Lecture 4: Quantitative Research

The what, the why & the how

Dr Danielle Evans (& Dr Milan Valášek)
21 October 2021

UNIVERSITY OF SUSSEX

Overview of Quantitative Research

Part 1: The What

- Quantitative Vs qualitative methods

Part 2: The Why

- Generalisable knowledge, theories, & hypotheses

Part 3: The How

- Quantitative research design
  - Observational/correlational
  - Experimental, quasi-experimental & natural experiments
  - Longitudinal & cross-sectional
- Measurement
  - Levels of measurement
  - Variables
  - Validity
  - Reliability

Quant Research: The What

- Two branches of investigation in psychology: **qualitative** & **quantitative**
- **Qualitative** is concerned with **meaning** & understanding subjective **experiences**
  - We will revisit qual in Year 2
- **Quantitative** aims to **quantify** & **measure**
  - We use numbers to represent different **constructs** & we can fit **mathematical** & **statistical models** to those data
- The decision to use either method should be based on your **research question**
- Quantitative research is the focus of today, & much of your research methods journey @ Sussex

Quant Research: The Why

- We use quantitative research to obtain **generalisable knowledge**
  - We can **generalise** our findings to other participants/contexts
- We can form a **scientific theory** which is a framework supported by observations
  - A theory should describe, explain, or predict phenomena
- We can develop **hypotheses** which are theory-driven predictions
- & we can then use **statistical models** to test those predictions

Are dogs the key to happiness?

Are dogs the key to happiness?

Not necessarily... 

Quant Research: The How

- The conclusions we can draw from our research depend on how that knowledge was generated & the research design used
    - But how do we actually test **hypotheses** appropriately?
    - How can we **generalise** our findings?
    - How do we **quantify** unquantifiable things?
    - & how can we be confident in our conclusions?

The answer lies in quantitative research design!

Quant Research: The How

- We can employ different types of research design to answer different **research questions** & to test different **hypotheses**
- Study designs can vary on the presence or absence of a **manipulation**, whether they use **between-** or **within-subjects** measurement, & the time frame used
- Some examples include:
  - **Observational**/**correlational** designs
  - **Experimental** designs (incl. **quasi-experiments** & **natural experiments**)
  - **Longitudinal** & **cross-sectional** designs
- Which design to choose depends on:
  - The hypothesis to be tested
  - Resources (time, money, facilities)
  - Logistical considerations
  - Ethical considerations

