19 Applied Research

Scientific research that attempts to maximize the external validity of findings works to apply the theories, methodologies and results outside the well-controlled laboratory environment. Like with non-experimental designs, giving up experimental control poses challenges for maximizing the internal validity of the scientific process. However, in many applied research studies, there can still be the key manipulated independent variable that allows for strong causal conclusions about how that manipulation affects the measured dependent variable.

In this chapter we will review three related types of research approaches that are all examples of quasi-experimental designs that mix elements of controlled laboratory designs with real world contexts. **Field research** typically refers to procedures testing psychological hypotheses in the conditions in which they are thought to apply. The term field research can also be applied to non-experimental approaches based on observation but when an environmental variable is surreptitiously controlled by a research team, this approach is very similar to experimental research within the laboratory.

Intervention research describes practical applications of research aimed at improving conditions or outcomes for populations related to the participants recruited into the study. These research projects are often very important

for policy questions and how organizations or institutions allocate resources. At their core, the basic methodology often depends on the simple designs described in Chapters 2-9 but ethical considerations about fair treatment of participants often require adding substantial complexity to the research process.

The term **translational research** is often used in medical or public health contexts to describe an attempt to bring a basic scientific finding, e.g., from biology, into practical medical applications aimed at improving health. In these areas, the term often serves to make a distinction among types of research where **basic science** is laboratory work aimed at understanding a core mechanism with an idea that this finding will later be used in translational science to provide some benefit. While these terms are used more in related fields like neuroscience and public health than psychological science, there is the same distinction between the value of well-controlled laboratory basic science with optimal internal validity and the messier process of applying a finding outside the lab with better external validity.

One of the major concerns with applied research is fitting this kind of science into the ethical framework for research. One of the benefits of laboratory work is that participants are aware they are voluntarily opting into the research process, can provide informed consent and choose whether they wish to continue through the study. Many of these factors are more difficult to achieve in applied research, necessitating careful discussions with the IRB providing oversight to ensure that the research properly balances the risks to participants and the benefits of the science. In addition, many kinds of intervention or translational research involve manipulating a factor that is thought to be a benefit to the participants. This is challenging to do and follow normal practice about justice, being fair to all participants. If the intervention being studied is effective, the control comparison condition is being unfairly deprived of this benefit. If the intervention is ineffective, the participants receiving it might be treated unfairly. However, there are a handful of common methodological approaches that minimize this issue as much as possible.

Example: Laptops in the classroom

Many teachers have been concerned that the use of laptop computers in classroom environments are not conducive to effective learning. Although laptops can help with notetaking, they also allow the possibility that students' attention might become distracted from the content being covered and interfering with learning. In laboratory environments, there is a robust literature on the phenomenon of *dual-tasking* that documents how attempting to do two cognitive tasks simultaneously affects performance on each task. There is also some research on a concept called *desirable difficulty* that suggests that memory might be better for material that one works a little harder to process. In a laboratory setting, taking notes manually (handwriting) has been shown to lead to better memory than typing (which is theoretically easier). However, neither of these phenomena have been convincingly demonstrated to lead to worse outcomes in a practical classroom environment.

Consideration of the challenges embedded in taking on a rigorous field research study to quantitatively test the effect of laptops illustrates why this study is deceptively difficult to carry out. On the surface this looks like a simple two-group design. Randomly assign half the students in some sample to having laptops available in class and the other half to relying solely on paper and pencil. This would be simple operational definition of our independent variable. As a dependent variable, we might be tempted to simply use final grade in the class at the end of the term as a measured operational definition of learning effectiveness.

As simple as this looks, as soon as we consider constructing a procedure to implement such a design, we run into problems. First, one of the concerns about laptops in the classroom is that they may actually impose their distracting effects on students sitting nearby. If somebody sitting in front of you is viewing non-class material, you might find yourself in an imposed dual-tasking condition where that information is competing with attention to the intended class content. As a result, creating properly controlled access to

laptops or freedom from their distraction might require assigning participants to conditions at the class level, i.e., one section of a class has laptop use allowed and another does not. However, if the two sections are not taught by the same instructor, we have an important uncontrolled extraneous variable (instructor) that is unfortunately confounded with our IV.

Further, if we pause to consider some of the other extraneous variables that might affect our study, we realize that our hypothesis may depend on the class content being studied. Humanities classes, STEM classes, small seminars and large lecture hall classes might all be affected by possible laptop distractions in different ways. Certain teaching styles might be more or less affected. It will also be difficult or impossible to control any possible bias due to expectancy effects from the students or the teacher, who will certainly be aware of the condition being imposed on the participants. Even though we have tried to conceptualize this research as a field study to improve external validity, all the choices about which classes are recruited into participation end up potentially limiting the applicability of our research.

Recruiting participants in a classroom context is also a potentially thorny issue. Students should feel free to take classes at their own choice, so it does not appear that any recruiting procedure is necessarily purely voluntary. And participants should always feel free to withdraw from the study, so the possibility of bias due to students being more likely to drop a class where they are unable to use their laptops for notetaking. If the intervention is effective, we also have to acknowledge that students who were required to take a class with laptops in use might have essentially have had their grades lowered by random assignment to conditions, which is generally considered unethical practice.

