Transforming Research Methods in Health Services Psychology: Applications for the Advocate ~ Practitioner ~ Scientist

Lynette H. Bikos, PhD, ABPP & Cirleen DeBlaere, PhD, Editors

Kiana Clay, Editorial Assistant

28 Jul 2022

Contents

BOOK COVER

This open education resource is available in the following formats, all available in the docs folder at the GitHub repository:

- Formatted as an html book via GitHub Pages available
- As a PDF
- As an ebook
- As a Word Document

All materials used in creating this OER are available at its GitHub repo.
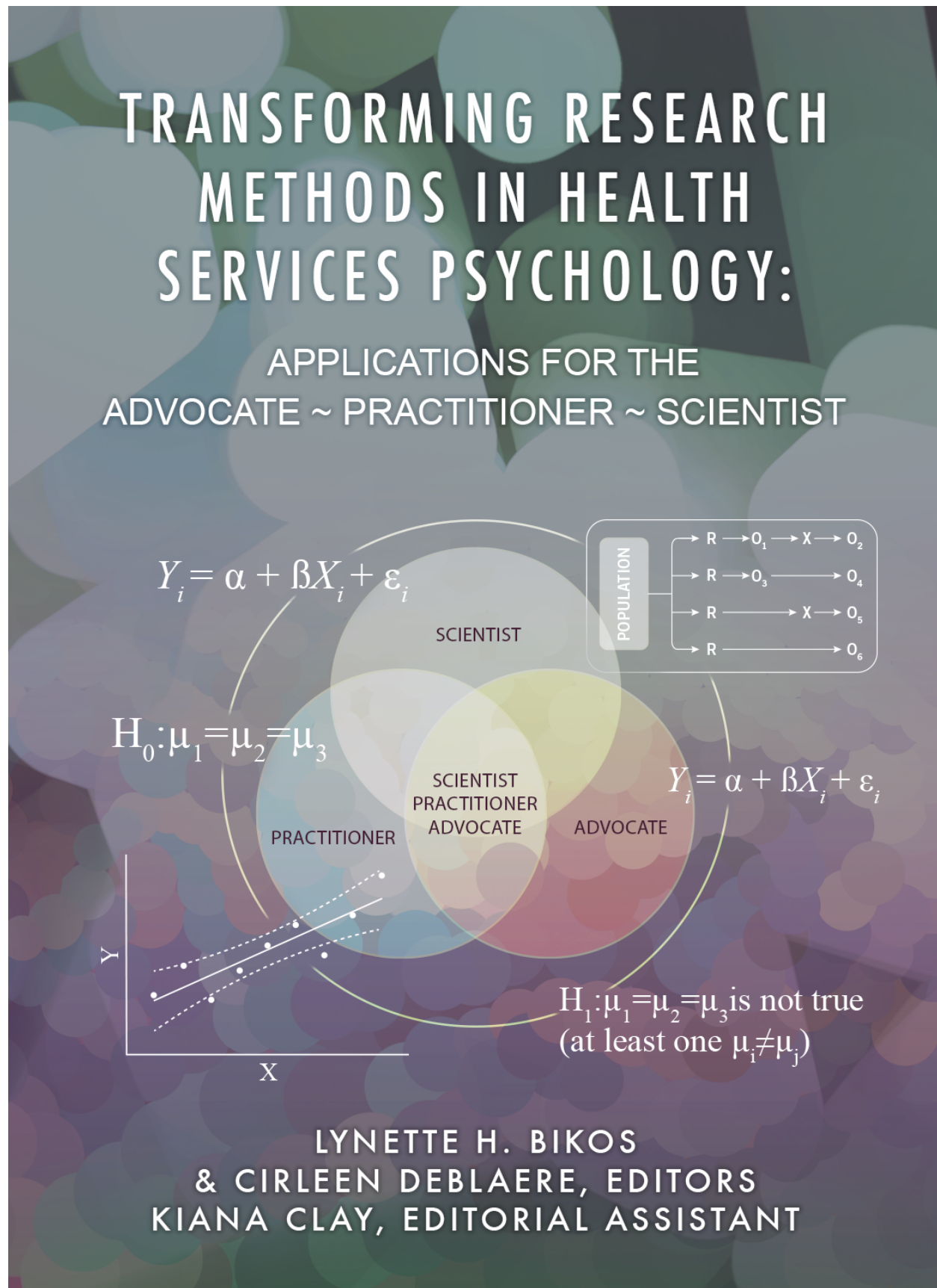
Figure 1: An image of the book cover. It includes three overlapping, pastel-colored, circles representing advocacy, practice, and science. These are surrounded by a handful of statistical symbols and formulae. methods.