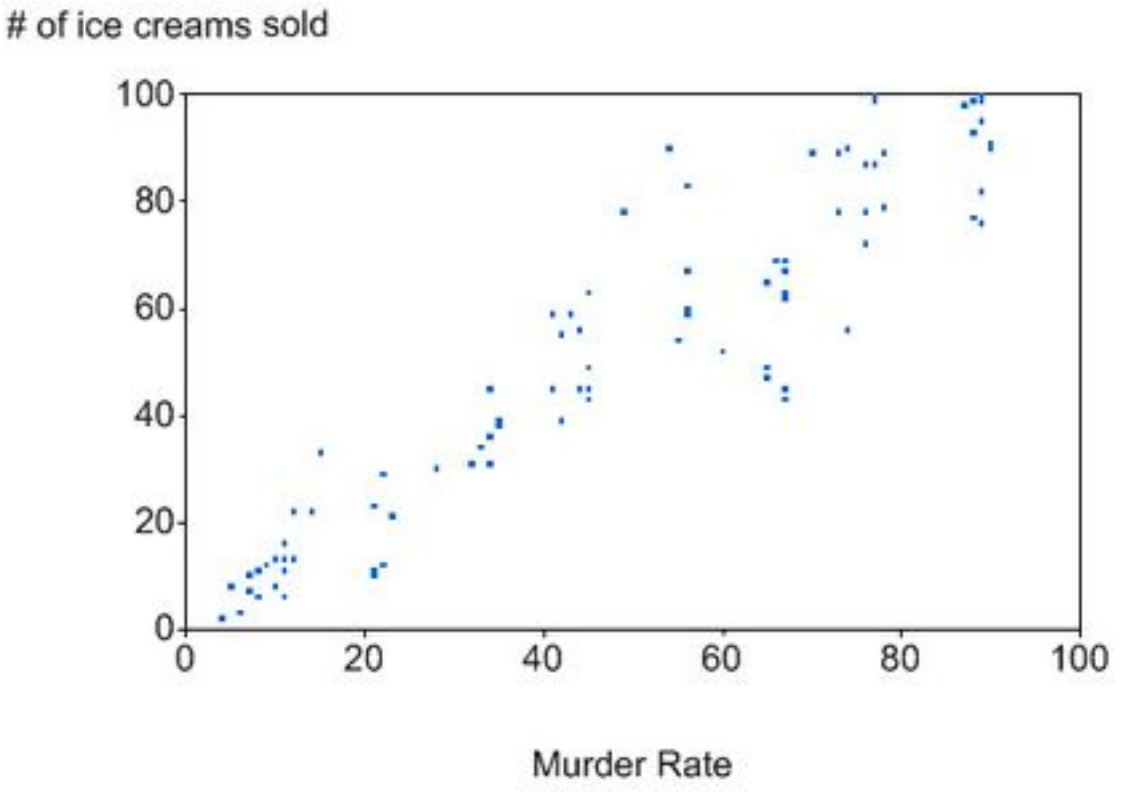
Manipulations

- The key feature of **experimental designs** is that there is some form of **manipulation** which changes an aspect of the study design to see whether it affects the study's outcome
- Designs that don't include manipulations are known as **observational** or **correlational** designs

Observational/Correlational Designs

- In correlational designs, we rely on **observed data** which are not subjected to any **experimental manipulation**
- Observation doesn't just mean *looking*, there is still measurement, questionnaires, laboratory tasks, etc.
- Correlational designs allow us to collect data from lots more people compared to experimental designs
- We can use these types of design to investigate **relationships** between **variables** but they cannot tell us anything about **cause & effect**
- Some experiments would be **unethical**, but observational data can be a solution
- But because we cannot infer **causality**, we have to be careful with our conclusions..

Ice Cream Murders in NY



But correlation does not equal causation!!

Conducting an Experiment

- Ignoring the *many* ethical issues for a sec, let's imagine this experiment...
- First, we would generate our **research question**: does eating ice cream make you more prone to murderous tendencies?
- Then we would come up with a **hypothesis**: eating ice cream increases the desire to commit murder
- To test our hypothesis, we could design an experiment where we invite a bunch of people into the lab:
    - Half of our participants eat some ice cream, the other half don't (our **manipulation**)
    - Then we get all participants to look at several images of people (the **stimuli**) & rate how much they want to *eliminate* them on a scale from 0 (no desire) to 9 (all the desire possible)
- After the experiment has finished, we would thank our participant, we would **debrief** them by describing the aims of the study in more depth, & send them on their way..

Then as they walk out we hear them say **'I could *kill* for an ice cream'** 😢

Conducting an Experiment

- In this study, we have one **independent variable (IV)**, or **predictor variable** which is the group people were assigned to (ice cream or no ice cream) & we have one **dependent variable (DV)** or **outcome variable** which is their total 'desire to eliminate people' score
- Once we have all our data, we could then compare the results between the ice cream **condition** and the no ice cream **condition** to see which group gave higher ratings for their 'desire to eliminate people' score (the **outcome**/**DV**)
- Overall, this is an example of a very unethical and a poorly designed study
- With a well designed experiment, we can be confident in saying that our **manipulation** caused the change in our **outcome variable**/**dependent variable** (**DV**)
- But this isn't the case for our ice cream study for many reasons...

Controls

- Our imaginary study didn't use *any* **controls**
- We recruited all kinds of people with no thought of how different characteristics might affect our results:
    - We could have children as well as adults in our **sample**
    - We didn't screen for any health conditions that may affect our findings (allergies, intolerance, dietary requirements)
- We also didn't give any **standardised instructions**: some participants could have arrived very hungry, and others very full
- We didn't control our **IV** appropriately: we often changed the brand, the flavour of ice cream, & the amount given - one day we ran out so had to use frozen yoghurt...
    - Now we don't know exactly what caused any changes in our **outcome**
    - It could be that only strawberry mini milks cause murderous tendencies
    - The poor control of our **IV** creates **noise** in our data
- We didn't control the lab environment it was conducted in, on some days the heating was up super high & on others we had the windows wide open

Randomisation

- We didn't **randomly assign** people to groups: maybe all our participants in the ice cream condition just so happened to have sensitive teeth, or already high levels of psychopathy
    - This would affect the results and may give us the impression that ice cream increases murderous tendencies when it doesn't
- Differences between participants can introduce unwanted **variability** into the experiment
- To deal with this, participants should be sorted into the groups **randomly**
- A well-designed experiment randomises both **participant allocation** and **stimulus presentation order** (in our experiment, this refers to the images of faces our participants rated)

Blinding

- In our study, we told them that we were interested in the effects of ice cream on murderous tendencies, & we also gave them the ice cream ourselves
- Participants may (consciously or not) modify their behaviour in accordance with or contrary to the hypothesis
- Because of that, it is crucial that they be kept unaware of what the hypothesis is & also what conditions they have been allocated to
- If participants are naïve to group allocation, the study is said to be **single-blind**
- If neither the participants nor the researcher know which condition the participants are put in, the study design is known as **double-blind**
    - Allocation is recorded but only revealed once the study is over and the data are being analysed

Theoretical Framework

- It's the researcher's job to decide which variable is the **predictor** (**IV**) and which is the **outcome** (**DV**) based on **theory**
- In this experiment, the decision to have ice cream as the **IV** & murderous tendencies as the **DV** was not based on any **theoretical model**
    - It could be that committing murder causes people to eat ice cream, they could be completely unrelated & the relationship between them is just a coincidence, or maybe something else is linked to them both...