The main point of this example is to consider what seems like a fairly simple and straightforward two-group experimental design and realize that as an intervention study, it is exceptionally difficulty to carry out a study with a very high level of internal validity. As a result, it is likely that any general policy decisions would necessarily rely on non-experimental methods (observation, correlational research) or laboratory studies that do not test the effect of the

intervention in a real context. In general with interventions, The elements that can not be approached with the best experimental design principles do not mean that valid inferences can not be drawn, but that our critical evaluation skills for identifying alternate explanations need to be applied to these studies. At the same time, a well-constructed and well-executed intervention design should be appreciated for the methodological care and effort that needs to be invested to make this type of research effective.

Quasi-Experimental Design

Quasi-experimental designs blend elements of experimental and non-experimental research designs. The goal in this approach to research is to try to capture aspects of the higher external validity of non-experimental designs, but still try to support strong causal statements about the relationship between the independent variable(s) and dependent variable.

A laboratory study is a study that is conducted in the laboratory environment where principles of experimental control such as constancy and counter-balancing can be used. In contrast, a field study is a study that is conducted in the real-world, in a natural environment. This increases external validity but at some cost to internal validity. Field studies can be real experimental designs with manipulated independent variables. In theory this allows for stronger causal inferences than non-experimental designs. However, the lack of full control over extraneous variables makes these less reliable and more vulnerable to Type 2 errors where the data are too noisy to reliably support strong conclusions.

Internal validity challenges in field research come from the same sources as all other experimental research, mainly arising from accidentally confounded variables or participant selection bias. In laboratory research, extraneous variables can be identified, and an experimental procedure constructed to minimize their effects (constancy) or at least de-confound these (counter-balance). In field research, there is often a very limited ability to employ these techniques, so it is more likely that conclusions will have to be

considered in the context of alternate explanations (like in non-experimental research).

Laboratory experiments that are designed properly typically have high internal validity. Internal validity refers to the degree to which we can confidently infer a causal relationship between variables. When we conduct an experimental study in a laboratory environment we have very high internal validity because we manipulate one variable while controlling all other outside extraneous variables. When we manipulate an independent variable and observe an effect on a dependent variable and we control for everything else so that the only difference between our experimental groups or conditions is the one manipulated variable then we can be quite confident that it is the independent variable that is causing the change in the dependent variable. In contrast, because field studies are conducted in the real-world, the experimenter typically has less control over the environment and potential extraneous variables, and this decreases internal validity, making it less appropriate to arrive at causal conclusions.

But there is typically a trade-off between internal and external validity. External validity simply refers to the degree to which we can generalize the findings to other circumstances or settings, like the real-world environment. When internal validity is high, external validity tends to be low; and when internal validity is low, external validity tends to be high. So laboratory studies are typically low in external validity, while field studies are typically high in external validity. Since field studies are conducted in the real-world environment it is far more appropriate to generalize the findings to that real-world environment than when the research is conducted in the more artificial sterile laboratory.

There are also field studies which are non-experimental in nature because nothing is manipulated. These will typically rely on observational methods and produce observed correlations between variables that are measured in non-laboratory contexts.

Extraneous variables in Field Research

In a field research protocol, the procedure is designed to be carried out in the real world. How are participants going to be identified to "recruit" into the study? In many cases they might not even know they are in the study, raising questions about ethics that must be addressed. Beyond ethics, the implications of the recruiting method may have important implications for the generalizeability of results if there are any concerns that the recruiting method accidentally biases the sample.

Many field research procedures effectively require experimenters to become actors to carry out the procedures of the study as if it were a script. The validity of the research can end up depending on the acting ability of the experimenters and/or confederates. Field research procedures can also put a lot of pressure on observational techniques that require scoring aspects of behavior that can have a significant subjective component. Where possible, standard techniques should be used such as multiple raters and assessing consistency, or to structure the procedure so that the raters are blind to the experimental condition. These are not always possible and can weaken the internal validity of the results.

The process of "recruiting" participants can also be highly sensitive to the place, time and date of planned experimental data collection in field research. These can inadvertently incorporate bias into the recruiting procedures that needs to be considered when interpreting the results. Recruiting bias typically affects the external validity of the research with the idea participants found at a different time or place might have behaved differently. While field research is generally higher in external validity, recruiting processes can pose a challenge.

Ethics in Field Research

Field research has several specific characteristics of ethical concerns that occur frequently. The most obvious is that data is being collected from

participants without their awareness that they are in a research study. That is an elevated level of deception that is almost always inherent in this type of research that needs to be evaluated by the IRB before data collection can occur.

Field research also often involves areas of psychology that are sensitive since these are the kinds of topics for which behavior can be markedly different in the laboratory. In a laboratory setting, when participants are asked about attitudes related to stereotypes or high-risk behavior, their responses may incorporate expectations about what they believe the experimenter things is most socially appropriate. Collecting data on these ideas in the field may provide a much more direct insight into what people actually do rather than allowing them to report what they think is commonly considered acceptable. This exacerbates the issues related to deception and lack of clarity about informed consent. However, since these are often question of great scientific interest, they can certainly be judged to be ethical in that the benefit of the research outweighs the costs to the participant in risk, or embarrassment.

Intervention research

A great deal of psychological science research aims to further our understanding of human behavior and the underlying cognitive processes. Studies that have a robust theoretical foundation and are designed to add new facts to what is known are termed "basic science" research. However, many psychological science research areas of study are aimed more directly at identifying ideas that might immediately and positively impact the world which are described as "applied research." These ideas are not necessarily mutually exclusive, but in practice, basic science tends to be carried out within the laboratory with maximum control and the best internal validity. Applied research is typically done as field research where there is a controlled independent variable that involves the intervention.

