Experiment Design

Main Parameters

Experimental design is not a linear planning process

Independent Variable

What will you manipulate either between subjects or within subjects?

Task

What will subjects be asked to do? What will the stimuli be?

Dependent Variable

What will you measure on the subject?

Sample

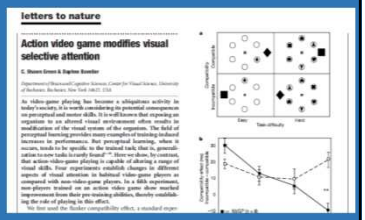
Who are the subjects, how will you recruit them, and how assigned to groups?

Mini summary of Green & Bavelier

What was the research question?

What did they find?

Why is that a big deal?


letters to nature

Action video game modifies visual selective attention

Hypothesis: Videogame playing enhances attention

Research Types

Controlled Experiment

*an **experiment** that isolates the effect of one variable on a system by holding constant all variables but the one under observation – via random assignment of subjects to experimental and control conditions or balancing of experimental and control conditions in within-subjects designs.*

Natural Experiment

*an empirical study in which individuals (or clusters of individuals) exposed to the **experimental** and control conditions are determined by nature or by other factors outside the control of the investigators.*

Case Study

a process of research in which detailed consideration is given to the development of a particular person, group, or situation over a period of time.

SOURCES: Encyclopedia of Psychology, Wikipedia

What type of research was the Green & Bavelier paper?

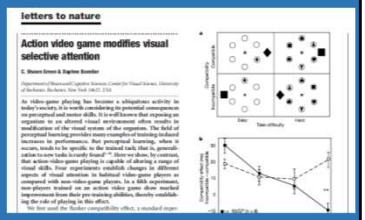
Controlled Experiment

People trained to play action videogames vs control (people trained non-action videogames)

Natural Experiment

People who play videogames versus people who don't

Case Study


letters to nature

Action video game modifies visual selective attention

What aspects of people's psychology/experience did Green & Bavelier manipulate or vary?

Videogame Experience

Attention

Independent Variables

the traits or factors that you manipulate or vary because changes in their value are expected to predict the value of the dependent variable (ie., the subjects' performance or behavior)

How to vary the independent variables:

Between-subject Designs (comparing different groups)

Each participant contributes to one condition
Baseline Group vs Treatment Group *("treatment" can be a drug, stimulus condition, environment, etc.)*

Advantage: Easy
Disadvantage: Group differences in personal characteristics (eg. IQ, age, culture, ability)

Within-subject Designs (comparing different conditions)

Each participant contributes to all conditions
Baseline Condition vs Treatment Condition

Advantage: Eliminates between-group subject differences
Disadvantage: Order Effects & Carryover Effects

Mixed Design (a little of each)

What type of experiment was the Green & Bavelier paper?

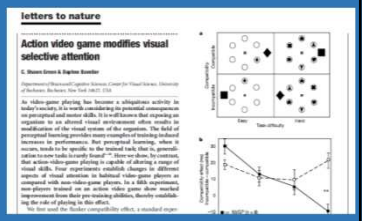
Between-subjects

Within-subjects

Mixed

Some comparisons were made between-subjects (eg., videogamers versus non videogamers; action game trained vs non-action game trained)

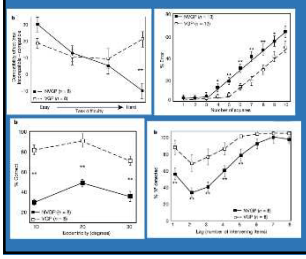
Some comparisons were within-subjects (pre- vs post-training)


letters to nature

Action video game modifies visual selective attention

Between-subject Designs (comparing different groups)

Non-videogame Players (NVGP) versus Videogame Players (VGP)
Comparing different groups of people who do and don't play videogames



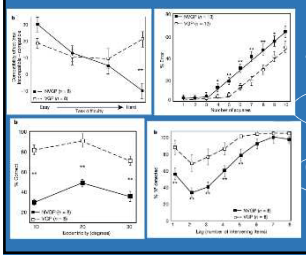
Which is the control group?

Equal N in the two groups? Does it matter?

How were subjects assigned to the control group?

Between-subject Designs: Define Control Group

Control Group (for between-subjects group characteristics)
Forms a baseline for comparison with experimental group



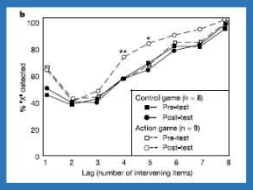
Random Assignment (subjects randomly assigned to groups)

Matched Groups (groups assigned to match average age, IQ, gender, etc)

Quasi-experiment (groups assigned by nature: age, gender)

Within-subject Designs (comparing different conditions)

No Videogame Experience (Pre) versus Videogame Experience (Post)

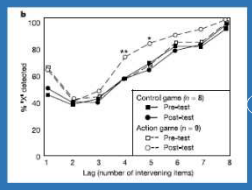


Within-subject is Pre vs Post = Squares vs Circles

Between-subject is Control Game vs Action Game = Solid vs Dashed lines

Within-subject Designs (comparing different conditions)

No Videogame Experience (Pre) versus Videogame Experience (Post)



Random Assignment (subjects randomly assigned to groups)

Matched Groups (groups assigned to match average age, IQ, gender, etc)

Quasi-experiment (groups assigned by nature: age, gender)

What aspects of people's psychology/experience did Green & Bavelier manipulate or vary?

