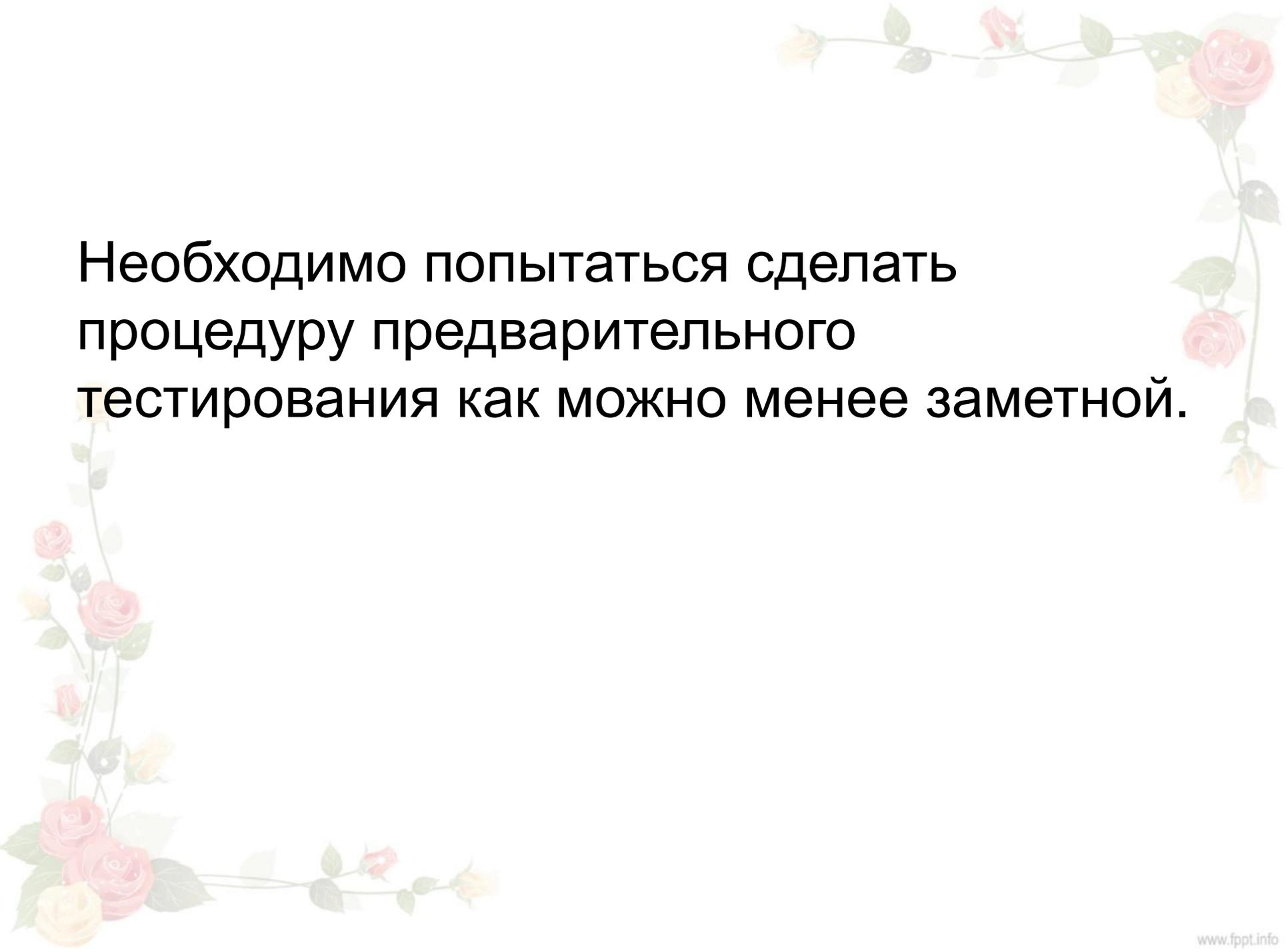
- 
- Реакцию на эксперимент может вызвать любой аспект процедуры исследования.

- 
- Предварительное тестирование
 - Содержательная сторона тестирования
 - Процедура рандомизации и распределение испытуемых по группам (когда уводят в другой класс, появление незнакомых учителей... все это вызывает ожидание чего-то необычного).
 - Экстраординарное X воздействие

- 
- Если есть только итоговое тестирование, то реакцию вызовет оно.



Вероятность эффекта тем больше, чем очевиднее связь между экспериментальным воздействием и характером тестирования.