PREFACE

If you are viewing this document, you should know that this is a book-in-progress. Early drafts are released for the Transforming Counseling Psychology Curriculum Showcase at APA 2022, for peer review, and for generating interest in collaboration. The document was last updated on 28 Jul 2022.

In her 2021-2022 term as President of the Society of Counseling Psychology, one of Dr. Amy Reynolds' Presidential Initiatives was *Transforming Counseling Psychology Curriculum and Praxis.* Dr. Reynolds invited counseling psychology faculty, practitioners, and doctoral students "to critically examine and deconstruct how various competencies, courses, and content are taught; how we socialize our students; and then re-imagine, dream, and reconstruct new and transformative ways to teach and train."

Strategies for a Social Responsivity

This open education resource (OER) is a product of the group devoted to *research.* There are a number of strategies we used to ensure that the OER moves in the direction of being socially and culturally responsive.

- Our authors committed to using the guidelines for a liberated syllabus found in the CCTC: Social Responsiveness in Health Services Psychology Education & Training Toolkit.
- We chose the format of OER because provides a zero-cost *textbook* to faculty and students.
- We sought authors and co-author teams that represent the diversity of health services psychology including discipline (counseling, clinical, educational),stage in career (students, early career professionals, mid- and late- career professionals), and identities that have been marginalized in higher education and our discipline.
- Each chapter works its way through an open peer review process where the chapter (with authors clearly identified)is hosted in a shared drive. At least two reviewers can mark up the same document and contribute to the same rubric. At any time the author(s) can see the review and, if desired, dialogue with the reviewers. At the outset, we specified the tone to be "formative not summative."
- Although we are still learning, we have attempted to use tools and techniques that are consistent with universal design. For example, we hope that the image captions and headers are marked such that text readers will identify them as such.

Perpetually in Progress

This book is being formatted in R Markdown, rendered into its "book" format with Bookdown, hosted on GitHub, and pushed to the internet (in its html format) through GitHub Pages. This set of tools allows the book to be *perpetually-in-progress.* This means that our authors can update their chapters at-any-time. It also means that we can add chapters at-any-time. If you are interested in contributing to the book, please contact us. It is one of our greatest hopes that this flexibility contributes to the socially and culturally responsive pedagogy that we intend.

Under Construction

At this stage in the OER's development, authors are still writing and revising chapters. The following designations will identify the chapters that have not been through the review process:

- *In-progress* means that the chapter is partially written (or perhaps outlined) and that the author(s) are continuing to work on the chapter.
- *Under review* means that the chapter is being (or has been) peer-reviewed.

Acknowledgements

Financial support, supporting the copy editing and desktop publishing for this project was provided by the Office of Education, Technology, & Media, Seattle Pacific University (Summer 2022).

The book cover was designed by Dominic Williamson, Senior Instructional Designer in Graphics & Illustrations, in the Office of Education, Technology, & Media at Seattle Pacific University.

Copyright with Open Access

This book is published under a a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. This means that this book can be reused, remixed, retained, revised and redistributed (including commercially) as long as appropriate credit is given to the authors. If you remix, or modify the original version of this open textbook, you must redistribute all versions of this open textbook under the same license - CC BY-SA.

A GitHub open-source repository contains all of the text and source code for the book, including data and images.

Chapter 1

Internal and External Validity in Health Service Psychology

Franco Dispenza, PhD & Alec Prince, MPA
Georgia State University

The focus of this lesson is to provide a review of internal validity, external validity, threats to validity, and current trends and considerations in relation to validity. Internal and external validity are foundational to experimental and quasi-experimental research. In experimental and quasi-experimental research designs, health service psychologists work thoughtfully and diligently to ensure that a line of systematic inquiry demonstrates some degree of internal and external validity. This is because psychologists and behavioral health researchers are concerned with making reasonable epistemological claims that could directly impact the lives of diverse communities and populations, especially in research studies attempting to show if particular interventions, treatments, or programs have a true effect on specific outcomes.