In medical research, studies that aim to establish the effectiveness of a new drug are done as intervention studies. These follow many of the same

methodological issues, but often do not involve psychological measures. In addition, many of these studies are done in the context of very specific methodological practices related to ethics and rigor (e.g., double-blind methods) that not always available in psychological studies. Research on improvements in educational practice captures more of the challenges of psychological research as there are important questions about efficacy of interventions that we want to observe in the context they would be used, e.g., a classroom. But in this work, it is challenging to implement all the preferred aspects of experimental control that we would want in laboratory research.

Basic Intervention Designs

In a **one-group post test only design**, a treatment is implemented (or an independent variable is manipulated) and then a dependent variable is measured once after the treatment is implemented. Imagine, for example, a researcher who is interested in the effectiveness of an anti-drug education program on elementary school students' attitudes toward illegal drugs. The researcher could implement the anti-drug program, and then immediately after the program ends, the researcher could measure students' attitudes toward illegal drugs.

This is the weakest type of quasi-experimental intervention design. A major limitation to this design is the lack of a control or comparison group. There is no way to determine what the attitudes of these students would have been if they hadn't completed the anti-drug program. Despite this major limitation, results from this design are frequently reported in the media and are often misinterpreted by the general population. For instance, advertisers might claim that 80% of women noticed their skin looked bright after using Brand X cleanser for a month. If there is no comparison group, then this statistic means little to nothing.

In a **one-group pretest-post-test design**, the dependent variable is measured once before the treatment is implemented and once after it is implemented. Let's return to the example of a researcher who is interested

in the effectiveness of an anti-drug education program on elementary school students' attitudes toward illegal drugs. The researcher could measure the attitudes of students at a particular elementary school during one week, implement the anti-drug program during the next week, and finally, measure their attitudes again the following week. The pretest-posttest design is much like a within-subjects experiment in which each participant is tested first under the control condition and then under the treatment condition. It is unlike a within-subjects experiment, however, in that the order of conditions is not counterbalanced because it typically is not possible for a participant to be tested in the treatment condition first and then in an "untreated" control condition. This is basically a non-experimental approach but gets used fairly often because of the ease of implementation.

The standard approach to ruling out the threats to internal validity in one-group designs is by designing the research protocol to include a **control group**, one that does not receive the treatment effect. A control group would be subject to the same threats from history, maturation, testing, instrumentation, regression to the mean, and spontaneous remission and so would allow the researcher to measure the actual effect of the treatment (if any). Of course, including a control group would mean that this is no longer a one-group design.

It can be difficult to identify an appropriate control task. The best control tasks manage demand characteristics for the participants. For example, a study on improving educational outcomes using intense extra training with personal tutors is going to have the issue that participants receiving tutoring are aware that something is being done to help them. If the control condition is simply additionally assigned, optional reading, the control condition may not respond to the post-test with the same expectations. However, a really well-matched control condition can be difficult to construct in many cases. As a result, it is not that uncommon to use the simplest alternative and simply not intervene. This is sometimes called a *wait-list control* condition to reflect the fact that recruiting is done among people who are seeking the benefit of the intervention, who are then randomly assigned to receive the intervention

or not. Those not receiving the intervention are designated the waitlist control condition.

Once a control condition has been identified this approach can be used as a simple two-group design comparing the intervention to the control condition. It can also be combined with a pre/post design where the DV is assessed both before and after either the intervention or control condition. Note that in this approach, the statistical tool that would be used to measure the intervention effect would be an ANOVA as this is now a 2x2 design. A successful intervention would cause an interaction where the difference between the first and second assessments would be larger for the group that received the intervention.

Using a control condition increases the internal validity of the design at the cost of an ethical problem. Participants being recruited into an intervention design study are often seeking help and randomly assigning them to the waitlist control condition is not entirely fair. A method for mitigating this problem is to use a **phased design** such as an **AB** or **ABA** design. In these designs, interventions alternate with baseline control conditions. The advantage of this approach is that the comparison group can alternate with the opposite conditions. The simplest design of this type is to use a waitlist control condition in the first phase and then in the second phase, apply the intervention to the control group. Contrasts on the DV at the end of the first phase are used as the core statistical test, but to manage ethical concerns the control group then obtains the potential benefit of the intervention in the second phase.

A more complex variation of the phased designs is to combine pre/post testing and multiple intervention phases into a **crossover interaction design**. In this design, all participants receive a pre-test at recruitment. They are then randomly assigned to intervention or control condition in the first phase. Then everybody receives a post-test. Then the intervention/control conditions are swapped in the second phase, which is then followed by a second post-test. While this is potentially complex and time-consuming to implement, a series of potentially interesting, related hypothesis are tested.

At the first post-test, differences in performance or differences in improvement from the pre-test across groups provide a measure of the intervention effectiveness. This is usually the primary goal of the research study. In addition, performance from the first post-test to the second for the intervention group provides a measure of whether the benefit of the intervention fades over time. This is very commonly an important question in interventions, not only whether they help but whether the benefits persist. Performance of the group that gets the intervention on the second phase can be compared as a pre/post design that effectively replicates the pre/post design of the early intervention group. This embedded replication, if successful, helps increase confidence in the efficacy of the intervention. And since both groups eventually get exposure to the intervention, ethical issues caused by random assignment to conditions are mitigated.