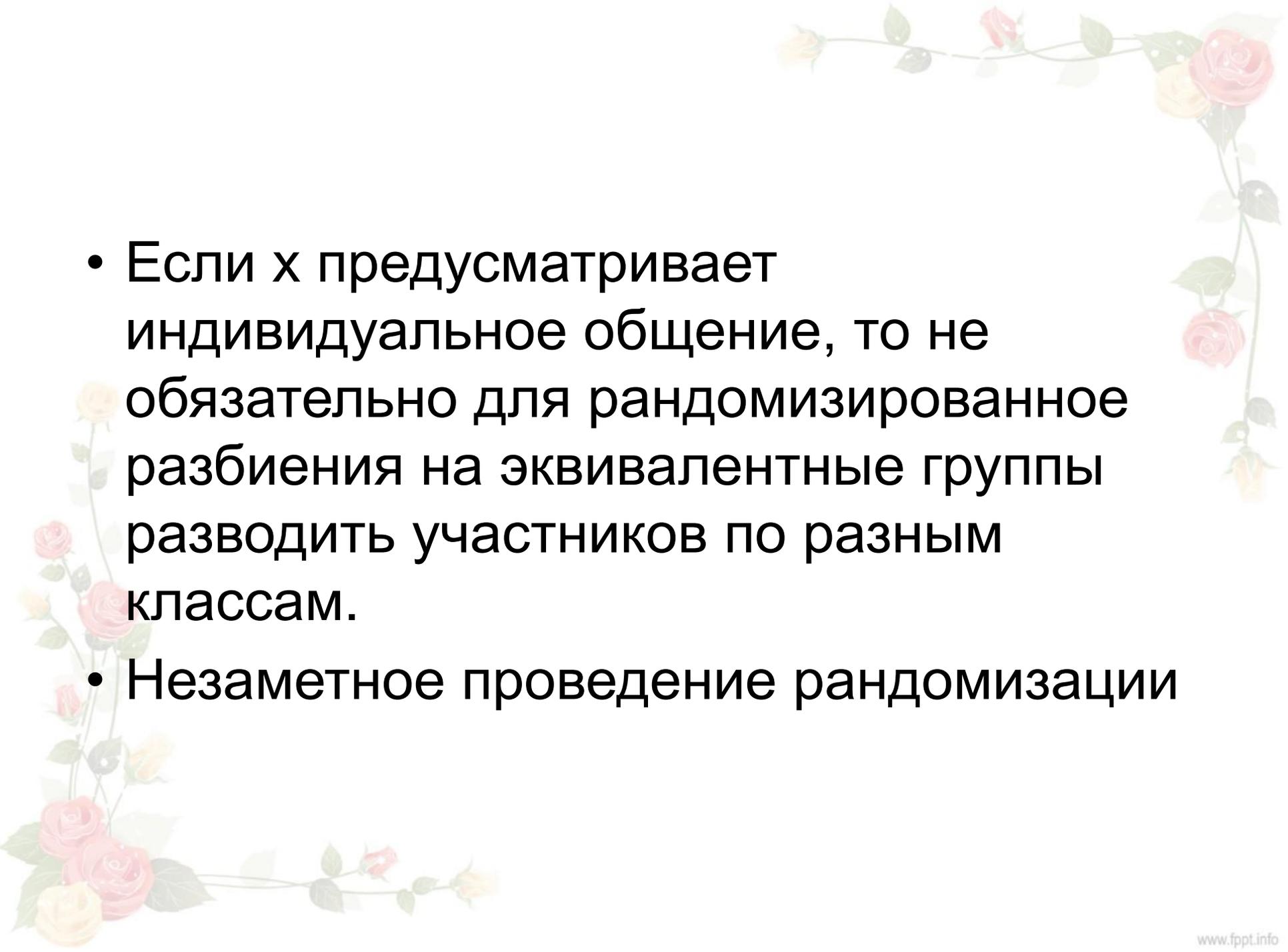


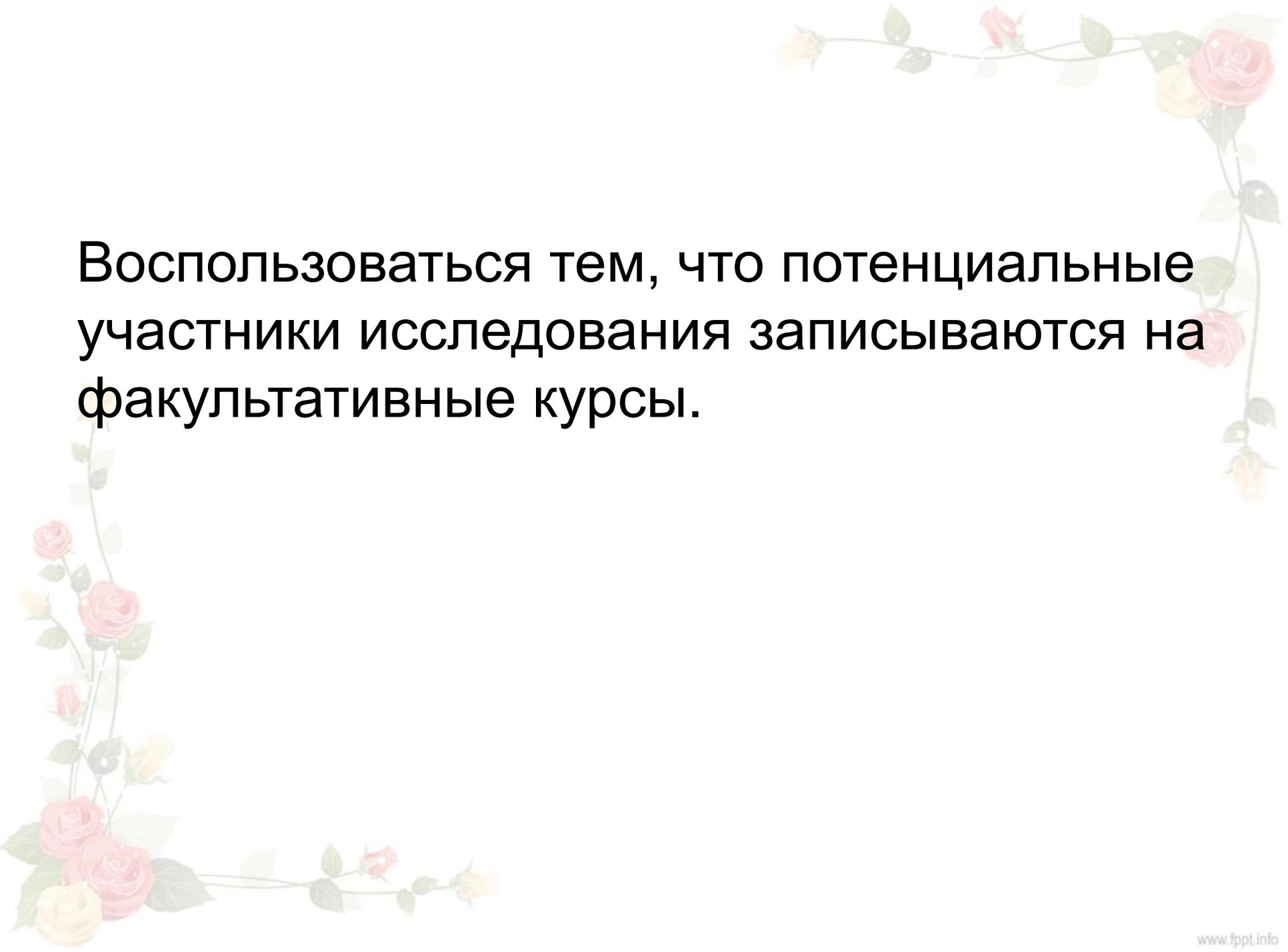
Необходимо попытаться сделать
процедуру предварительного
тестирования как можно менее заметной.

- В процедуре обучение можно скрыть изучение новых способов преподавания.

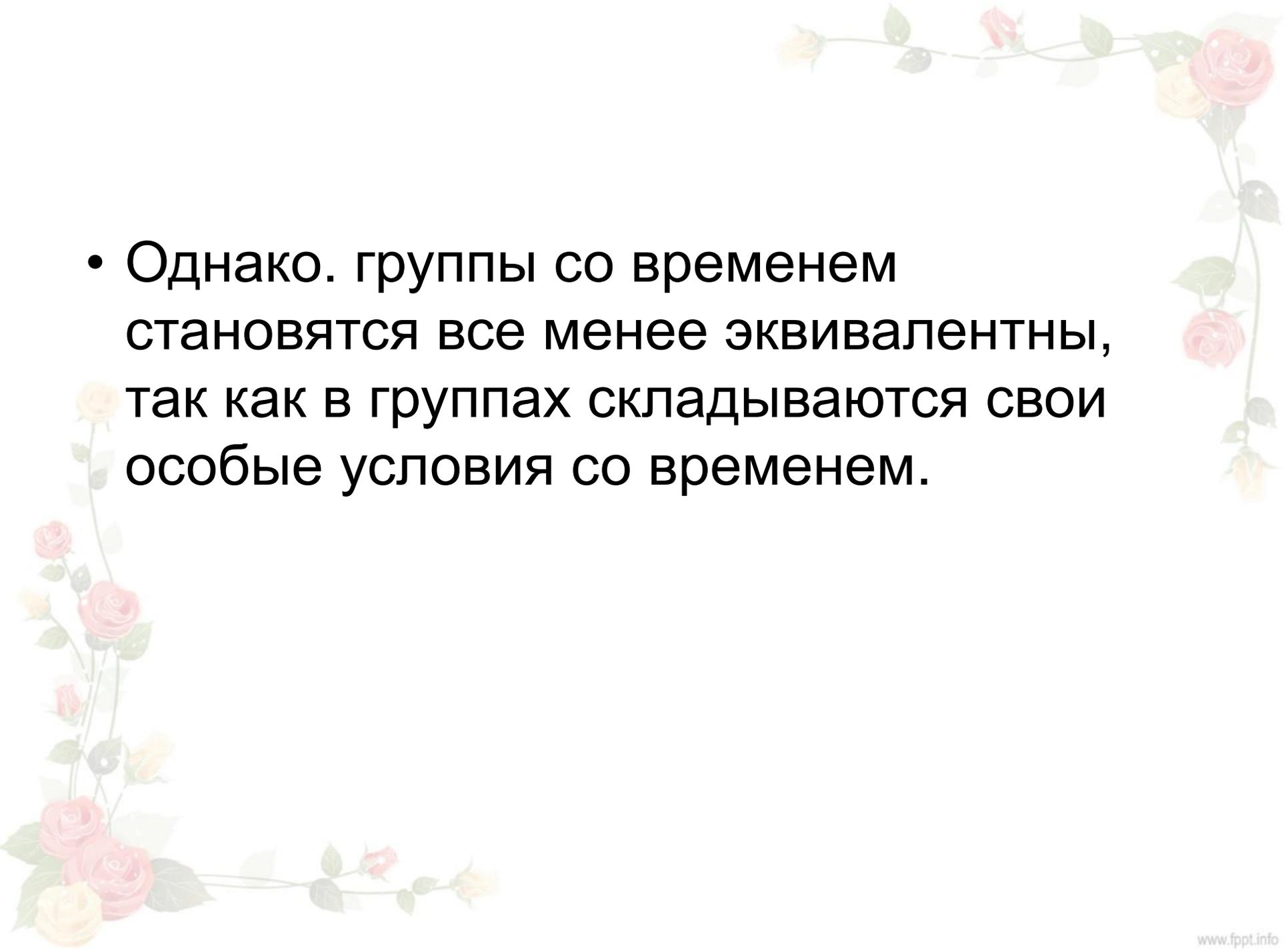
маскировка

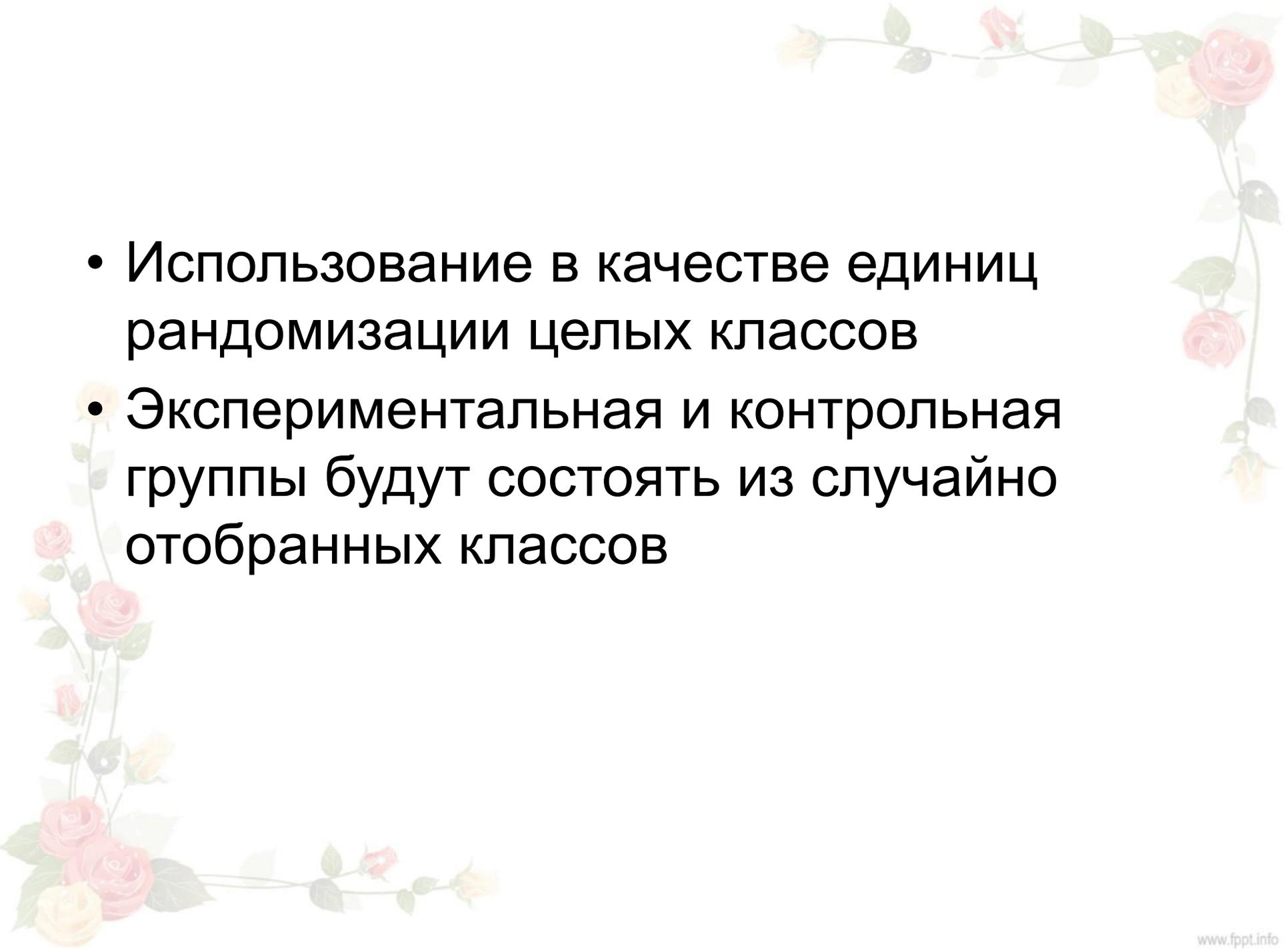
- Если эксперименты – это варианты обычных мероприятий, проводимых в соответствующее время.
- Не оповещать о введении экспериментальных воздействий.
- О можно проводить под видом регулярных контрольных работ