Videogame Experience

Attention

Independent Variables

the traits or factors that you manipulate or vary because changes in their value are expected to predict the value of the dependent variable (ie., the subjects' performance or behavior)

Task-based Manipulations of Attention

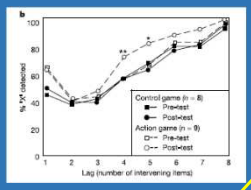
Between-subjects (Exps 1-4):
- **Flanker Task**
- **Enumeration Task**
- **Useful Field of View Task**
- **Attentional Blink Task**

Within-subjects (Exp 5):
- **Enumeration Task**
- **Useful Field of View Task**
- **Attentional Blink Task**

Define the Independent Variable

Factors & Levels

What are the specific **factors** that define the independent variable?
How many **levels** does each factor have?



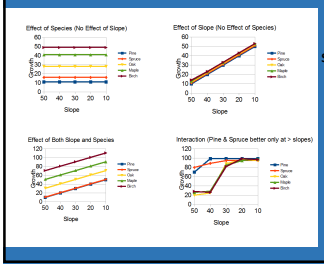
Factors
- **Videogame Exposure**
  - 2 levels: Pretest vs Posttest
- **Videogame Type**
  - 2 levels: Action vs Control
- **Number of Intervening Items to Target**
  - 8 levels: 1-8 items

Between-subject Factor — **Within-subject Factors**

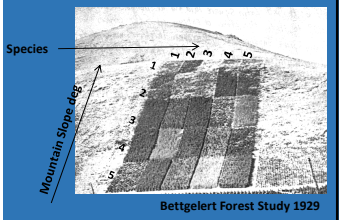
Mixed Design (a little of each)

Factorial Design

means it's a balanced design because all levels of one factor are represented at each level of another factor



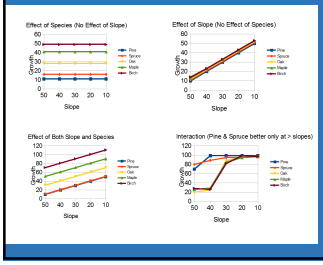
5 x 5 Factorial



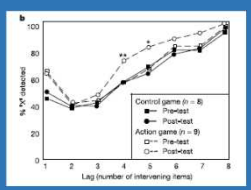
Bettgelert Forest Study 1929

Factorial Design

means it's a balanced design because all levels of one factor are represented at each level of another factor



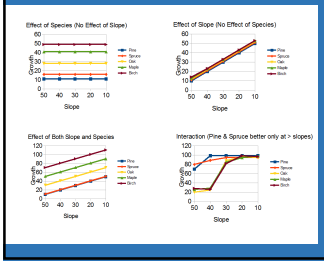
2 x 2 x 8 Factorial



Interaction: Post > Pre only for action videogame and only at intermediate num of intervening items

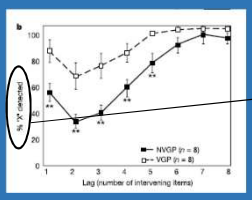
Possible & Predicted Results

A good thing to do during design is to graph out the possible patterns of results that you might get from your experimental conditions



Define the Dependent Variables

A **dependent variable** is what you measure in the experiment and what is affected during the experiment. It is called **dependent** because it "depends" on the **independent variable**.



Minimize the Error and Noise

minimize performance & measurement error
many measurements (trials)
many subjects

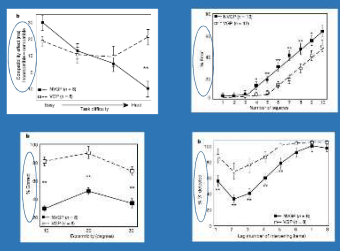
Dependent Variable
% Xs Detected (a type of Accuracy)

Types: Accuracy and Response Time are common

What were some of the other dependent variables in the paper?

SOURCE: https://www.ncsu.edu/labwrite/po/dependentvar.htm

Define the Dependent Variables



What unique information does each measure provide?

For example, if you have accuracy and RT will they provide redundant information? Or will RT be more useful than accuracy or vice versa?

Task Details

Instructions
What will subjects be told about the task (the task instructions)?
Will the instructions be written or spoken (script)?