Longitudinal Research

Intervention research is often done over a more extended time period that our simpler experimental designs run in an hour or two in a laboratory. This feature requires consideration of some history effects like those that apply to within-participant designs. One such effect is the phenomenon of regression to the mean, which is a natural consequence of variability in human performance. A well-known example was reported in discussions with instructors who provided feedback to pilots who felt that it was most important to harshly criticize poor performance and ineffective to compliment good performance. If some of the daily variation in performance is simply due to random variance, today's poor performers are likely to improve to perform near the mean on the next session. At the same time, better performers will fall back to their average performance. The tendency for performance to center around a mean level plus variability leads to the phenomenon of regression to the mean. In this case, today's top performers are most likely to do nominally worse next time, and the bottom performers ought to do better even without a robust learning effect. This effect can look a lot like a response to feedback, as in this example, but with the data

described here, it is impossible to tell which account for the finding.

A related idea in clinical or medical intervention research is the idea of spontaneous remission, which is the rate of improvement that is not influenced by an intervention. This idea is another rationale for placebo control groups. Participants in the control group may improve due to demand characteristics (knowing they are in a study) or may simply improve due to other effects related to the passage of time. Intervention research aimed at treatment very rarely relies on one-group designs due to the risk of improvement being due to effects unrelated to the intervention.

Mortality Effects

The term **mortality effect** refers to a challenge to the internal validity of a study based on participants dropping out of the study. In this context mortality often does not refer to death of participants but can simply reflect a decision not to continue with participation in the research project. Since all participation in research should be voluntary, if one of the conditions is seen as unpleasantly onerous and not worth the time required or compensation offered, participants have the right to opt out of the study. If one of the experimental conditions experiences a higher rate of dropout, then this can bias the subsequent comparison between groups.

To illustrate this problem, consider a simple research design comparing the effect of reward and punishment on learning. Participants in the reward condition are asked to do a moderately challenging perceptual task and are rewarded with 50 cents for each correct response. A separate group of participants attempts to learn the same task but are punished with a mild electric shock for each incorrect response. If the study runs over several sessions and several days, a very plausible outcome is that the participants receiving shocks will exhibit higher levels of performance. However, without detailed information about participant dropout, we can draw no conclusion about the relative value of reward versus punishment. It is very likely in this simple design that participants in the punishment group who were poor at

learning the task would drop out of the study. If more of the poor learners drop out of the punishment than the reward group, any differences at the end of the study could be entirely due to differential mortality across groups.

This effect does not only happen in intervention research but can occur with any longitudinal design where data is collected across many days, sessions or even weeks or years. Research using this approach will generally need to report detailed information on participant retention across the study and compare this across all independent variables to try to rule out this problem. Intervention research with a challenging intervention can be weakened by participants opting out of the study because the intervention is unpleasant (e.g., diet or exercise interventions). Even participants in the control condition can opt out of participation if they become aware they are temporarily in a control or placebo condition. In some designs, recruiting and compensation practices are design explicitly with reducing participant dropout by including compensation bonuses for completing the study. These techniques are always reviewed carefully by the IRB for any accidental coercion to participation that they might imply.

Mortality effects in research are closely relate to the idea of survivorship bias in which inferences about data available at the end of a study or other ongoing process can be distorted by the missing data unavailable at the end. The most famous example of this was an analysis of military airplane damage done in World War II on planes that returned from combat. The key inference was that the undamaged parts of the planes that returned were likely the most vulnerable and critical regions based on the inference that the planes that had been hit in those regions failed to return. This idea has been raised as an issue in highly competitive fields (business, athletics, acting) where studying only people who have succeeded may not reveal critical differences that actually predict success for which information from less successful people needs to be included, essentially a sampling issue in the language of experimental research.

Maintaining Good Scientific Practice

Intervention research is also an area where challenges can arise related to conflict of interest and experimenter bias and can make maintenance of best rigor practices difficult. Because intervention research is challenging to implement, it is generally done with a strong hypothesis and expectation about the outcome by the research staff. Research procedures should be designed with attention to the possibility of bias in scoring the dependent variable or any aspect of implementing the conditions manipulated for the independent variable(s). In addition, large scale intervention research is resource intensive to do effectively, often meaning that a funding agency is supporting the research, usually with the idea of wanting to find effective interventions to address social problems. This can create the appearance of conflict of interest in that successful research can have fairly immediate direct benefits for researchers.

The existence of a potential conflict of interest does not mean that research cannot be carried out with a high degree of rigor and integrity. Standard practice is to disclose all information related to possible conflicts of interest so that reports of results can be interpreted in that context. Results in this context may be considered with a heightened degree of skepticism, but should not be treated as if they cannot be accurate. As with all research, confidence in conclusions accrues over time, replication and extension to novel operational definitions, samples and research contexts.

Key Takeaways

- Applied research methodologies use research methodologies outside the well-controlled laboratory environment.
- Field research aims to strengthen the external validity of scientific findings by establishing that the independent variable affects behavior in real-world contexts.
- Intervention research is designed to test manipulations that are hypothesized to have positive benefits for participants. These studies are often part of translational research programs to bring basic science to practice.
- These quasi-experimental designs involve the manipulation of an independent variable in order to draw causal conclusions. However, there are greater challenges from uncontrolled extraneous variables, and may not be able to fully use random assignment to conditions or counterbalancing. For these reasons, quasi-experimental research is generally higher in internal validity than non-experimental studies but lower than true experiments.
- Intervention research frequently uses within-participant methods to help control for participant variables and particular care must be taken to address the usual history effects and also any influence of the intervention on participant dropout or mortality effects.
- Ethical practice in intervention research requires using designs that provide fair access to the beneficial intervention to all participants whenever possible. This often leads to phased designs or a crossover intervention design approach.
- Ethical practice in field studies often does not permit the standard recruiting and informed consent process for participants. This kind of research needs to be carefully reviewed by the IRB to ensure that the research still meets expectations for balancing risk and benefits of research.