- 
- Если x предусматривает индивидуальное общение, то не обязательно для рандомизированное разбиения на эквивалентные группы разводить участников по разным классам.
 - Незаметное проведение рандомизации



Воспользоваться тем, что потенциальные участники исследования записываются на факультативные курсы.

- 
- Однако. группы со временем становятся все менее эквивалентны, так как в группах складываются свои особые условия со временем.

- 
- Использование в качестве единиц рандомизации целых классов
 - Экспериментальная и контрольная группы будут состоять из случайно отобранных классов

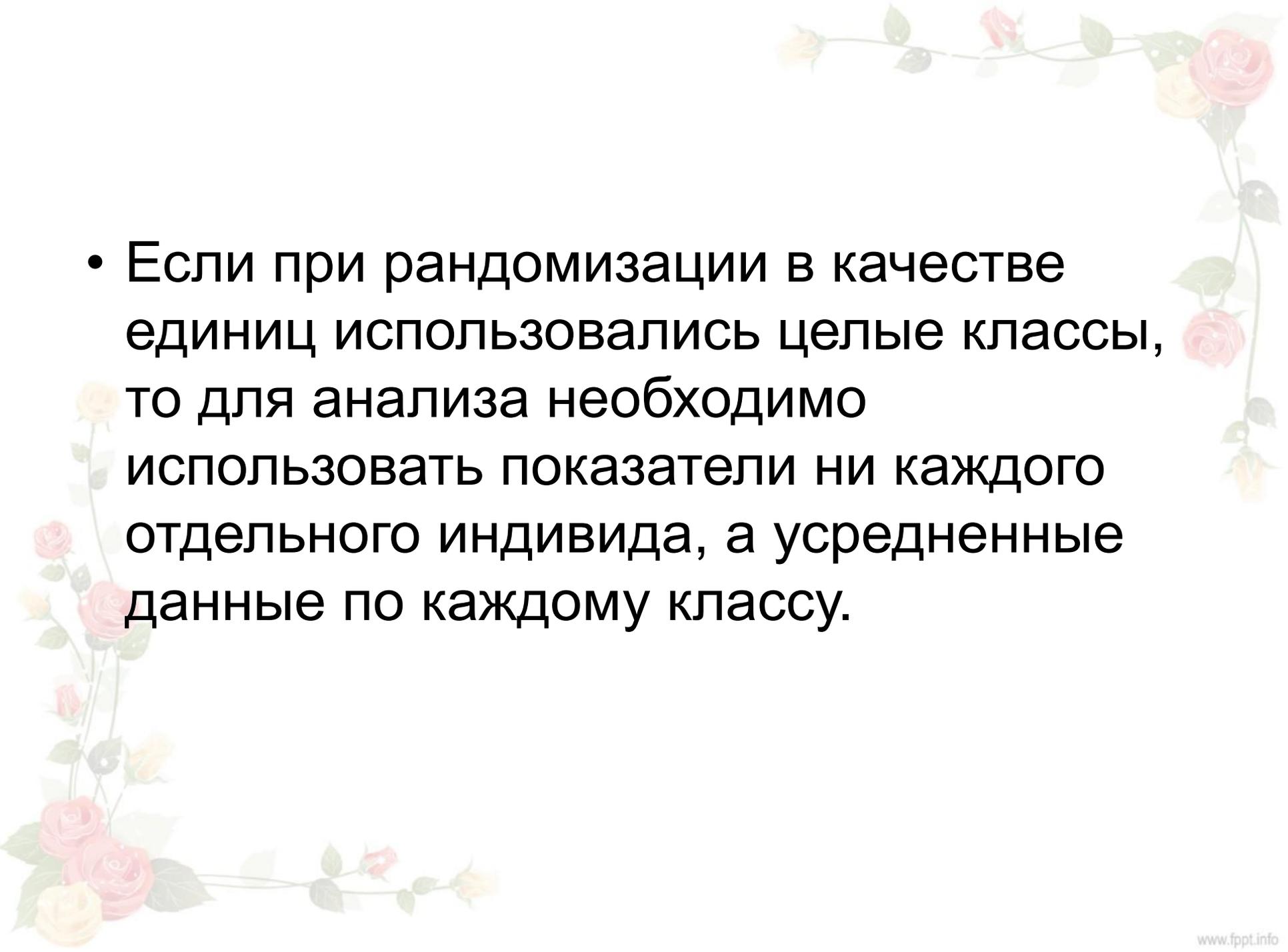
- Нельзя объединять данные по всем учащимся поскольку они не подлежали рандомизации.

Статистический анализ

- 1. Вычисляются приращения (разница между показателями тестирования до и после X).
- 2. Для сравнение приращений используется t-критерий

ковариационный анализ

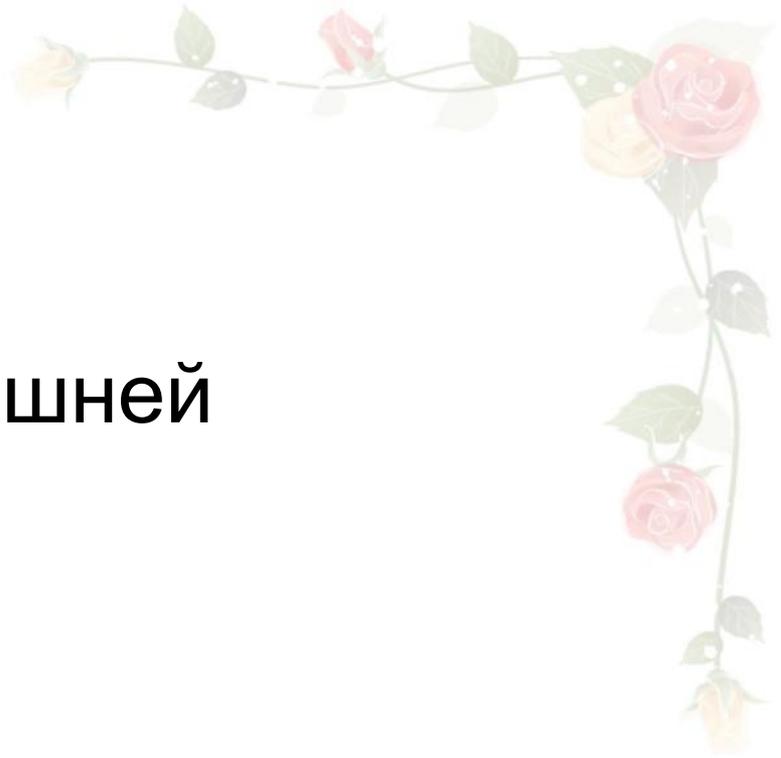


- 
- Если при рандомизации в качестве единиц использовались целые классы, то для анализа необходимо использовать показатели не каждого отдельного индивида, а усредненные данные по каждому классу.

