True Experiments

Cause and Effect Relationships

- Can be clear relationship but not a causal one – e.g., between getting dressed up and having a headache the next morning
- Four basic elements to establish cause and effect:
  1. Manipulation
  2. Measurement
  3. Comparison
  4. Control

Defining Characteristics of the Experiment

- Manipulation of the independent variable
  - Experimenter creates the conditions to be studied
  - If IV is measured rather than manipulated, it is *not* a true experiment
- Holding all other variables constant (Control)
  - Participants in all conditions have as similar an experience as possible, except for IV

Defining Characteristics of the Experiment

- Ensuring that participants in all conditions
  - have equivalent personal characteristics
  - are equivalent with respect to the DV before taking part in the study
  - attained by
    - holding the characteristics constant
    - random assignment to condition
    - matching experimental and control participants on selected characteristics
    - having the same participants take part in all conditions

Forms of Experiment

- In its simplest form, an experiment has two conditions
- To test how people behave in presence of treatment, researcher compares
  - experimental condition, which represents IV
  - control condition, which provides a baseline against which effect of treatment is assessed
- To compare the effects of two treatments, two comparison conditions are used

Mill's Methods

Experimental Group	Control Group
If *X*, then *Y*	If not-*X*, then not-*Y*
(Method of Agreement)	(Method of Difference)

- Method of agreement
  - If *X (drink)*, then *Y (headache)*
  - *X* is a sufficient condition of *Y*
- Method of difference
  - If not–*X*, then not–*Y*
  - *X* is a necessary condition of *Y*

When Are More Than Two Experimental Conditions Needed?

- When hypothesis being tested has more than one component
- When the baseline control condition is not sufficient to rule out all alternative explanations
- E.g., Does watching violence on TV cause more aggression?
  - Would violent TV versus no TV suffice?

Characteristics of a Good Manipulation

Has good construct validity

- The manipulation accurately represents the construct
- E.g., Does having confederate flirt with partner create threat of infidelity?

Two Forms of Manipulation Checks

- Post-experimental interview
  - Participants are asked questions that indicate whether manipulation had intended effect
  - Checking it against a measure of the construct
    - E.g., those in infidelity threat condition worried more about infidelity in self report measure
- DVs that assess the construct being manipulated
  - Ensuring that participants in different experimental conditions have different experiences (related to IV)
    - E.g., Participants primed for aggression should be more likely to complete word strings with synonyms of aggression than participants primed for altruism
    - E.g., In study where a perpetrator's motives are manipulated, participants report what they believed the motive to be
    - If participants report a different motive than the manipulation, construct validity is poor

Characteristics of a Good Manipulation

Reliability: Manipulation is applied in the same way every time

- Every participant in experiment experiences it in the same way
- Manipulations with low reliability have low validity

Characteristics of a Good Manipulation

Salience: Manipulation is noticeable

- Stands out from background
- If manipulation is complex, should be presented in different ways to ensure it is understood

Characteristics of a Good Manipulation

Strength: With a strong manipulation, the conditions of IV differentially affect behavior

- Best if extreme levels of IV are also
  - realistic
  - similar to situations in everyday life
  - ethical

Extraneous Variance

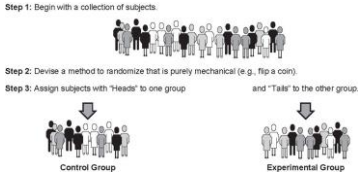
- Variance in DV not caused by IV is error variance
  - variations in how the experimenter treats the research participants
  - factors systematically related to DV, but not of interest to researcher
  - extraneous variables that are part of the research situation but can be controlled
    - Hold absolutely constant or
    - Limit to restricted range

Between-Subjects Designs

- Participants take part in only one experimental condition, so contribute only one data point (between groups)
- Data points should be independent (independent samples)

Between-Subjects Designs

- Must assure participants in different experimental groups are equivalent in their personal characteristics
- Accomplished by
  - simple random assignment