Exercises

TBA

20 Development and Neuropsychology

The patient known by the initial H.M. is one of the most famous case studies in psychology. H.M. suffered from severe epilepsy, a syndrome noted for severe and unpredictable seizures, that was intractable to all available medical treatment. Epileptic seizures arise from neural dysfunction that is often associated with a specific brain region, or foci, that sets off a brain-wide electrical storm. Treatment is usually pharmacological, but in severe cases where drug treatments are found to be ineffective, surgical resection (removal) of the foci region is sometimes done to reduce or eliminate subsequence seizure activity. The surgical approach can be very effective as the dysfunctional region is typically not contributing to general cognitive function in any meaningful way, so removal has no long-lasting effects on any general cognitive process. However, the surgery has to be done carefully to avoid removing any cortical brain regions that might still be contributing to cognition, known as *eloquent* cortex. Epileptic foci often occur in or near parts of the brain associated with memory so modern treatment approaches include a long period of cortical mapping to attempt to identify all the regions near the epileptic foci that need to be left intact.

H.M.'s surgical treatment predated the modern understanding of critical

brain regions for specific cognitive functions and the tools to do very careful mapping of still-functioning cortical regions had not yet been developed. In fact, at the time, the prevailing theory of cortical function was one of *equipotentiality*, the idea that all brain regions contributed together to complex cognitive functions such as memory. It was not thought that any one region of the brain would support a function as complex as memory, although there were some hints from other neuropsychological studies that there might be more *modular* organization to cognitive function in the brain.

To treat H.M., Scoville & Milner (1957) described a large, bilateral resection of cortical regions around the medial temporal lobe of his brain. For the epilepsy and subsequent seizure activity, the treatment was a success, however it soon became clear that H.M.'s cognitive function had been substantially altered in an unexpected way. H.M. exhibited a pattern of memory impairment that came to be called anterograde amnesia. The key feature of this syndrome is the inability to form new memories. Previously acquired knowledge and memories were intact, but nothing new could be added to his memory store. Almost all other aspects of general cognitive function were completely intact. He could carry on conversation, carry out problem solving and decision-making exercises. His knowledge of language and basic semantic information about the world was fully functional. However, nothing about his experiences after the surgery was stored, resulting in the peculiar experience that if you left the room and returned, he was unaware that

Henry Molaison

After his passing, the famous patient H.M.'s real name was revealed to the broader scientific community. He was referred to by his initials in the scientific literature for the majority of his life as a measure to protect his privacy. This was his family's preference instead of the potential notoriety associated with one of the most famous individual case studies in the history of neuropsychology.

he had ever seen you before.

This case was revolutionary in the understanding of the organization of memory and the brain overall. His selective impairment indicated that complex cognitive processes such as memory were localized to or at least dependent on specific structures. Once the specific pattern of cognitive impairment exhibited by H.M. had been described, many other patients with similar patterns of memory impairment were identified. Most of these patients have memory deficits much less severe, but exhibit the same distinction between acquisition of new memories for facts and events with largely intact retrieval of older memories. These patients also exhibit intact learning of procedural skills acquired through repetitive practice even though they do not remember the event of practicing. The study of patient HM also inspired a wide range of further research looking for other cognitive processes that could be strongly connected to specific brain regions, an area of research known as cognitive neuropsychology.

In this chapter we will review the methodology associated with analysis of famous single cases such as H.M., **case study** research. In addition, we will review some techniques associated specifically with **developmental research**, which aims to characterize changes in psychological processes across the lifespan. Each of these methodologies will have elements that are fundamentally non-experimental. Particularly interesting single cases are not assigned to that condition and age-related change is not something possible to control experimentally. However, techniques to study these areas often involve using experimental methods together with the non-experimental aspects of research to better understand the underlying phenomena.

Case Study Research

A case study is an in-depth examination of an individual. Sometimes case studies are also completed on social units (e.g., a cult) and events (e.g., a natural disaster). Most commonly in psychology, however, case studies provide a detailed description and analysis of an individual. Often the individual has a rare or unusual condition or disorder or has damage to a specific region of the brain. These studies can bear some similarity to non-experimental research approaches that are qualitative or observational.

Like many observational research methods, case studies tend to be more qualitative in nature. Case study methods involve an in-depth, and often a longitudinal examination of an individual. Depending on the focus of the case study, individuals may or may not be observed in their natural setting. If the natural setting is not what is of interest, then the individual may be brought into a therapist's office or a researcher's lab for study. Also, the bulk of the case study report will focus on in-depth descriptions of the person rather than on statistical analyses. With that said some quantitative data may also be included in the write-up of a case study. For instance, an individual's depression score may be compared to normative scores or their score before and after treatment may be compared. As with other qualitative methods, a variety of different methods and tools can be used to collect information on the case. For instance, interviews, naturalistic observation, structured observation, psychological testing (e.g., IQ test), and/or physiological measurements (e.g., brain scans) may be used to collect information on the individual.