Whereas researchers utilizing qualitative frameworks may be more interested in methodological integrity (e.g., credibility and transferability; Levitt et al., 2017), researchers employing quantitative-based paradigms are especially interested in internal and external validity. You will notice that the term "validity" is used in many concepts and research frameworks, including construct validity, content validity, predictive validity, criterion validity, and statistical conclusion validity. These are all specific types of validity attempting to establish a degree of accuracy or truthfulness in research. Further, the aforementioned forms of validity often have statistical computations and procedures that accompany them (e.g., bivariate correlations, beta weights, etc.). This chapter will not address those forms of validity. Instead, we will focus on internal and external validity, conceptual constructs that rely on methodological procedures and considerations versus statistical calculations.

1.1 Learning Objectives

Learning objectives for this chapter include the following:

- Students will be able to explain the dimensional features of internal and external validity in experimental and quasi-experimental research.

- Students will be able to list and discuss threats associated with internal and external validity in experimental and quasi-experimental research.
- Students will be able to discuss and apply established methods for controlling the various threats associated with internal and external validity in experimental and quasi-experimental research.
- Students will be able to identify and apply their knowledge of internal validity, external validity, and their associated threats to current trends and issues in psychological research.
- Students will be able to critique and distinguish the strengths and limitations of internal and external validity when applied to socially responsive research.

1.2 Recommended Readings

The following served as critical references in the development of this chapter. We encourage you to review them.

- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi- experimental designs for research. Chicago, IL: Rand McNally.

This classic and foundational text introduces readers to internal validity, external validity, and threats to validity as originally proposed by Campbell and Stanley. It also provides readers with an overview of the various experimental and quasi-experimental methodological designs, and how various methodological designs could be used to minimize threats to validity.

- Ferguson L. (2004). External validity, generalizability, and knowledge utilization. Journal of Nursing Scholarship, 36(1), 16-22. https://doi.org/10.1111/j.1547-5069.2004.04006.x

Although aimed for a nursing audience, there is much to be gained from Ferguson's (2004) review of external validity, generalizability, and evidence-based practice. Ferguson reviews much of the major conceptual tenets of external validity, including its threats and control strategies. Ferguson also discusses ways in which researchers and practitioners could enhance external validity of research.

- Schmuckler, M.A. (2001).What Is Ecological Validity? A Dimensional Analysis. Infancy, 2(4), 419-436. https://doi.org/10.1207/S15327078IN0204_02

Schmuckler (2001) introduces readers to a subset of external validity, namely ecological validity. Schmuckler provides an historical review of ecological validity, and a discussion of the various dimensions, advantages, and criticisms of ecological validity.

1.3 Internal Validity

Health service psychologists are committed to accuracy, honesty, and truthfulness when engaging in the research process. Whether representing research findings fairly, capturing the essence of people's lived experiences with precision, or using research to advocate against harmful psychological practices, researchers are compelled to uphold the integrity of the knowledge claims they make

through their scholarship. Researchers are committed to ensuring that their research is justifiable, well-grounded, and internally valid.

Internal validity is the extent to which a researcher can infer cause-effect relationships between a set of variables, while simultaneously excluding the influence of confounding or extraneous variables (Campbell & Stanley, 1963; Cook & Campbell, 1979). In an attempt to establish internal validity, it is important to rule out the effects of confounding or extraneous variables. *Confounding or extraneous variables* can serve as potential rivals to inferred relationships, leading a researcher to be less confident about the conclusions and implications that can be made between an independent (or treatment) variable and a dependent (or outcome) variable.

Methodological control is also paramount in experimental and quasi-experimental research. In an ideal research scenario, a researcher will have identified and controlled for all possible confounds and extraneous variables that may be inherent in the methodological design of a study. Since determining causality is the goal in experimental and quasi-experimental research, such a scenario would render a study to be high in internal validity. However, we know in practice there is no such thing as a perfect study, especially when there exist numerous *threats to internal validity*. Threats consist of factors or conditions that endanger a researcher's capacity to amplify a study's level of internal validity (Onwuegbuzie & McLean, 2003).