План Соломона для 4-х групп

R	O ₁	X	O ₂
R	O ₃		O ₄
R		X	O ₅
R			O ₆

- Учитывает факторы внешней валидности

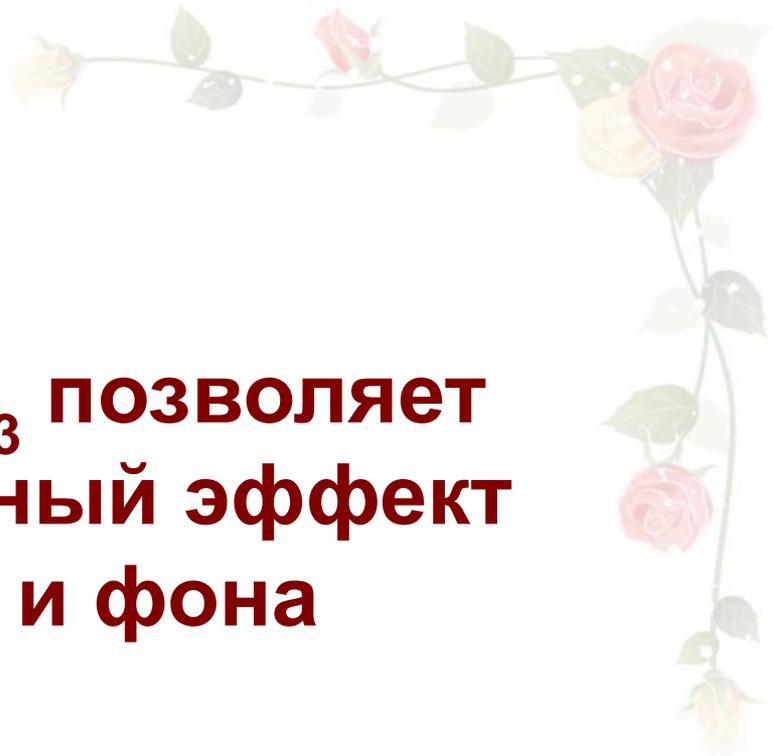


- Можно оценить не только эффект тестирования, но и его взаимодействие с X.

Эффект Х воспроизводится 4-мя различными способами

- $O_2 > O_1$; $O_2 > O_4$; $O_5 > O_6$; $O_5 > O_3$

- Если эти 4 соотношения выполняются, то правомерность вывода возрастает.



**Сравнение O_6 с O_1 и O_3 позволяет
выявить комбинированный эффект
естественного развития и фона**

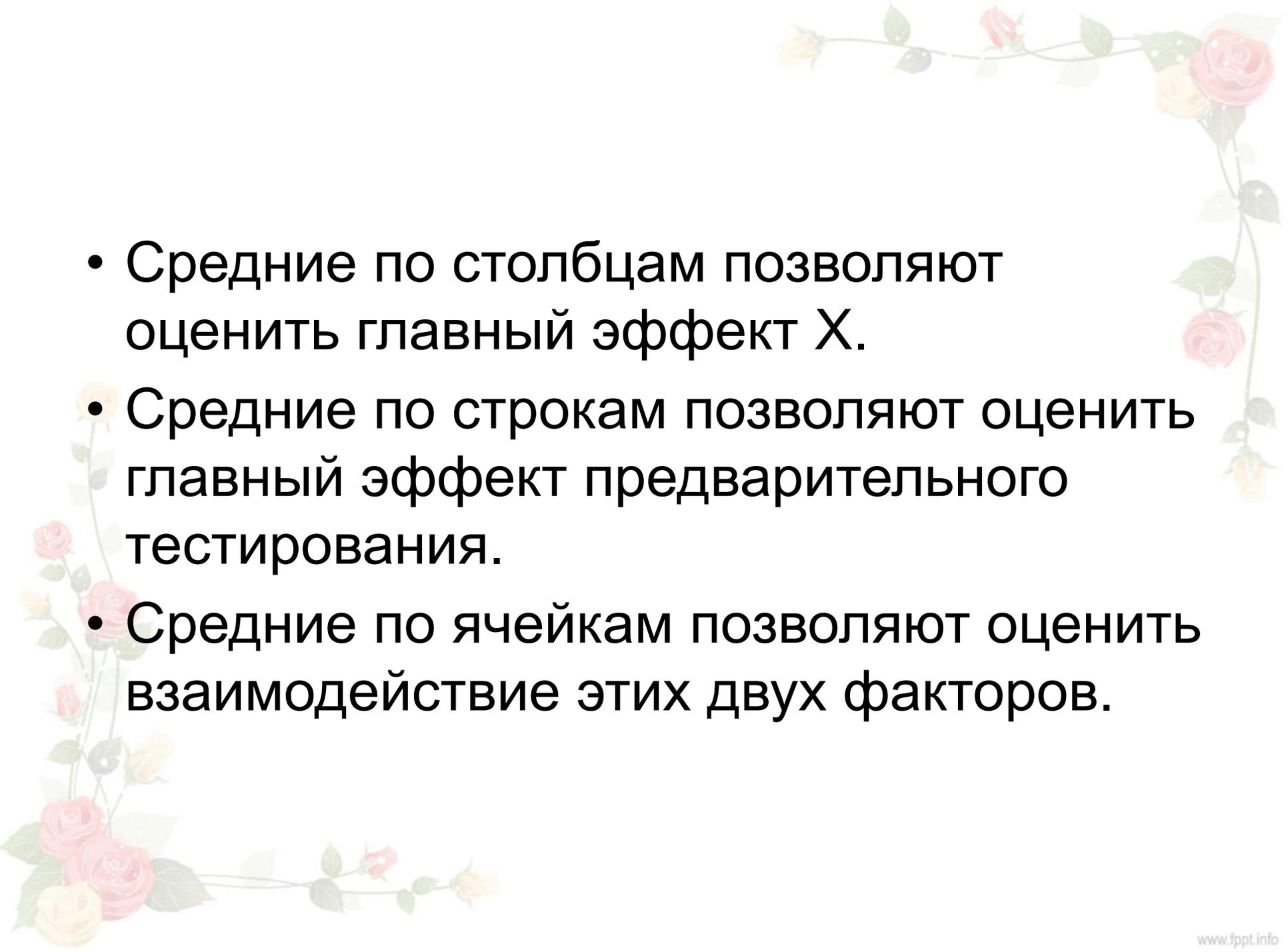


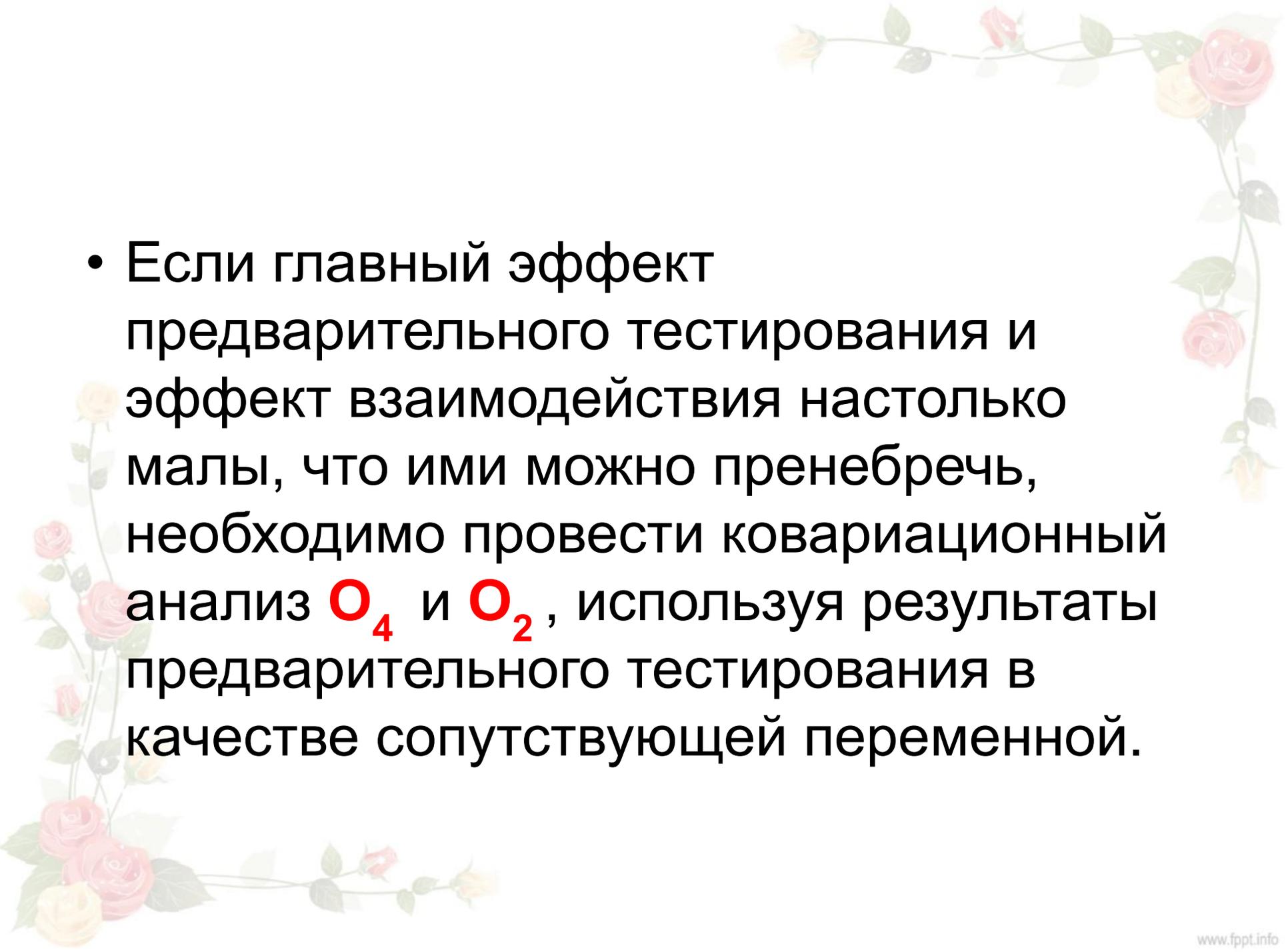
Статистический анализ

- Проводить дисперсионный анализ приращений невозможно.