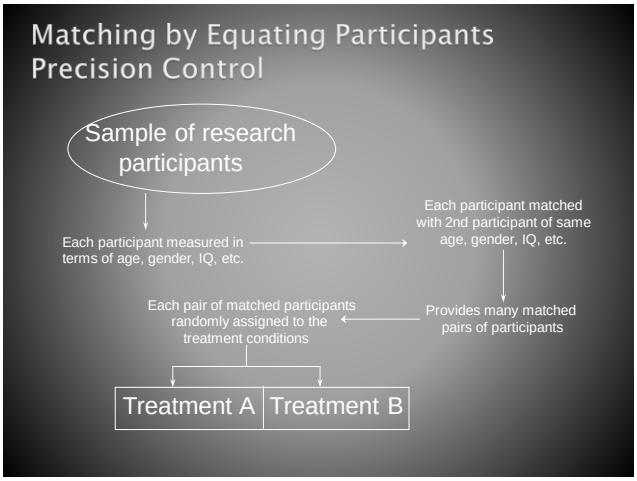


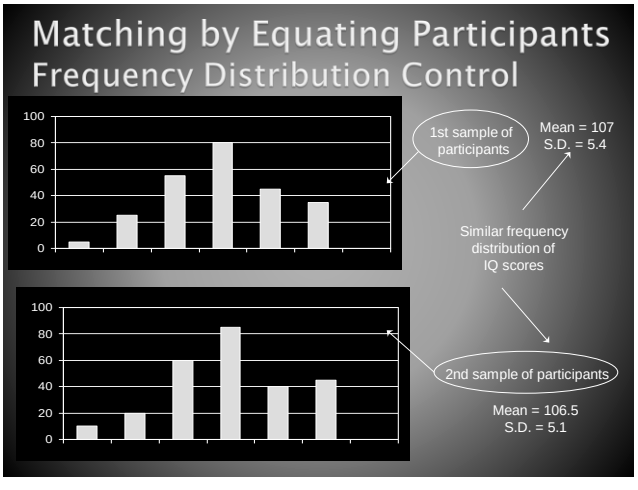
Methods of Controlling Extraneous Variables

- Randomization
  - Disrupts any systematic relation between EVs and IV – prevents EVs from becoming CVs
  - Unpredictable, unbiased procedure to distribute different values of each EV across treatment conditions
  - All possible outcomes equally likely
  - But chance CAN produce biased outcomes – e.g., all heads with 10 coin tosses

Matched Random Assignment

- Ensures that the members of the experimental and control groups are equivalent on one (or more) characteristic(s)
  - Researcher measures the characteristic to be controlled via a pretest
  - Participants are rank-ordered on their scores
  - Members of adjacent pairs of participants are randomly assigned to experimental or control condition



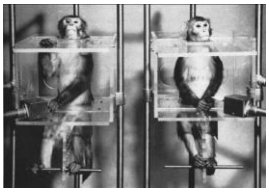


Participants' Level of Introversion by Condition (with Random Assignment)

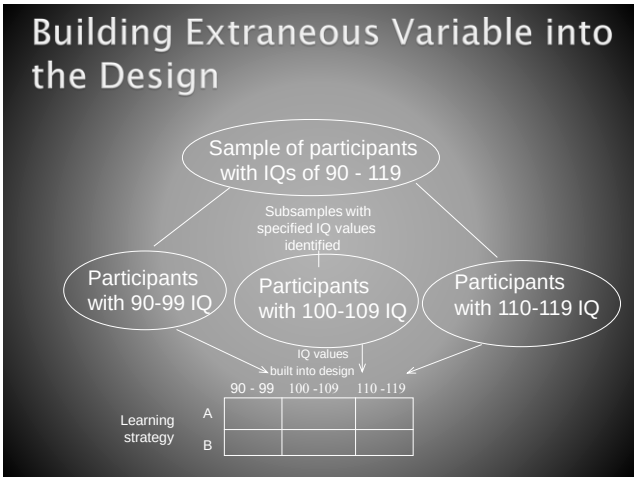
Participant	*Introversion Level (0–100)*	*Condition*	*Average Introversion Level by Condition*
A	36	Experimental	
B	45	Experimental	50
C	69	Experimental	
D	29	Control	
E	51	Control	50
F	70	Control	

Matching by Yoked Control

- Controls for possible influence of temporal relationship between event and response
- Match on events that happen within experiment
- Ulcers due to physical or psychological stress of shock in monkeys (Brady, 1958)



Building Extraneous Variable into the Design



Multiple Groups Design

- Experiments that consist of more than the traditional two groups, experimental and control
- Can involve quantitative IVs
  - Conditions referred to as levels
  - Represent greater or lesser amount of the IV
  - Amount of alcohol consumed
- Can also involve qualitative IVs
  - Conditions of IV represent different types or aspects of IV
  - Nature of video; graphic, comedic, cartoon