Stimuli
What will subjects see/hear/taste etc. in each condition (the stimuli)?
How many stimuli per condition? How many trials per stimulus?
How many secs/mins will subjects be exposed to each stimulus (stim duration)?
In what order will the stimulus conditions be presented (counterbalanced)?

Response
How will subjects make a response (mouse, keyboard, other peripheral)?
How many secs/mins will subjects be allotted per response (response window)?

Counterbalancing Design

Example: Measuring ability to detect light flash with 3 levels of flash duration: Short, Medium, Long
We need to **counterbalance** the order that subjects experience these levels

Possible Orders (3! = 6):
1) Short, Medium, Long
2) Short, Long, Medium
3) Medium, Short, Long
4) Medium, Long, Short
5) Long, Short, Medium
6) Long, Medium, Short

Fully Counterbalanced test all possible orders making sure each condition follows every other condition and each condition precedes every other condition once.

Can test different subgroups of subjects on each order type (total N will be multiple of 6).

Can test each order type within each subject's session (total trials will be multiple of 6).

Full counterbalancing isn't always practical, depending on variable type & levels

Counterbalancing Design

4 levels: Short, Medium, Long, and Longest

4! = 24 orders
5! = 120 orders
7! = 5040 orders

Latin Square select a subset of orders in which each condition (short, medium, etc.) is tested once in each position (1st, 2nd, 3rd, and 4th).

Balanced Latin Square select a subset of orders such that each condition is tested once in each position AND each condition both precedes and follows each other condition once.

Make an n x n table. The first row of the Latin Square will follow the formula 1, 2, n, 3, n-1, 4, n-2..., where n is the number of conditions. For subsequent rows, you add one to the previous values, circling back to 1 after n.

	1st	2nd	3rd	4th
Order 1	Short	Medium	Long	Longest
Order 2	Medium	Long	Longest	Short
Order 3	Long	Longest	Short	Medium
Order 4	Longest	Short	Medium	Long

	1st	2nd	3rd	4th
Order 1	Short	Medium	Longest	Long
Order 2	Medium	Long	Short	Longest
Order 3	Long	Longest	Medium	Short
Order 4	Longest	Short	Long	Medium

http://rintintin.colorado.edu/~chathach/balancedlatinsquares.html

Define the Sample

http://www.uv.es/uvetica/files/McCrum_Gardner2010.pdf

You'll already know some characteristics of your sample from the research question and independent variables: typical adults, children, people with certain experiences, etc.

How many subjects?
Replication Study: Same as original
Original Study: Power Analysis
Power is the probability of detecting a true significant difference
Effect size is the numerical difference between groups/conditions in uits of standard deviation
Alpha level is the cut-off for determining statistical significance
How recruited?
Is sample representative (age, gender, race, IQ, etc.)?
Group assignment

Warning about sample size

http://www.uv.es/uvetica/files/McCrum_Gardner2010.pdf

Larger is not automatically better

Do people with curly hair weigh more than people with straight hair?

Curly Hair	Straight Hair
150.009 pounds	150.001 pounds

In a t-test:
N = 10, **NO**, p = .9
N = 1,000,000, **YES**, p < .001

Is a difference of .008 pounds meaningful?

Always have to **think**:

What number of subjects do I need to answer my question? Is 1 enough, 10, 100?

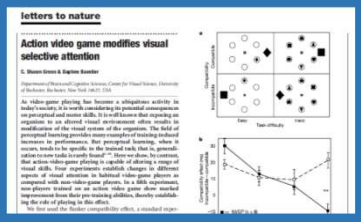
What amount of difference between groups or conditions would be meaningful? How many subjects would be needed for that difference to be detected? **POWER ANALYSIS**

What was the sample size in Bavelier & Green?

Between-subjects

Within-subjects

Problematic? Or not?


letters to nature

Action video game modifies visual selective attention

Pilot testing

Test procedures and determine optimal values of the experimental variables before moving on to the main experiment

Helps identify weaknesses in the experiment

Typical Pilot Test is 5 Subjects per group

Debriefing Subjects

Ethics
Inform subjects about the hypothesis and nature of the study.

Allow participants to express discomfort.

Methods
Evaluate whether subjects guessed the hypothesis and whether any of their ideas influenced their behavior in the experiment.

Check effectiveness of manipulation and subjects' impressions of the task.

Subjects can be excluded on the basis of idiosyncratic strategies that they adopted during the experiment.

Control Condition Design

Very important…will talk about it in depth in class on MEASUREMENT which covers the *subtraction method*

Design of Shepard & Metzler

What type of research (natural, controlled, etc)?
Between- or within-subject design?
How many subjects?
How were subjects assigned to study conditions?
What were independent variables? How many levels?
What were the dependent variables?
Was it a factorial design?
Were conditions counterbalanced?
What were control conditions?