Можно провести дисперсионный анализ 2x2.

X	Да	Нет
Предварительное тестирование		
Да	O₂	O₄
Нет	O₅	O₆

- 
- Средние по столбцам позволяют оценить главный эффект X .
 - Средние по строкам позволяют оценить главный эффект предварительного тестирования.
 - Средние по ячейкам позволяют оценить взаимодействие этих двух факторов.

- 
- Если главный эффект предварительного тестирования и эффект взаимодействия настолько малы, что ими можно пренебречь, необходимо провести ковариационный анализ O_4 и O_2 , используя результаты предварительного тестирования в качестве сопутствующей переменной.

контролируется

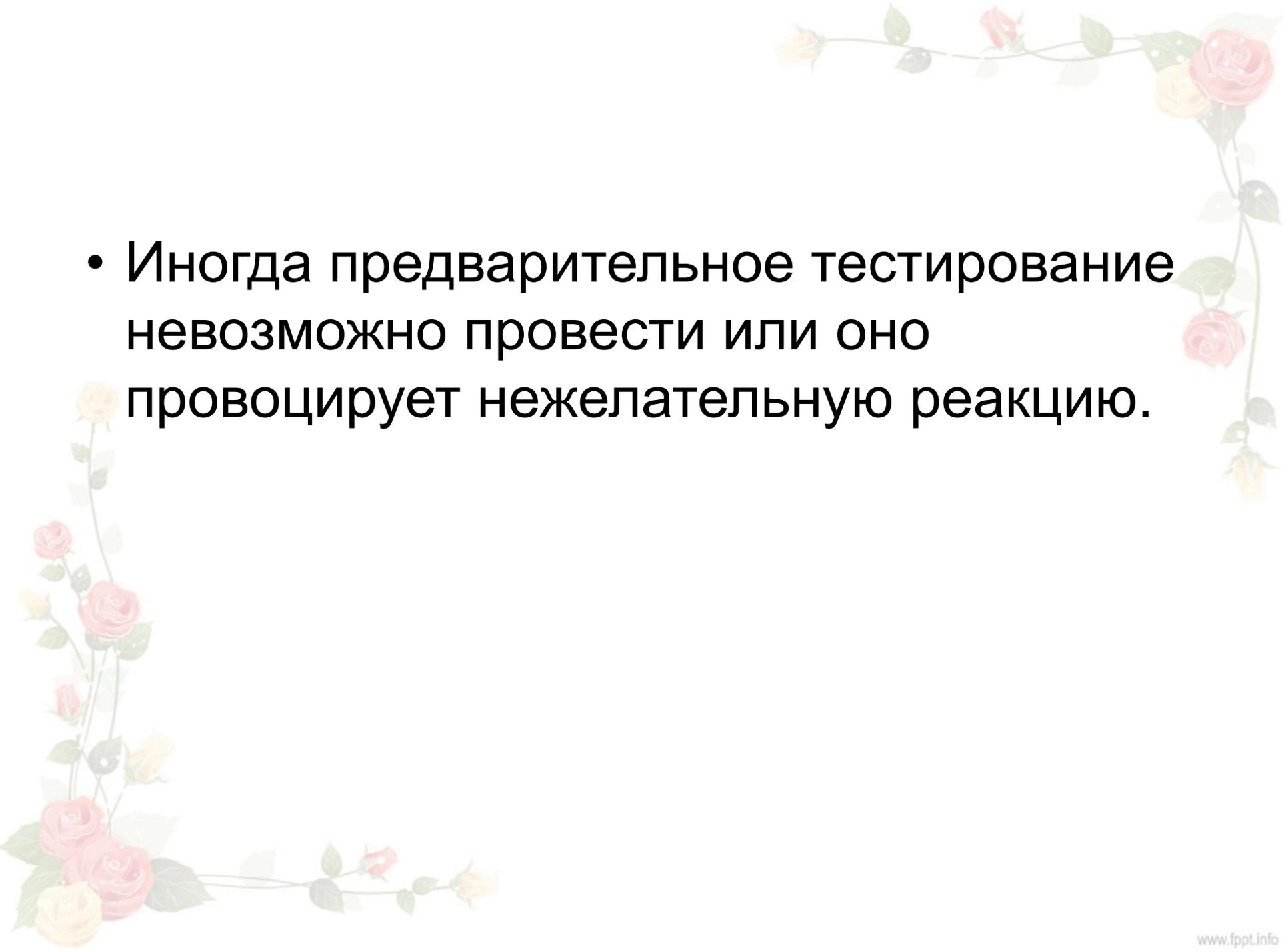
- + Фон
- + Естественное развитие
- + Эффект тестирования
- + Инструментальная погрешность
- + Регрессия
- + Состав групп
- + Выбывание
- + Взаимодействие состава групп с естественным развитием

Источники внутренней невалидности не выявлены

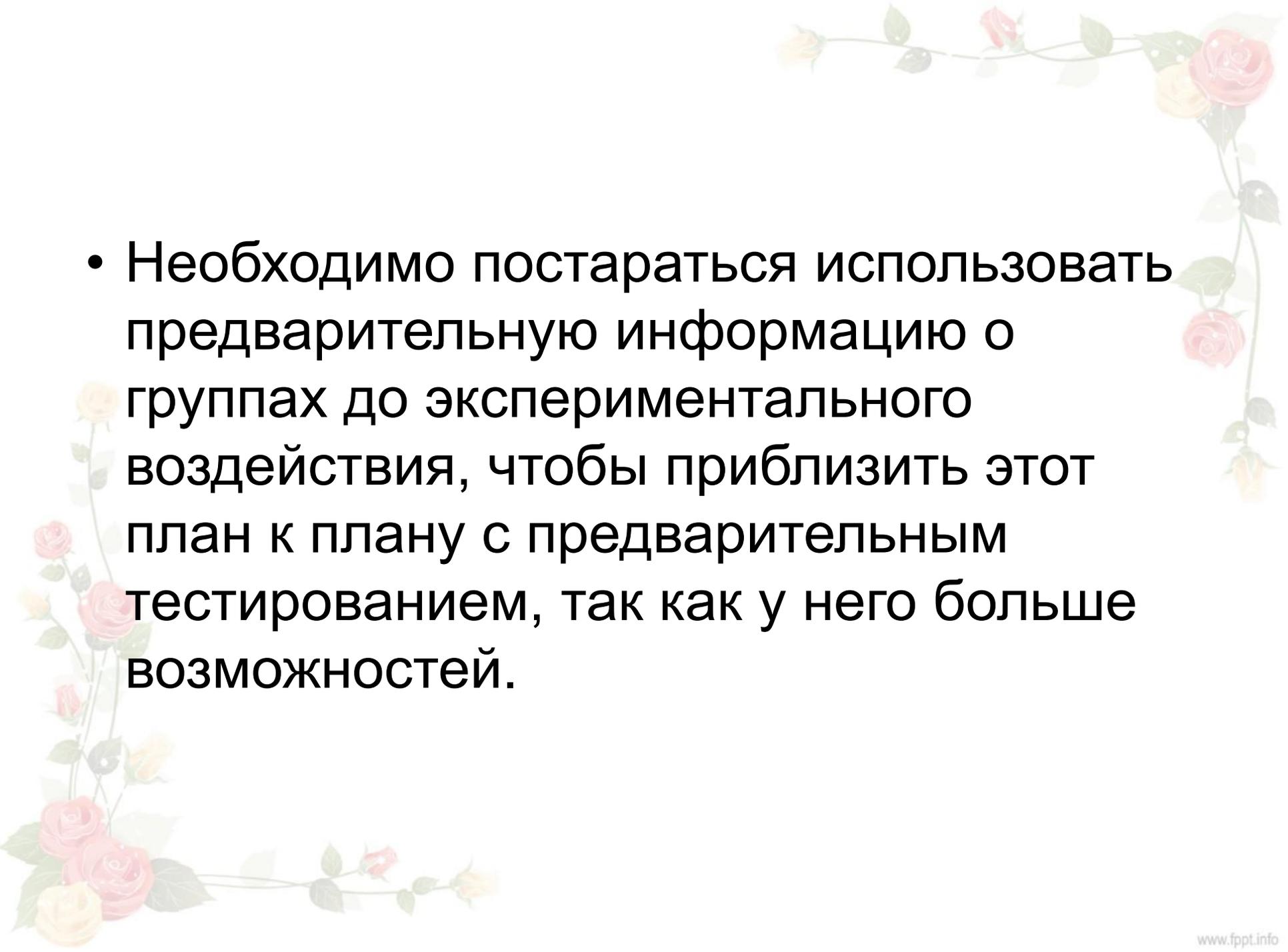
План с контрольной группой и тестированием только после воздействия