Experimental Design

- The experimental setting with its tight controls is always at least somewhat **artificial**
    - It presents an abstraction from reality
- Just because something is true in the lab, doesn't necessarily mean it will be true in "the real world"
- The experimental design provides the most rigorous methodology to investigate **causal relationships**
- However, there are often methodological, logistical, and ethical obstacles to randomisation, manipulation, and controls that render designing an experiment infeasible
    - This is when we might use **quasi-experiments** & **natural experiments**

Quasi-Experiment

- Conforms to all the requirements of the experimental design **except for participant randomisation**
- Used in situations where it is not possible to randomise the allocation of participants into group
    - E.g., the effectiveness of attending summer school - one school offers the intervention, the other does not
- But conditions should be **matched**
    - Participants are matched so that the groups don't differ on any relevant characteristics, other that the one being investigated

Natural Experiment

- Manipulation and randomisation occur through some **natural** or **socio-political** means
- A good example are **twin studies**:
  - Identical twins share essentially 100% of their genes
  - Fraternal twins share on average 50% of their genes
  - Both kinds of twins tend to share the same home environment (raised together)
  - Comparing similarities between identical twins and similarities between fraternal twins, we can estimate the role of genes and environment in all sorts of things (physical/mental health, personality, cognitive ability, etc.)
- Other kinds of natural experiments may be made possible due to policy changes (smoking ban, length of compulsory education...) or natural events
- Neither **natural experiments**, nor **quasi-experiments** are true experiments - they can't provide evidence of causality of the same strength as experiments

Within- & Between-Subjects

- Studies can also vary on whether the manipulation or measurement of interest occurs **between groups** of participants or **within** each participant's data
- **Between-subjects** also known as **independent designs** are where we compare different groups of participants
    - This is where different participants are assigned to (or naturally fall into) different conditions
- **Within-subjects** also known as **repeated measures** designs are where we take the same measurement from the an individual multiple times
    - This is where the same participants are assigned to (or naturally fall into) *all* conditions of the study
    - Cons include **order effects** such as fatigue, practice etc.,
- **Mixed designs** have both within- and between-subjects components

Time Frame

- The type of design we have can also vary on the time frame used
- **Cross-sectional designs**
  - Take a cross-section of a sample at a single time point
  - Cross-sectional studies are logistically easier to run compared to other types of studies
  - Not very useful for telling us how things change over time because different groups of participants are compared at the same time
- **Longitudinal designs**
  - Involves repeated measurement of the same characteristic of the same participants at multiple different time points
  - Logistically very difficult to do, some longitudinal studies have been going for several decades, & are very expensive to run
  - They are very time-consuming for the researcher & the participant, & so the amount of missing data can be very high

Missing Data

- Missing data can occur in all types of study - not just longitudinal studies though the large-scale of them makes it more likely
- Missing values can result from different causes:
    - Tech issues meaning that an answer doesn't get registered or recorded
    - Participants skip a question or a task by accident
    - Participants don't wish to disclose a piece of information (more common with sensitive topics)
- Missing data are often assigned the value **NA**, for "**N**ot **A**pplicable"
- The presence and pattern of missing data can be random, but sometimes it isn't
    - Maybe the participants skipping your over income question are from a low SES background
- So by exploring the pattern and possible causes of missing responses we can learn more about our data

Measurement

- Measurement refers to the way we **record behaviours, feelings, attitudes, & characteristics in a quantifiable sense**
- We typically refer to these different measures as '**variables**'
    - **A variable is anything that can be measured which differs across participants or across time**
- Even things that we have not measured or manipulated can be variables:
    - **Confounding** & **extraneous variables** are variables that might affect the results of our study
- We can measure variables like height & weight, but we can also measure 'unquantifiable' constructs like 'happiness' or 'lecture-induced boredom'

Measurement Issues

- With measurement there are some issues to be aware of:
  - In quantitative research we can use numerical values to represent pretty much any construct, but what we can do with our data depends on how we measured these 'things'
  - We must carefully consider any variables that might affect the interpretation of our findings (i.e., **confounding** & **extraneous variables**), and ideally, we should design our study to eliminate or reduce their impact
  - We also have to consider how appropriate or *good* our measures are