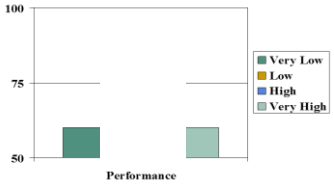
Multiple Groups Design

When additional levels of IV are examined, results can distinguish between two general categories of relationships between IV and DV

- Linear relationships: Scores on DV change constantly with level of IV
  - Positive relationship: Scores on DV increase as level of IV increase
  - Negative relationship: Scores on DV decrease as levels of IV increase

Multiple Groups Design

- Curvilinear relationships: Relationship between IV and DV takes a form other than a straight line
  - If only two levels of IV are examined, these relationships can be overlooked



Interpreting Results of Multiple Group Experiments

- Results of ANOVA do not tell you which of all possible two-group comparisons are statistically significant
  - Only know that at least one is
- To determine significant difference, you conduct a follow-up analysis
  - When you do not have a specific a priori hypothesis, use a post-hoc test
  - If you have an a priori hypothesis, use a priori contrasts

Advantages of Between Designs

- Each score is independent
- Not susceptible to:
  - Practice or experience gained in other treatments
  - Fatigue or boredom
  - Contrast or carry-over effects

Disadvantages

- Large # of participants
  - Esp. problematic with special populations
- Environmental Confounds
  - Characteristics of environment that might vary between groups
- Individual Differences
  - Can become confounding variables
    - Assignment bias
  - Can produce high variability in scores

Threats to Internal Validity

- Biased assignment of participants to conditions
- Differential attrition
- Diffusion of treatment
- Compensatory behavior
  - Rivalry (John Henry effect)
  - Equalization
- Resentful Demoralization

Within-Subjects Designs

- Each research participant experiences all of the conditions/levels of the IV (Repeated Measures)
- Results in perfect equivalence of participants across conditions (Dependent samples)
  - Reduces error variance

Advantages of Within Subjects Designs cont.

- Reduces number of participants needed
- When individual differences are consistent across treatments, can measure them and separate effects from the rest of the variance
  - Treatment effects easier to see when individual differences removed
- So within design more powerful than between design

Between vs. Within

Student	Speeded Test	Untimed Test
John	78	
Mary		74
Peter		68
Paul	80	
Average	79	71

Between vs. Within

Student	Speeded Test	Untimed Test
John	78	88
Mary	62	74
Peter	60	68
Paul	80	93
Average	70 (79)	81 (71)

Disadvantages of Within Subjects Designs

- Time demand
- Participant Attrition
  - Volunteer Bias
- Time Related Factors/Threats to IV
  - History
  - Maturation
  - Instrumentation
  - Regression
  - Pretest sensitization
  - Testing
    - Order Effects
    - Carry-Over

Solutions to Time-Related Threats

- Reducing time between treatments
  - But can increase risk of carry-over etc.
- Switch to between design
- Counterbalancing
  - Matching treatments with respect to time

Counterbalancing

- Treatments given in different orders
- Balances but hides order effects
- NOTE: does NOT make it a between design
  - Groups balanced on order but NOT on IV itself

Group 1	Treatment A	Treatment B
Group 2	Treatment B	Treatment A

Intrasubject (within subject) Counterbalancing

- The ABBA Technique
- Administer treatment conditions to each participant in more than one order
- Coke pepsi pepsi coke
- Based on assumption that order effects are linear
- If not linear – use each treatment condition in every possible position in sequence
  - Also use BAAB pepsi coke coke pepsi
  - Half participants assigned to each sequence

Intragroup Counterbalancing

- Less time-consuming
- Groups of participants rather than individuals counterbalanced
- Different groups take each of the sequences

Complete Counterbalancing

- All possible treatment sequences are presented.
- You can calculate the number of sequences by using the formula *n!* (*n* factorial).
- With n = 6, *n!* = 720!
  - 6 X 5 X 4 X 3 X 2 X 1
- Might require too many participants

Incomplete/Partial Counterbalancing

- Only a portion of all possible sequences are presented
  - Three basic requirements:
    - Each treatment must be presented to each participant an equal number of times.
    - Each treatment must occur an equal number of times at each testing or practice session.
    - Each treatment must precede and follow each of the other treatments an equal number of times.