Case studies as illustrating descriptions

The history of psychology is filled with influential cases studies, such as Sigmund Freud's description of "Anna O." Sigmund Freud used the case of a young woman to illustrate many principles of his theory of psychoanalysis (Freud, 1961). (Her real name was Bertha Pappenheim, and she was an early

feminist who went on to make important contributions to the field of social work.) Anna had come to Freud's colleague Josef Breuer around 1880 with a variety of odd physical and psychological symptoms. One of them was that for several weeks she was unable to drink any fluids. According to Freud,

She would take up the glass of water that she longed for, but as soon as it touched her lips she would push it away like someone suffering from hydrophobia.…She lived only on fruit, such as melons, etc., so as to lessen her tormenting thirst. (p. 9)

But according to Freud, a breakthrough came one day while Anna was under hypnosis.

[S]he grumbled about her English "lady-companion," whom she did not care for, and went on to describe, with every sign of disgust, how she had once gone into this lady's room and how her little dog—horrid creature!—had drunk out of a glass there. The patient had said nothing, as she had wanted to be polite. After giving further energetic expression to the anger she had held back, she asked for something to drink, drank a large quantity of water without any difficulty, and awoke from her hypnosis with the glass at her lips; and thereupon the disturbance vanished, never to return. (p.9)

Freud's interpretation was that Anna had repressed the memory of this incident along with the emotion that it triggered and that this was what had caused her inability to drink. Furthermore, he believed that her recollection of the incident, along with her expression of the emotion she had repressed, caused the symptom to go away.

As an illustration of Freud's theory, the case study of Anna O. is quite effective. As evidence for the theory, however, it is essentially worthless. The description provides no way of knowing whether Anna had really repressed the memory of the dog drinking from the glass, whether this repression had caused her inability to drink, or whether recalling this "trauma" relieved the symptom. It is also unclear from this case study how typical or atypical Anna's experience was.

Case studies are useful because they provide a level of detailed analysis not found in many other research methods and greater insights may be gained from this more detailed analysis. As a result of the case study, the researcher may gain a sharpened understanding of what might become important to look at more extensively in future more controlled research. Case studies are also often the only way to study rare conditions because it may be impossible to find a large enough sample of individuals with the condition to use quantitative methods. Although at first glance a case study of a rare individual might seem to tell us little about ourselves, they often do provide insights into normal behavior. The case of H.M. provided important insights into the role of the hippocampus in memory consolidation.

However, it is important to note that while case studies can provide insights into certain areas and variables to study, and can be useful in helping develop theories, they should never be used as evidence for theories. In other words, case studies can be used as inspiration to formulate theories and hypotheses, but those hypotheses and theories then need to be formally tested using more rigorous quantitative methods. The reason case studies shouldn't be used to provide support for theories is that they suffer from problems with both internal and external validity. Case studies lack the proper controls that true experiments contain. As such, they suffer from problems with internal validity, so they cannot be used to determine causation.

Neuropsychological Case Studies

When case studies are based on known specific damage to brain regions provide some ability to draw a causal inference – damage to that area leads to the observed impairment – the difficult question to answer is about generalizeability. It can never be fully determined if the same damage to any brain region will end up having the same impact for everybody else. Essentially, we cannot tell with current methods how similar brain organization is across people. The power of case studies is the ability to document even for one person that the observed pattern of damage and

intact function is possible to occur.

Another particularly impactful case study for this idea is the case of **Phineas Gage**. In the mid-19th century, Phineas Gage suffered an injury in a factor that led to an iron bar being blasted through part of his skill causing the loss of an eye and damage to the prefrontal cortex just above the eye. He survived the injury and additionally surprisingly exhibited very little impairment in general cognitive function. He did not exhibit difficulty in speaking (aphasia), recognizing objects (agnosia) or any pattern of motor impairment that is often associated with brain damage due to stroke (that impact motor control regions). However, reports at the time documented a very robust change in his personality and mannerisms. His behavior became extremely rude, marked by gross profanity and he appeared to unable to follow plans.

Unlike H.M., other patients with damage to the general region impacted by Gage's injury do not exhibit precisely the same pattern of impairment. Deficits like Gage are termed problems with *comportment* that do occur in some syndromes associated with prefrontal cortex damage. Modern studies of cognitive neuroscience have found a wide range of cognitive operations that depend on this brain region and also documented that there is great variety across individuals in how this system is organized. This part of the human brain is also notable as being the most distinct from all other great apes, likely reflecting the parts of the brain most greatly changed in the evolution of modern humans. As a case study, Gage clearly indicates a relationship of comportment and planning to these regions and then served to motivate studies since to identify the neural basis and operation of these processes.

Cognitive Neuroscience

The field of Cognitive Neuroscience reflects research on the understanding of the relationship of the neural operation of the brain to human cognitive processes. Many modern studies of this relationship have been inspired

by case studies of specific damage and constrained patterns of cognitive impairment. Some of this research is done with neuropsychological methods of looking for new cases with similar patterns or differing patterns associated with damage to similar brain regions. A great deal of this research aims to build from the neuropsychological studies to studies of cognitively healthy operation using methods of neuroimaging.