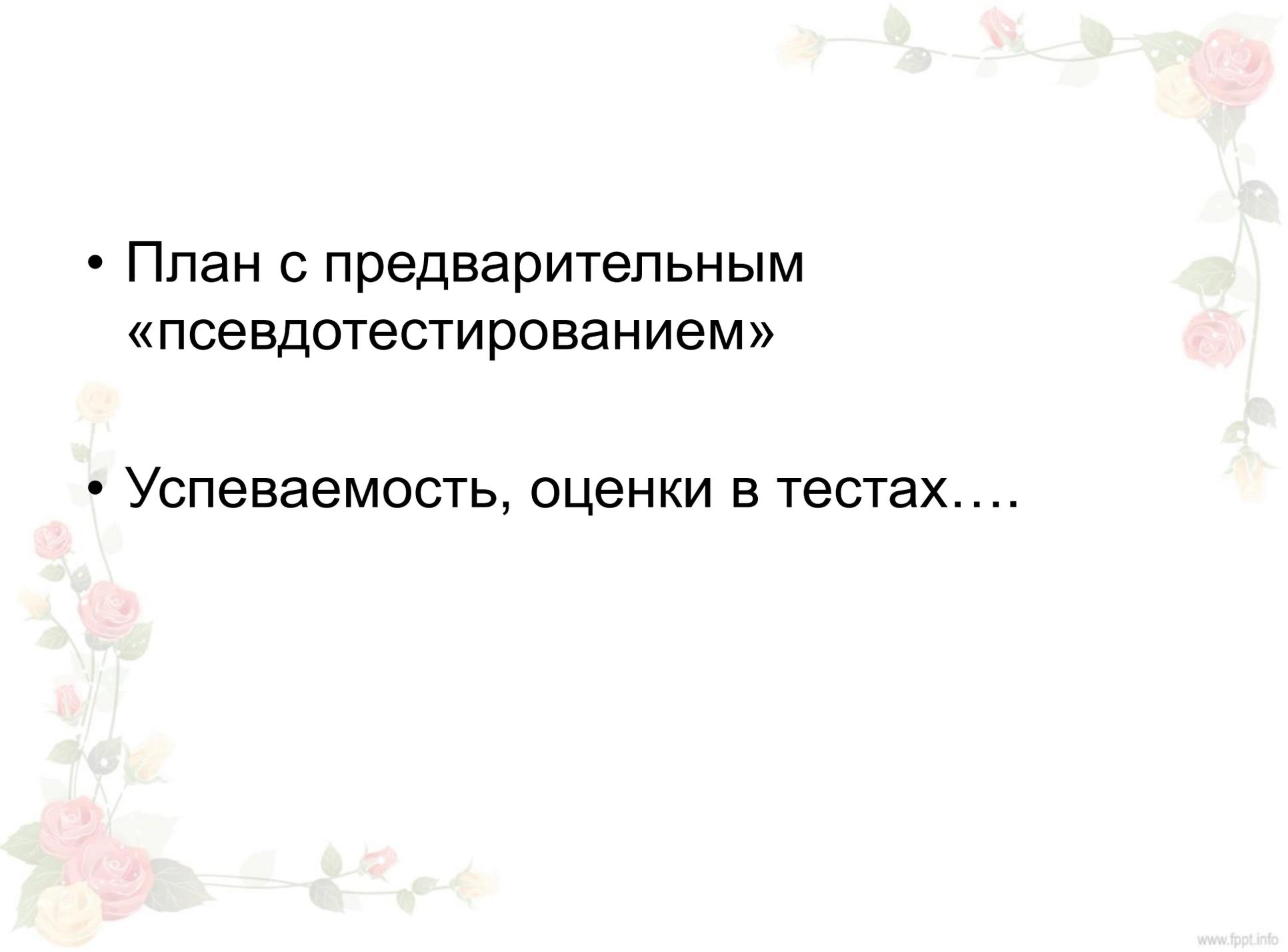
R X O
R O

Рандомизация является гарантией
отсутствия смещения между группами.

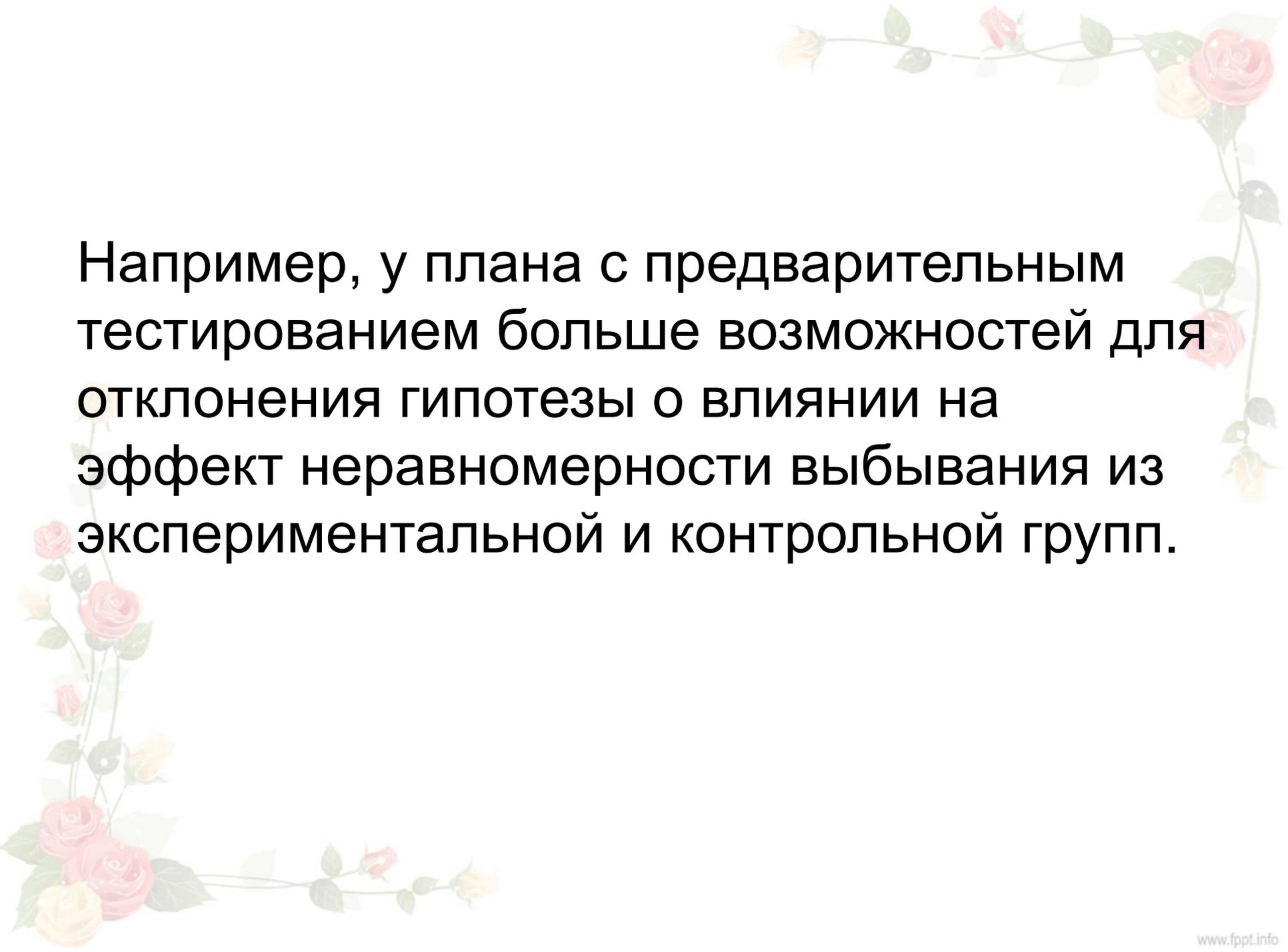
- 
- Иногда предварительное тестирование невозможно провести или оно провоцирует нежелательную реакцию.

- Контролируется эффект тестирования и его взаимодействие с X, но при этом они не измеряются.

- 
- Необходимо постараться использовать предварительную информацию о группах до экспериментального воздействия, чтобы приблизить этот план к плану с предварительным тестированием, так как у него больше возможностей.

- 
- План с предварительным «псевдотестированием»
 - Успеваемость, оценки в тестах....