Counterbalancing

A	B
B	A

With Two Conditions

A	B	C
A	C	B
B	A	C
B	C	A
C	A	B
C	B	A

With Three Conditions

Latin Square Designs

The number of orders is equal to the number of conditions
Each condition appears in each place in the order

Basic

1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

Balanced

1	A	B	C	D
2	B	D	A	C
3	C	A	D	B
4	D	C	B	A

Counterbalancing

- Sequence or Order Effects
  - Sequence or order effects are produced by the participants being exposed to the sequential presentation of the treatments.
    - The sequence or order effect depends on *where* in the sequential presentation of treatments the participant's performance is evaluated, not *which* treatment is experienced.

Counterbalancing

TABLE 5-7 EXAMPLE OF SEQUENCE OR ORDER EFFECTS IN A COUNTERBALANCED EXPERIMENT

The performance increase shown in parentheses below each sequence indicates the effect of testing reaction time to red (R), green (G), and flashing white (FW) lights on an instrument panel at that particular point in the sequence. Thus, second and third testings result in increases of 4 and 3, respectively, regardless of the experimental task.

	ORDER OF TASK PRESENTATION		
	R	G	FW
Performance Increase →	(0	4	3)
	R	FW	G
	(0	4	3)
	G	R	FW
	(0	4	3)
	G	FW	R
	(0	4	3)
	FW	R	G
	(0	4	3)
	FW	G	R
	(0	4	3)

Counterbalancing

- Carryover Effects
  - The effects of one treatment persist or carry over and influence responses to the next treatment.

TABLE 5-8 EXAMPLE OF CARRYOVER EFFECTS IN A COUNTERBALANCED EXPERIMENT

Carryover effects occur when a specific preceding treatment influences the performance in a subsequent treatment. In this example, experiencing Treatment G prior to Treatment R results in a decrease of 2 (i.e., −2), whereas experiencing Treatment R prior to Treatment G results in an increase of 2 (i.e., +2). Experiencing Treatment G prior to FW or Treatment FW prior to G does not produce a unique effect. However, experiencing Treatment R prior to FW results in an increase of 3, whereas experiencing Treatment FW prior to R results in a decrease of 3.

	SEQUENCE OF TREATMENTS		
Effect on Performance →	G	R	FW
	(0	−2	+3)
	G	FW	R
	(0	0	−3)
	R	G	FW
	(0	+2	0)
	R	FW	G
	(0	+3	0)
	FW	G	R
	(0	0	−2)
	FW	R	G
	(0	−3	+2)

Categories of Order Effects

- Sensitization effects: A form of reactivity in that experiencing one condition affects performance in another condition
  - Example: Participants evaluate the résumés of both a physically attractive and physically unattractive job applicant
  - Results in demand characteristics if participants:
    - form hypothesis about purpose of experiment (Does physical attractiveness affect evaluations?)
    - respond in socially desirable way
  - Harder to control than practice or carryover effects

No Order Effects

- No difference if treatment is presented first or second
- Where for Group A, Treatment 1 occurred 1st and for Group B Treatment 2 occurred 1st
- Difference of 5 points between Treatments regardless of when presented

Group	Treatment A	Treatment B
1 (A, B)	20	15
2 (B, A)	20	15

Group	Order 1	Order 2
1 (A, B)	20 Treatment 1	15 Treatment 2
2 (B, A)	15 Treatment 2	20 Treatment 1

Symmetrical Order Effects

- Order matters and is the same regardless of what the treatment is
- E.g., second treatment score always raised by 10 points regardless of which treatment it is

Group	Treatment A	Treatment B
Group 1 (A, B)	20	30
Group 2 (B, A)	34	24

Group	Order 1	Order 2
Group 1 (A, B)	20 Treatment 1	30 Treatment 2
Group 2 (B, A)	24 Treatment 2	34 Treatment 1

Nonsymmetrical Order Effects

- Specific treatments determine the type of order effects, e.g., fatigue vs. practice (Differential order effects)
- Group 1 does better on Treatment B when receiving it second, but Group 2 does the same on both treatments when receiving treatment B first

Group	Treatment 1	Treatment 2
Group 1 (A, B)	20	30
Group 2 (B, A)	24	24

Group	Order 1	Order 2
Group 1 (A, B)	20 Treatment 1	30 Treatment 2
Group 2 (B, A)	24 Treatment 2	24 Treatment 1

Example of Differential Order Effects

	Condition		
	Experimental	*Control*	*Difference*
True performance level	**40**	**30**	**10**
Order Effect for E to C	**0**	**10**	
Order Effect for C to E	**20**	**0**	
Observed Score	**60**	**40**	**20**