Neuroimaging methods are techniques for collecting data about brain function noninvasively from humans. Examples of these techniques include functional magnetic resonance imaging (fMRI), electro-encephalography (EEG) and transcranial magnetic stimulation (TMS). Detailed discussion of these methods is beyond the scope of this text but depends on the same underlying research method processes used here. Constructs must be implemented the same way we draw inferences from simpler behavioral measures in the experimental psychology methods described here.

Developmental Research

In this chapter, we will review a few specialized research methodologies that apply to research across human age range. Research with young children or infants depends on many of the basic experimental design techniques discussed previously but requires some additional consideration of the kinds of dependent measures that are available and the ethics of research with *vulnerable populations* (Chapter 8). In addition, research comparing changes across age introduces methodological terminology contrasting cross-sectional and longitudinal designs.

Developmental research often refers to experimental methodology applied to young children. Obviously, the operational definitions for this kind of research need to account for the participant population. The research protocol and especially any task instructions need to be prepared to be age appropriate so that the participants understand the experiment details. The dependent variable needs to be a task that the children can perform.

In research with very young infants, measuring behavior can be challenging. Infants can indicate preference or in some cases familiarity by measures of preferential looking. In these paradigms, the infant is presented with a display that contains two different kinds of information, e.g., on two sides of a display. The infant is carefully observed to identify which direction they are looking or for how long they look. This can be used to infer babies' understanding of the physics of the world around them by demonstrating that they look longer at displays constructed to contain apparent physical violations of collisions or violations of numeracy.

Among the challenges of implementing this technique is the need to make subjective ratings of the direction of looking. This is often done with a blind-rater technique where a camera records the direction of looking from an angle where the rater cannot see what is visible to the infant. The rating process done this way avoids any potential bias in rating based on knowing what the infant was supposed to do based on the experimental hypothesis. This research is still challenging to carry out because of the potential for many extraneous variables unrelated to the experiment affecting looking. These studies also frequently have a challenge in accumulating enough participants to support robust statistical inference. Not every child is comfortable in the experimental situation and research is often highly restricted to very specific age ranges (e.g., 3-6 months), making the available population very limited.

An additional practical challenge is that in many of these procedures, it is necessary to have the implicit participation of an adult who is a parent or guardian of the child. It is common in these paradigms to have the child seated on the lap of a parent in order to maintain comfort so that behavior can be observed. However, this can introduce concerns about the child's behavior being biased by parental expectations. The parent has effectively become part of the research staff and needs to try to act in a consistent, unbiased manner so that behavioral differences solely reflect the child.

Cross-Sectional and Longitudinal Research

When psychologists wish to study change over time (for example, when developmental psychologists wish to study aging) they usually take one of three non-experimental approaches: cross-sectional, longitudinal, or cross-sequential. **Cross-sectional** studies involve comparing two or more pre-existing groups of people (e.g., children at different stages of development). What makes this approach non-experimental is that there is no manipulation of an independent variable and no random assignment of participants to groups. Using this design, developmental psychologists compare groups of people of different ages (e.g., young adults spanning from 18-25 years of age versus older adults spanning 60-75 years of age) on various dependent variables (e.g., memory, depression, life satisfaction). Of course, the primary limitation of using this design to study the effects of aging is that differences between the groups other than age may account for differences in the dependent variable. For instance, differences between the groups may reflect the generation that people come from (a cohort effect) rather than a direct effect of age. For this reason, longitudinal studies, in which one group of people is followed over time as they age, offer a superior means of studying the effects of aging. However, **longitudinal studies** are by definition more time consuming and so require a much greater investment on the part of the researcher and the participants.

A third approach, known as cross-sequential studies, combines elements of both cross-sectional and longitudinal studies. Rather than measuring differences between people in different age groups or following the same people over a long period of time, researchers adopting this approach choose a smaller period of time during which they follow people in different age groups. For example, they might measure changes over a ten year period among participants who at the start of the study fall into the following age groups: 20 years old, 30 years old, 40 years old, 50 years old, and 60 years old. This design is advantageous because the researcher reaps the immediate benefits of being able to compare the age groups after the first assessment. Further, by following the different age groups over time they can subsequently

determine whether the original differences they found across the age groups are due to true age effects or cohort effects.

Research across the lifespan also needs to be sensitive to the implementation of best ethical practices across age ranges. Participants under the age of 18 cannot provide written consent to participate in research but are generally consulted in addition to a parent or guardian consenting to their participation. The research protocol still needs to be arranged to acknowledge that the child still understands that participation is voluntary. Interactions between parents and children can be complex to anticipate, especially for older children such as adolescents who have different expectations about their preferences for research or privacy of data. All participants under 18 are considered *vulnerable populations* and oversight of research practices is elevated compared with research on cognitively healthy adults.

Research that aims to understand changes in behavior or cognition as a result of clinical syndromes is often termed neuropsychological research. Much of this work is done at the other end of the lifespan development with older adults. This work can be done as systematic research across groups of patients with clinical syndromes such as Alzheimer's disease or Parkinson's disease (and also with younger patients for syndrome such as schizophrenia). The cognitive impairments associated with some neuropsychological or neurological research means that these patients are sometimes also treated as vulnerable populations, requiring consent of guardian, etc. In general, these studies just require attention to the basic element of respect for persons for these older adults and acknowledgment of their challenges. This area of research is also one where the method of single case studies has been used to characterize particularly interesting single patients with unexpected patterns of cognitive function.