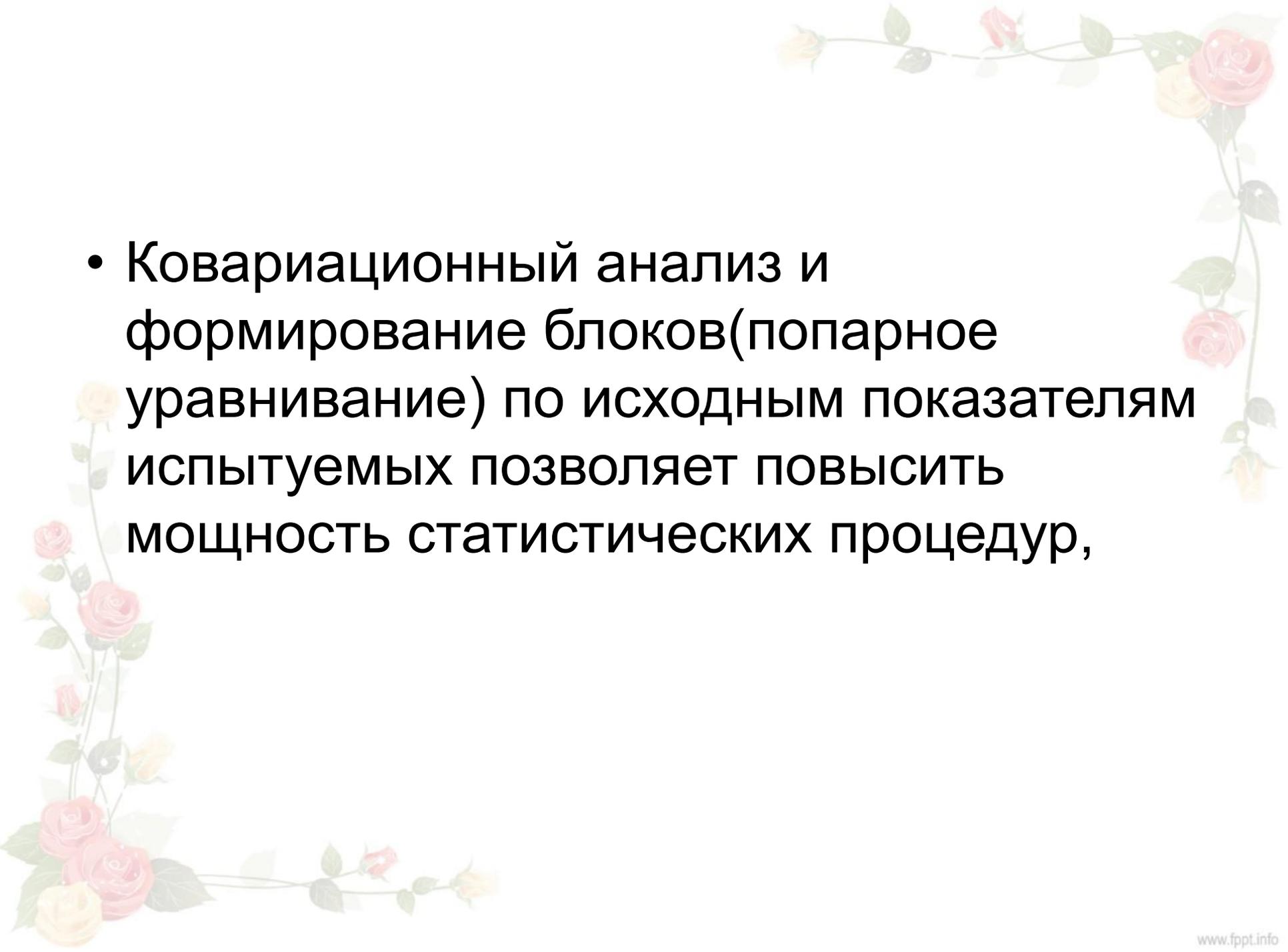
- Такая информация позволит проверить наличие взаимодействия X с начальным уровнем.



Например, у плана с предварительным тестированием больше возможностей для отклонения гипотезы о влиянии на эффект неравномерности выбывания из экспериментальной и контрольной групп.

Статистический анализ

Для этого плана применим t-критерий

- 
- Ковариационный анализ и формирование блоков (попарное уравнивание) по исходным показателям испытуемых позволяет повысить мощность статистических процедур,

- Оценка этого плана в отношении реакции не эксперимент выше, чем у других планов.

контролируется

- + Фон
- + Естественное развитие
- + Эффект тестирования
- + Инструментальная погрешность
- + Регрессия
- + Состав групп
- + Выбывание
- + Взаимодействие состава групп с естественным развитием

Источники внутренней невалидности не выявлены

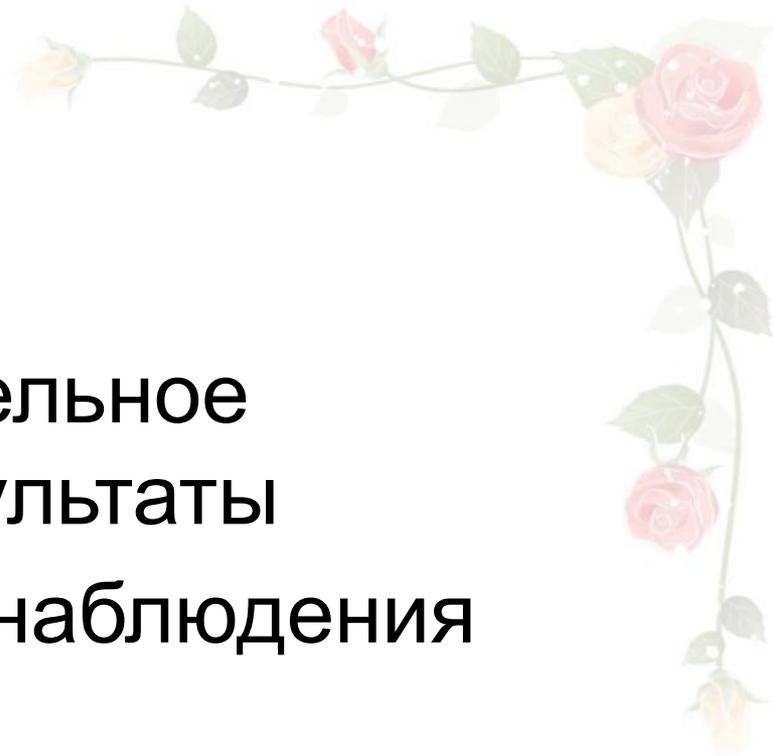
Источники внешней невалидности

❖ взаимодействие тестирования и х

Эффект предварительного тестирования

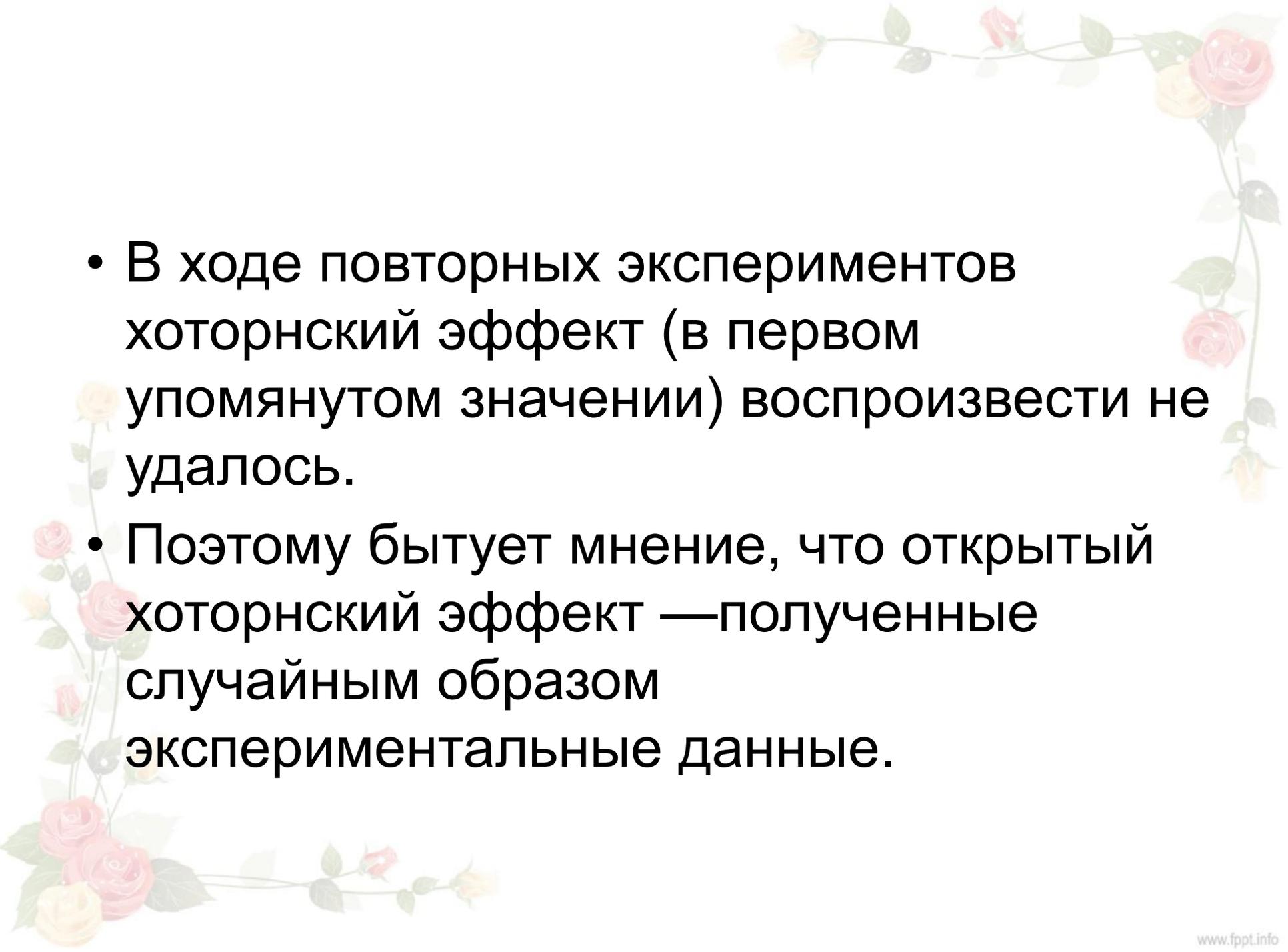
- хоторнский эффект(Hawthorne effect)
- Observation bias