Measures as Tools

- All measurement involves the use of a tool or a **measure**
  - Tape measure
  - Stopwatch
  - Questionnaire
  - Exam
  - Experimental task
- These measures/tools must work in the way we expect them to (**validity**) and they must work the same each time (**reliability**)

Construct Validity

- A measure is **valid** if it measures what it is supposed to measure
- In psychology, we often want to measure things that may be difficult to observe directly or **quantify**
    - Things like happiness, cognitive ability, personality...
- We attempt to measure these unobservable things by using a range of different tools including questionnaires, experimental tasks, tests etc.
- We design these tools by using the theoretical underpinnings behind the constructs we are trying to measure
- **Construct validity** is the extent to which a tool can be justifiably trusted to actually measure the construct it is supposed to measure

External Validity

- A study has **external validity** if its findings can be applied to the entire population of people with relevant characteristics & in real-world conditions
    - If study uses a sample of white women in western cultures, the findings may only be true for white women in western cultures
    - They may not apply or **generalise** to *all* people
- The amount to which findings apply in the real world outside of the study is **ecological validity**, this is often an issue for experiments
    - Just because something is true in the lab doesn't mean it is going to be true in the real world

Reliability

- **Reliability** refers to the **consistency** of a measure
- A measure is considered to be reliable if it produces consistent results each time it's used by the same participant
- For example, we could measure someone's maths anxiety with a questionnaire
- Our questionnaire could be considered reliable if we tested the same participant on different occasions and they got similar scores each time
- This stability over time is known as **test-retest reliability**
    - Doesn't work for measures that we expect to change over time

Levels of Measurement

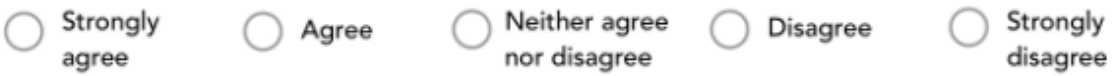
- The level of measurement refers to the kind of information we are working with when measuring attributes of interest
- There are four levels:
  - **Nominal/categorical**
  - **Ordinal**
  - **Interval**
  - **Ratio**
- Sometimes a construct can fall into many of these levels, and it's on the researcher to decide what measurement level is the most appropriate to use

Nominal/categorical

- Refers to **names**, **categories**, **labels** or **group membership**
- Examples include:
  - Eye colour (e.g., green, brown, blue)
  - Occupation status (e.g., FT employed, PT employed, unemployed, student...)
  - Study condition (control, experimental)
  - Marital status
  - Even age can be nominal if we wanted it to be (under 50s vs over 50s)
- When using nominal variables, we cannot compare the different groups in any quantifiable way
  - It makes no sense to say that green is more than blue when it comes to eye colour

Ordinal

- At the ordinal level, individual observations of the measured attribute can be **ordered** in a meaningful way
    - This could be marathon runners ranked in order of who came 1st, 2nd, 3rd...
- However, it doesn't provide any information about the differences between individual points
    - We don't know how much faster the winner was compared to the runner-up
    - The distance between the 1st and 2nd doesn't have to be the same as the distance between the 2nd and 3rd, or between any adjacent pair of runners
- An example of this type of measure in psychology is the **Likert scale**:



- These scales are classed as **ordinal** because we can't say that the difference in agreement between "Agree" and "Neither agree nor disagree" is the same as the difference between "Disagree" and "Strongly disagree"

Interval

- At the **interval** level of measurement, the differences (intervals) between pairs of adjacent values are the same
    - The difference in temperature between 20 and 21 degrees Celsius is the same as that between 35°C and 36°C
    - The intervals marked by the degrees are evenly spaced or equidistant
- But there is not an absolute zero point so we cannot say that 40°C is twice as hot as 20°C
- In psychology, a good example is IQ:
    - Someone with an IQ of 200 is not twice as smart as someone with an IQ of 100
    - There's also no such thing as an IQ of zero

Ratio

- Similar to interval, but with **ratio** data there is a **meaningful 0 point**
- Examples of the ratio level of measurement in psychology:
  - Reaction time
  - Number of correct responses
  - Score on an exam
  - Income

Variable/Data Types

- When using quantitative methods, we represent variables with a numerical value, but we can have different types depending on the type of data it represents
- **Continuous variables** can contain any numerical value within a certain range
    - e.g., time, height, weight
- **Discrete variables** can only contain some values
    - e.g., number of children: hypothetically you can have any number of children, but it has to be a **whole number**
    - There's no such thing as a 2.4 children
- **Binary variables** can only take one of two possible values
    - A special case of **discrete variables**
    - Used for variables with two responses/categories: yes & no, heads Vs tails, pass/fail...

That's all for today!