1.3.1 Threats to Internal Validity

Researchers often execute considerable forethought when identifying and eliminating potential threats to internal validity. Psychological researchers are encouraged to use a variety of logic-informed heuristics, "what if" sensitivity analyses, or consider any new data or findings to rule out rival hypotheses (Kenny, 2019). Researchers are also encouraged to consider more classically identified threats to their research design (Campbell & Stanley, 1963). Some of the most common threats to internal validity consist of history, maturation, instrumentation, statistical regression, selection, attrition, and experimenter bias (Campbell & Stanley, 1963; Cook & Campbell, 1979; Shadish et al., 2002). Threats to internal validity can occur, or be discovered, at any point during the research process, including design, implementation, data collection, and analyses. There can exist multiple threats in a given study, and these threats can also intersect with one another. Keep in mind that threats are never something a researcher just "checks off" in a box, but rather a researcher continuously monitors their methods to ensure the credibility of the concluded findings (Onwuegbuzie & McLean, 2003). Each threat to internal validity is discussed in more detail below.

1.3.1.1 History

No researcher can control the potential of geological, sociopolitical, financial, cultural, climate-related, pandemic, and other national and global events from impacting individuals in a research study. But inevitably such events happen, and unfortunately with what seems to be at higher frequencies these days. Historical events can influence the manner in which participants respond to a study's dependent variable, making it difficult for researchers to determine whether an observed outcome was the result of a study's independent variable or the historical event. In some instances, the historical event could directly intersect with objectives of the study itself making it even more difficult for researchers to make valid conclusions. Imagine you are evaluating the effectiveness of a new stress-reduction intervention for survivors of catastrophic hurricanes in a low-income, rural community in the southeast United States. In the middle of implementing the intervention, a

tornado hits the same community, leading to severe damage and loss of life. What impact will this historical event have on the research study?

1.3.1.2 Maturation

Development and growth are natural processes that occur for individuals during their lifespan. These processes could be physiological, cognitive, social, or emotional. Given the length of a particular study, sometimes results from a study may be more indicative of naturally occurring growth and developmental factors, versus any manipulated independent variable.

1.3.1.3 Instrumentation

No tool, test, or measure used in research can be entirely valid and reliable one hundred percent of the time. Instruments can produce low reliability scores or even produce inadequate psychometric properties (e.g., construct and predictive validity) with a study's particular sample. This is an important consideration as some tests and measures were never validated or normed with diverse samples. Consider a measure of marital satisfaction that was created using heterosexual couples in the Netherlands. The measure may have great face validity, and even demonstrate adequate levels of content, criterion, and construct validity. But it may not produce the same psychometrics if used in married couples in the United States. It may be entirely inappropriate to use this measure if the United States sample also includes same-sex couples. Lastly, human beings administering instruments could contribute additional error, making instrumentation a serious threat to research studies. For example, some researchers may become fatigued when administering some measures, may not be consistent with measurement, or may not accurately observe a phenomenon during a research study.

1.3.1.4 Testing

It is common to test participants multiple times throughout the course of a research study. Researchers often use the same test, measure, scale, or inventory when testing participants, but it begs the question as to whether changes in a test—that had been given multiple times—reflect true change? Sometimes referred to either as a *practice effect* or *fatigue effect*, changes in scores on the same test may be the result of familiarity with the test or becoming tired with the test.

1.3.1.5 Statistical Regression, or Regression to the Mean

Sometimes researchers select participants based on high or low scores on a particular test or measure (Campbell & Kenny, 1999). For example, researchers may be interested in examining people who score high on measures of academic or cognitive functioning. If retested, those same participants may continue to score high or low. However, not all would score as high or low because of statistical regression, or regression to the mean. When participants are retested, researchers find that scores are less extreme and center toward an average score. Imagine a researcher was interested in testing a career counseling intervention with first generation college students who reported high scores on career indecision and anxiety in a career battery of questionnaires. After the intervention is complete, the researchers issue a final battery of questionnaires. These same first generation

college students may appear less anxious and indecisive due to statistical regression to the mean and not necessarily the intervention.