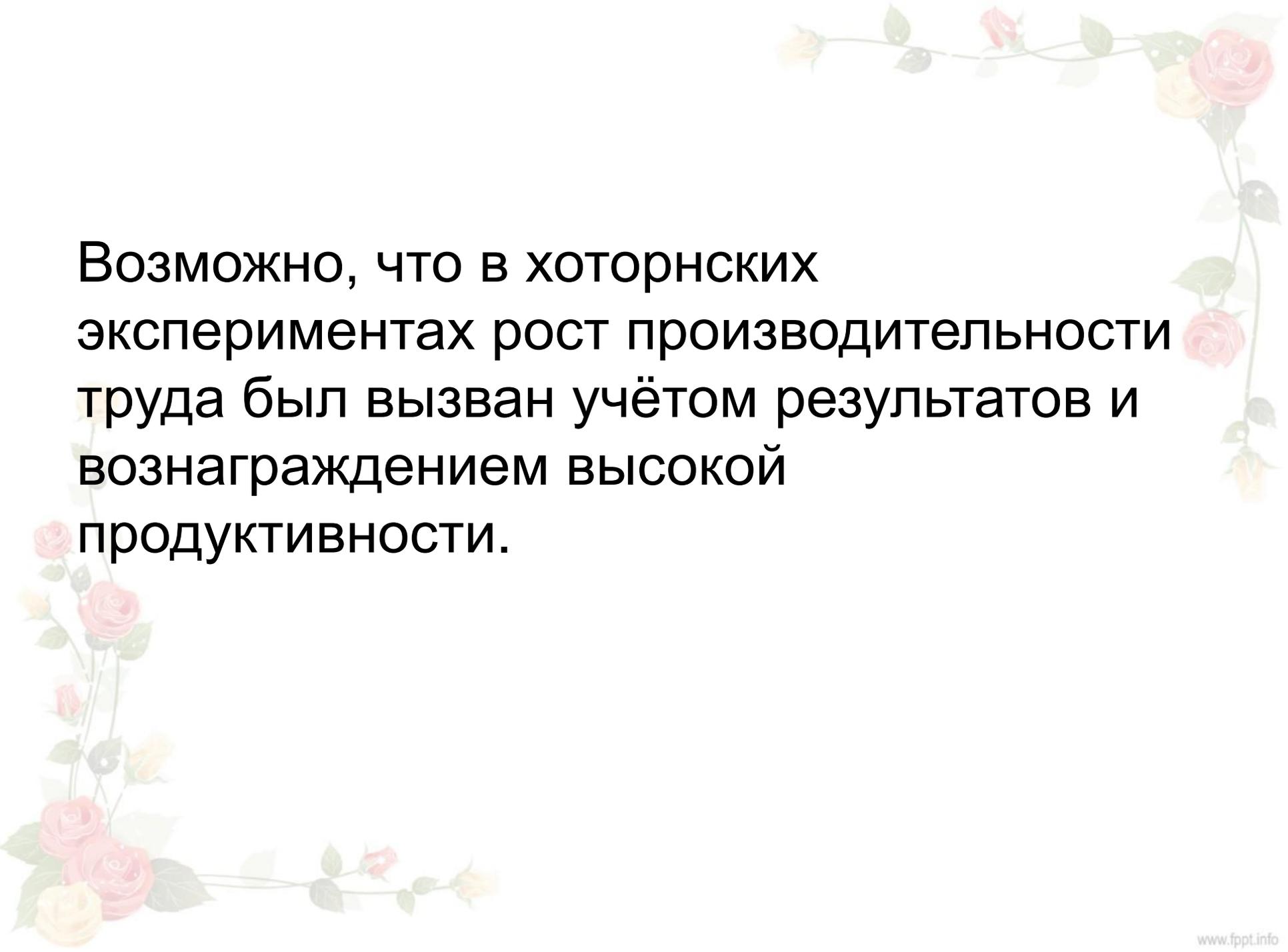




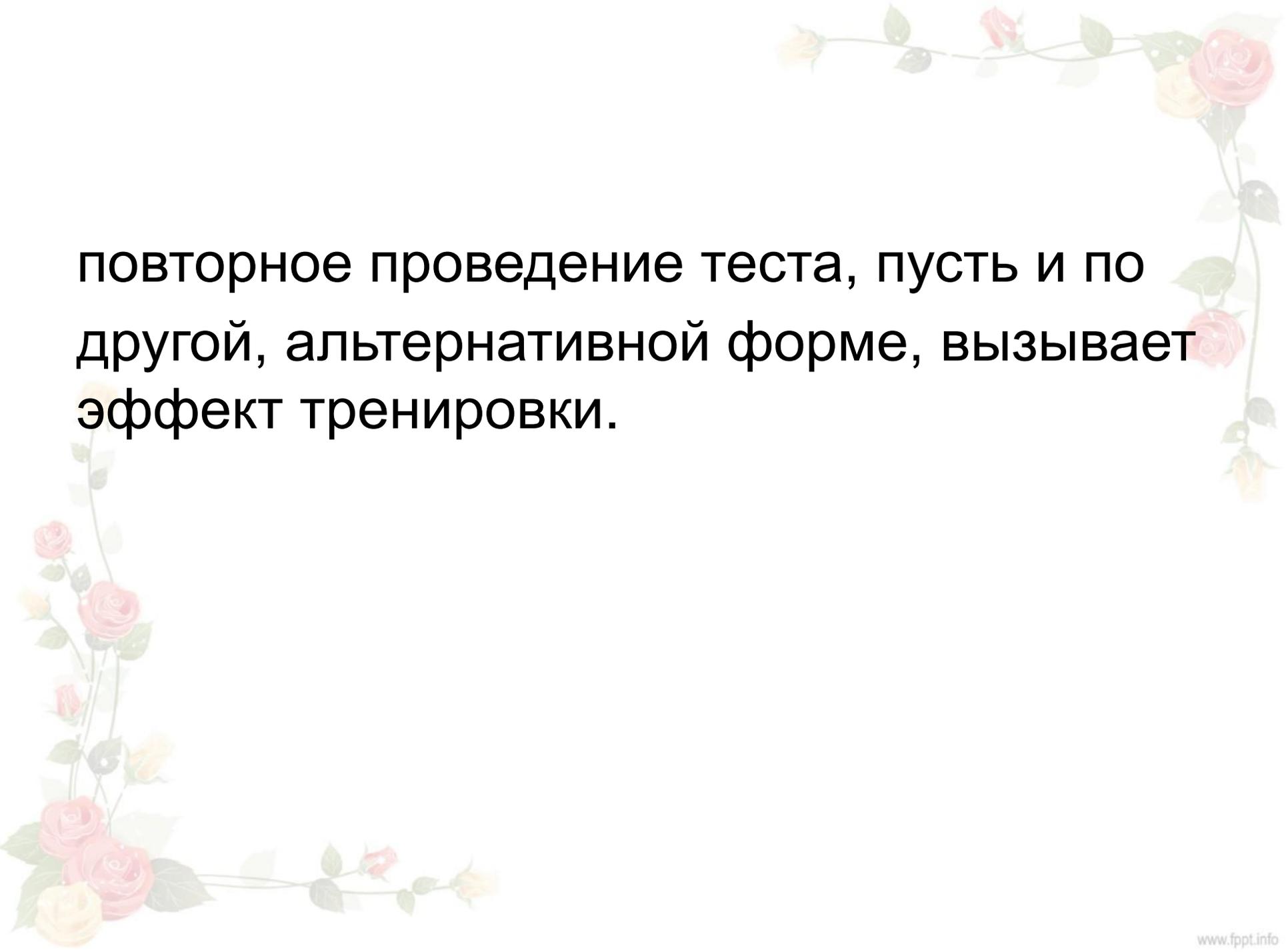
1. Внимание, благожелательное отношение улучшает результаты
2. Эффект от процедуры наблюдения



- 
- В ходе повторных экспериментов хоторнский эффект (в первом упомянутом значении) воспроизвести не удалось.
 - Поэтому бытует мнение, что открытый хоторнский эффект — полученные случайным образом экспериментальные данные.



Возможно, что в хоторнских экспериментах рост производительности труда был вызван учётом результатов и вознаграждением высокой продуктивности.



повторное проведение теста, пусть и по другой, альтернативной форме, вызывает эффект тренировки.