Developmental stages

A common use of cross-sectional studies in young children is to identify and characterize developmental stages as children grow from infancy. In these designs, participants of specific age ranges will be compared. For example, measures related to object permanence might been seen to be different for infants at 4 months of age compared with 8 month olds. This research will usually follow the familiar form with manipulated independent variables and a measured dependent variable. Age, of course, is a participant variable not under experimental control in cross-sectional studies. Developmental studies will often use fairly creative dependent measures because very young participants may not provide accurate verbal answers.

For pre-verbal infants, measures of behavior often focus on subtle shifts based on eye-gaze or attention. These are examples of measures that need rigorous scoring procedures to ensure accuracy of data and to avoid bias. These techniques often rely on video recordings of the participants that are scored by blind raters unaware of the experimental condition of the measure. Further complicating these methods is that very young infants must often be tested while on a parent's lap in order to observe attentive behavior. In these cases, bias can creep into a design through expectations of the parents. In some cases, it can even be necessary to occlude the vision of the parents, e.g., through a blindfold, to keep them unaware of the experimental condition in which the infant is in.

Developmental studies also are notably difficult and time-consuming the carry out due to the complexity of designs, unpredictability in behavior of children and the need to have well-controlled face to face interactions. Recruiting can also be a significant challenge when very specific age ranges are targeted. Challenges in scheduling the research protocol can lead to a participant no longer meeting the intended age requirements for the planned study.

Research on children also notably requires informed consent from a parent or guardian. However, the willingness of the parent to have the child participate is not necessarily the same as willingness of the child, which makes confident

compliance with general ethical procedures complex. For older children, dual consent processes may allow children to consent verbally in conjunction with written consent from the parent. Any discomfort or disinterest in the research protocol has to be considered as a possibly interest in withdrawing from the protocol. Particular care must be paid to participant privacy, especially with video recordings of behavior which have to be handled so that access is restricted to study personnel and these are stored securely.

In addition, recent concerns about creating a robust culture of child protection raised the question of whether researchers are required to report any evidence they might observe of mistreatment of children by the parent or guardian. A substantial debate arose around this question. Research team members are trained in research and are not necessarily experts in detecting issues related to child welfare and safety. Mistakes leading to accusations of parents inappropriately would raise a new form of harm emerging from research participation. However, not acting on obvious evidence of abuse seems entirely unethical. Common current practice is to consider researchers in many developmental labs as *mandatory reporters* who are required to report obvious evidence of mistreatment. However, detailed reporting mechanisms aim to provide expert review of concerns to best maintain fair treatment of all participants.

Life-span Development

While much developmental research is aimed at understanding growth and maturation from infancy to childhood and then adulthood, research on late-life aging has some similar methodological challenges. Ethical concerns about voluntary consent, patient privacy (especially for health records) and even mandatory reporting (elder abuse) are present although to a much lesser extent. The primary methodological common element is the intention to characterize changes in cognition or behavior across aging, which is a variable that cannot be experimentally controlled. Comparing across large age ranges also compares across cohort effects caused by broad environmental or social

changes over the lifespan age range. For example, comparing some aspects of health in participants who are currently in their 60s to participants in their 30s varies a substantial period of exposure to lead in the atmosphere (in the 1970s) in addition to age. Social, historical and cultural changes over time are extraneous variables all potentially confounded with age that provide challenges to drawing robust conclusions.

As an example of the difficulty in drawing causal inferences, it has long been observed that many more women than men are diagnosed with Alzheimer's disease. One hypothesis is that this emerges from some key genetic difference related to the X chromosome, such as hormonal differences across the lifespan. However, it is also the case that women currently tend to live longer than men and age is a very strong predictor of Alzheimer's disease in that the older you are the more likely you are to express symptoms. Current aged populations grew up and lived in a social context where women live longer than men, putting them potentially at greater risk for later life onset. This creates the usual problem with confounded variables in that we have two perfectly reasonably explanations for an observed relationship and no easy way to tell which is true. Additional scientific work aimed at a better understanding of the underlying neurobiology of Alzheimer's disease may eventually clarify why this difference emerges.

A more recent similar example of a more positive orientation is recent work in aging identifying a sub-population of older adults called **super-agers**. These are adults who at the age of 80 or more exhibit a high level cognitive functioning that would match participants under the age of 50. These individuals were originally identified as a series of case studies. More systematic research then identified larger numbers of these individuals who may have been previously overlooked in health research because their functioning is entirely intact. Ongoing research aiming to identify "their secret" for aging so successfully has not yet found any robust lifestyle, activity or genetic differences that account for their retained high level of cognitive ability. This work will necessarily be correlational at first but could potentially lead to some future intervention based research to improve aging

outcomes.

Key Takeaways

- Case studies are an observational technique with particularly interesting individuals that often illuminate new areas of science
- Neuropsychological studies of unusual patients have been very influential in developing theories of brain-behavior relationships
- Cognitive Neuroscience is a research area within psychology and neuroscience that uses research methodology to understand cognitive functioning in the brain based on patient studies and neuroimaging methods
- Developmental psychology relies on using traditional research methods together with the participant variable of age
- Comparisons across ages can be done as a cross-sectional design between-participants or as a longitudinal design within-participants over time
- Developmental methods with young children pose specific challenges for developing operational definitions that can be implemented with rigor and without bias
- Ethical research with children, who are a vulnerable population, requires additional oversight and care
- Life-span development examines age-related change in older adults, often through correlational research methods which have the usual challenges for drawing inferences from non-experimental designs

Exercises

TBA