

- → Pages 31 – 42
- → Does listening to classical music make you smarter?
- Governor Zell Miller of Georgia placed \$105 000 in state budget towards purchasing classical music
- Believed that providing young children with classic music would make them smarter
- May be true: observations that most ppl we know that listen to classical music are seen as intelligent/sophisticated
- Around the same time that Georgia took this step, consumers being bombarded with advertisement about “the Mozart Effect”
- Classical sections at music stores moved to front of the store – drawing customers’ attention to the intelligence boosting effects of the CDs
- Parents considered that it was never too early for children to listen to classical music/ even when in the womb!
- Some U.S. states dedicated to handing out classical CDs along with hospital birth certificates
- Enthusiasm for the Mozart effect died down – other scientists unable to replicate results
- -Turns out hype surrounding the Mozart effect was based on the results of *one* study
- Study:
- -Twelve participants listened to Mozart/performed better than other adults on a test of spatial ability
- -These temporary differences in spatial intelligence were then inflated by the popular press to mean intelligence in *general* (there is a big difference!)
- Highlights that there should be greater scientific literacy in our society
- -Researchers did not make any unethical claims, media trying to present an interesting story to their audience
- -But because of a lack of scientific literacy/critical thinking → these events have now become a cautionary tale
- Shows that asking tough questions about how research was designed/conducted is never a bad thing
- Most important aspect of scientific research is that it strives for objectivity
- Objectivity: Assumes that certain facts about the world can be observed and tested independently from the individuals who describe them (e.g., the scientist)
- -Everyone → not just the experts → should be able to agree on these facts given the same tools, the same methods, and same context
- Soon as ppl observe an event, their interpretation of it becomes subjective

- Subjective: Meaning that their knowledge of the event is shaped by prior beliefs, expectations, experiences, and even their mood
- A scientific objective approach to answering questions differs from a subjective one *
- **→ FIVE CHARACTERISTICS OF QUALITY SCIENTIFIC RESEARCH:**
- Quality scientific research meets the following criteria:
 1. It is based on measurements that are *objective, valid, and reliable*
 2. It can be *generalized*
 3. It uses techniques that reduce *bias*
 4. It is made *public*
 5. It can be *replicated*

→ Scientific Measurement: Objectivity, Reliability, and Validity:

- Foundation of scientific methodology is the use of objective measurements
- Objective Measurements: The measure of an entity or behaviour that, within an allowed margin of error, is consistent across instruments and observers
- The way that a quality or behaviour is measured must be the same regardless of who is doing the measuring and the exact tool they are using
- → Example: weight is measured in lbs/kg
- -1kg in St. John's is the same as 1kg in Victoria – researchers don't get to choose how much mass a kilogram is worth
- -Your weight will be the same regardless whether you're using the scale in your bathroom or a scale in a change room
- However, your weight *will* vary slightly from scale to scale – this is the margin of error mentioned in the definition of objective measurements
- Scientists have to agree upon how much variability is allowable
- -Most ppl are comfortable if their weight differs 1/2kg depending on the scale being used
- -But, if you weigh 70kg on one scale and 95kg on another → then you know one of your measurement tools is inaccurate
- Weight would be considered a variable in this example
- Variable: The object, concept, or event being measured
- Variables can be utilized in perceptual processes, learning/memory, to how we interact with each other, etc. (Each variable can be described and measured)
- Most of psychology's history, measurements involved observations of behaviour in different situations/examinations of how participants responded on a questionnaire/to stimuli presented on a computer

- As technology advanced though, so did the ability to ask psychological questions in new/interesting ways
- High-tech equipment → such as functional magnetic resonance imaging (fMRI) – allows researchers to view the brain/see which areas are activated while you perform different tasks such as remembering words/viewing emotional pictures
- Other physiological measures: involve gathering samples of blood/saliva, which can be analyzed for enzymes, hormones, and other biological variables that relate to behaviour/mental functioning
- Due to greater number of measurement options, now possible to examine the same variable (e.g., anxiety) using many different techniques
- -Allows us to understand the different elements of behaviour
- Any method used by a researcher to measure a variable needs to include carefully defined terms*
- -How would you define personality, shyness, or cognitive ability?
- Researcher must answer this type of question carefully not for only planning/conducting a study, but also when sharing the results of that research
- Operational Definitions: Are statements that describe the procedures (or operations) and specific measures that are used to record observations
- - Example: Depression could be operationally defined as “a score of 21 or higher on the Beck Depression Inventory” (A measure that is common/widely accepted clinical questionnaire)
- Operational definitions would have been useful for the Mozart effect
- -“How do researchers define the outcome of their study? Do they mean listening to classical music makes you *smarter*, or just that you remember better? Do they claim the effect is permanent, or does it occur only while listening to Mozart?”
- Mozart Effect in Georgia (spent over 100 000 on CDs)
- -Researchers used many objective measures of thinking/reasoning in studies of listening to classical music, including objective behaviour measures
- Based on these measures →the only improvement is to be in one type of thinking called spatial reasoning → The ability to look at objects and mentally manipulate them
- Improvement is also short-lived→disappears after about 10mins
- Therefore, search through the existing research literature before you make any decisions *
- After researchers have defined their terms, they then pay attention to the tools they plan to use to measure their variable(s) of interest
- Behavioural measurements that psychologists make must be valid/reliable

- Validity: Refers to the degree to which an instrument or procedure actually measures what it claims to measure
- It is difficult: researchers can't simply ask people a few questions and then randomly decide that one score qualifies as depressed while the other does not (using the depression scenario beforehand pg.33)
- For the measure to be valid, a particular score would have to differentiate depressed and non-depressed people in a way that accurately maps on how these people actually feel (depressed and non-depressed people will score differently)
- Measurement tool must also be reliable
- A measure demonstrates reliability: when it provides consistent and stable answers across multiple observations and points in time
- Test-retest reliability: Examines whether scores on a given measure of behaviour are consistent across test sessions
- -Example: Scores on a test of depression vary widely each time when you take the test, it is unlikely that your test is reliable
- Alternative-forms reliability: Examines whether different forms of the same test produce the same results (many cases a person will be tested on multiple occasions)
- -Example: Individuals with brain damage might have their memory tested soon after they arrive at the hospital and then one/more points during rehabilitation
- -Giving these individuals the same exact test, possible that any type of improvement is due to practice
- -Having multiple versions of a test that produce the same results (e.g., 2 equally difficult lists of words as stimuli for memory tests), researchers/hospital workers can test individuals on multiple occasions/know that their measurement tools are alike
- Third type of reliability: When observers have to score/rate a behaviour/response
- -Example: Psychologists interested in the effects of nonverbal behaviour when people interact, so they videotape participants solving a problem and then have trained raters count the # of touches/the amount of eye contact that occurred during the experiment
- -Example: participants write down lengthy, open-ended responses to an experimenter's questions; these responses would then be rated on different variables by laboratory personnel
- The catch: More than one person must be rating; otherwise it impossible to determine if the responses were actually measured/if the results were due to the single rater
- Having more than one rater gives you inter-rater reliability
- Inter-rater reliability: Meaning that the raters agree on the measurements that were taken

- Designing an experiment with clear operational definitions and criteria for the raters = have a high inter-rater reliability
- Important that your results are not limited to a small group of people in a single lab.
- Ideal for these results to relate to other groups/situation →so it can be generalizable (making it valid/reliable)

→Generalizability of Results:

- Psychologists primarily interested in understanding behaviour *in general*
- Examining trends/patterns that allow us to predict how *most people* will respond to different stimuli/situations
- Generalizability: Refers to the degree to which one set of results can be applied to other situations, individuals, or events
- -Example: one person told you that a memory-improvement course helped her raise her grades...How useful is this course?
- You would initially view the course favourably – however after further reflection, you'd realize that a # of other factors could have influenced your friend's improvement →she is suddenly paying more attention to her grades
- You would wait until you've heard more about the course before investing your money
- But, if you found out that several 100 people in your city had taken the same course and received benefits
- -Results will appear more likely to predict what would happen if you/other people took the course
- →They are generalizable *
- To increase the possibility that research results will generalize is to study a large group of participants
- Examining/reporting an average effect for that group, psychologists can get a better sense of how individuals are *likely* to behave
- Ideally best to study the entire population
- Population: The group that researchers want to generalize about
- -However it is impossible to find all population members, and persuade them to participate
- Psychologists instead typically study a sample
- Sample: A select group of population members (Once sample has been studied, then results may be generalized to the population as a whole)
- How a sample is selected will determine whether your results are generalizable

- Researchers use a random sample: A sampling technique in which every individual of a population has an equal chance of being included
- -Example: Studying the population at your school → to obtain a true random sample would be to have a computer generate a list of names from the entire student body
- -Your random sample – a subset of this population – would then be identified, with each member of the population having an equal chance of being selected regardless of class standing, gender, major, living situation, etc.
- Not always possible to use random sampling, especially if the results are to be generalized to a large population/to all of humanity
- In these situation researchers use convenience samples: Samples of individuals who are the most readily available
- Psychological research should generalize across time/location
- Research should ideally have high ecological validity: Meaning that the results of a laboratory study can be applied to or repeated in the natural environment
- -Example: Computer based studies testing your ability to pay attention to different stimuli on a computer screen → these artificial situations are assessing human abilities that are used in common situations (driving/finding a friend in a crowded classroom)
- We need to be careful not over-generalize
- -Example: The Mozart Effect → Most of the studies involved adults, yet companies/politicians assumed – with little evidence – that the results would generalize to children/infants
- Scientific literacy involves thinking critically about when it is *appropriate* to generalize results to other groups, times, and locations

→Sources of Bias in Psychological Research:

- Creating objective, reliable, and valid measures important in quality research → many types of bias can be unintentionally introduced by the researcher (called researcher bias)
- -Example: Experimenter treats participants in different experimental conditions differently → making it impossible to know if any differences were due to the experimental manipulation being tested/were instead due the experimenter's behaviour
- Possible for participants (even animals) to introduce their own bias → effects known as subject biases or participant biases
- Sometimes involves a participant trying to figure out what the experimenters are testing/trying to predict the responses that the researchers are hoping to find
- Bias can be introduced by the act of observation itself

- -Example: At Western Electric Company's Hawthorne Works: Researchers went to the factory to study the relationship between productivity/working conditions
- When researchers introduced some minor change in working conditions – adjustment to lighting →workers became more productive for a period of time
- When they changed another variable in a diff. study →having fewer but longer breaks→productivity increased again
- Not obvious to the researchers that any change in factory conditions brought increased productivity →changes were always followed by close attention from the factory supervisors
- Results due to participants noticing that they were being observed, instead to the variables being manipulated
- Hawthorne Effect: A behaviour change that occurs as a result of being observed
- Most psychological research, participants know that they are being observed
- Challenging for experimenters to limit the effect that they have on the results of their own study so that the results are due to variables being studied instead of the participants responding to cues from the researcher

→Working the Scientific Literacy Model:

- Results of psychological studies should provide uncontaminated behaviour
- People who participate in psychological studies usually enter the research environment with curiosity about the subject of the study
- Researchers must withhold as much detail as possible (while still being ethical) to get the best/least biased results
- →What do we know about how bias affects research participants?
- Major concern to studying human behaviour is demand characteristics
- Demand characteristics: Inadvertent cues given off by the experimenter or the experimental context that provide information about how participants are expected to behave
- Demand characteristics range from very subtle to obvious influence on the behaviour of research participants
- Can respond based on social desirability (also known as socially desirable responding)
- Social desirability: Means that research participants respond in ways that increase the chances that they will be viewed favourably (ex. study involves an interview, where the researcher has face-to-face contact with volunteers)
- Demand characteristics causes problem in lab. studies, for participants try to figure out the purpose of the experiment →subtle/obvious cues by researcher gives away the purpose of the study

- -Example: psychologist asks you to put on a heavy backpack, shows you a ramp, asks you to estimate how steep it is
- -Obvious that the experimenter wants to know whether wearing a backpack will affect your judgment about the steepness of the ramp
- In demand effect people are searching for a reason for their behaviour
- If there is no explanation for performing a given set of actions, people tend to draw their own conclusions
- Demand characteristics also presented by researchers/observers introducing biases/unwittingly drawing out the responses they desire
- -Example: Researchers told teachers in 18 diff. classrooms that a group of children had an “unusual” potential for learning →realistically they were just a random selection of students
- After 8 months of schooling, the children singled out showed significant gains in grades, intelligence test scores →relatively stable (but how would this occur if the students were randomly selected?)
- The observers (teachers) assumed those students would do well, thus more likely to pay attention to those students/give them positive/encouraging feedback
- Positive experiences with the teachers motivated these students to improve themselves
- Experimenter bias can be found when people work with animals too
- -When research assistants told that they were working with “bright” rats, learned significantly faster
- -When assistants told they were working with “dull” rats learned slower
- →Researchers made subtle changes in how they treated animals/how they observed/recorded behaviour making possibly making them perform differently
- →How can we critically evaluate the issue of bias in research?:
- Completing rigorous training/following careful scripts when explain experimental procedures to participants reduces experimenter effects (unintentionally manipulating their participants – usually occur accidentally/subtlety)
- To see whether participants’ expectations are influencing the results is to create an additional manipulation → researchers give different groups of participants different expectations of result