1.3.1.6 Attrition, or Differential Mortality

Ethically, participants have the right to withdraw from participating in a research study at any given time during the research process. When participants withdraw, or drop from a study, researchers refer to this as attrition. A major concern of attrition is that it leads to potential biases in scores between groups (or observations in the case of longitudinal studies) that may not reflect whether an independent variable had any effect on the dependent variable. If there is substantial attrition in one group, or across the entire study, it leads a researcher to question why participants are withdrawing. Researchers may further wonder if participants dropping from the study are characteristically different from those who are remaining in the study. Attrition limits the type of conclusions that can be made because observations made across time or between groups may not reflect true differences as a function of the independent variable. Rather, there may be some other underlying issue with the research study.

Imagine a researcher has been evaluating the effects of a six-session mindfulness cognitive behavioral group intervention for the reduction of race related stress among Black and African American women employed in a large healthcare setting. The researcher used a longitudinal, between group experimental design with an intervention group (treatment) and a control group (education), but found that attrition rates were higher for those in the control than for the intervention group. Equivalent and adequate comparisons could not be reliably made between the two groups at completion of study, so the researcher decided to send surveys to those dropped from the study and inquire why they dropped. The surveys return and the researcher finds that a significant portion of the control group were made up of on-call, ambulatory nurses who were unable to commit to the scheduled sessions. Therefore, the attrition was the result of some characteristic that differentiated the two groups.

1.3.1.7 Experimenter Bias

Researchers need to ensure they do not engage in verbal or nonverbal behaviors that inadvertently alter the results of a study. Sometimes this is referred to as an *experimenter expectancy*, and it is even known as the *Rosenthal Effect*. This becomes a threat when a participant's response in a study is the result of the experimenter's expectations versus the manipulated or independent variable. Examples include emphasizing particular words when reading prompts or scripts, or excessive nodding and smiling when certain favorable responses are solicited by study participants.

1.3.1.8 Selection Bias

Differences between groups in experimental studies can sometimes be the result of characteristics of the participants themselves, versus the manipulated or independent variable. For instance, a researcher may have accidentally grouped people along the same characteristic, such as sex, gender, sexual orientation, race, ethnicity, or age group. This potentially sets up nonequivalent groups in experimental or quasi-experimental research, and makes it difficult for researchers to determine whether any changes in dependent or outcome variables were the result of an independent variable

or the characteristic itself. This also pertains to self-selection bias commonly seen in survey and questionnaire research, in which participants self-select to participate in a study.

1.3.2 Controlling for Threats to Internal Validity

Researchers have identified a number of methodological procedures that could be used to control for threats to internal validity (Campbell & Stanley, 1963; Cook & Campbell, 1979; Fabrigar et al., 2020; Onwuegbuzie & McLean, 2003; Shadish et al., 2002). Below we discuss the importance of considering control groups, random assignment, matching, blocking and holding variables constant.

1.3.2.1 Control Groups

Control groups are commonly used in experimental and quasi-experimental human subjects research, and there is an important logic to its use. Individuals are assigned to a group (or condition) in which they do not receive the treatment or manipulated variable. However, participants do partake in similar tasks and conditions (e.g., complete surveys, questionnaires, physiological markers, etc.) as those in the experimental condition. Upon completion of the study, researchers then compare how the intervention or experimental condition performed alongside the control condition or group. Control groups are particularly effective at controlling for the effects of history, maturation, testing, instrumentation, and regression toward the mean (Campbell & Stanley, 1963; Cook & Campbell, 1979).

1.3.2.2 Random Assignment

Considered one of the most robust methods to control against threats to internal validity, researchers randomly place volunteer participants in various study conditions at the very beginning of a research study. This assures the researcher that each participant had the same or equally probable chance of being placed in either an experimental condition (or group) or a control condition (or group), while helping to decrease any unknown or intentional influence on assignment of participants to different groups. It further assures the researcher that the groups are equitable in terms of various characteristics of the participants (e.g., demographics, temperament, etc.; Fabrigar et al., 2020). Any observed differences or changes seen among the groups or conditions could then be accounted for by the manipulated independent variable or the applied intervention. More importantly, any observed difference between the groups is not the result of any sort of systematic bias that might have occurred during the initial phases of the research study.

Take a hypothetical scenario in which a researcher is evaluating the ways in which implicit sexist messaging influences women's responses to a cognitive motor task. Women are recruited from the community to participate in the study. One group receives the implicit sexist messaging while the other group does not. However, nearly all members in one group happened to be women between the ages of 21 and 29, while the majority of those in another group were women between the age of 43 and 56 This could constitute a systematic bias since there are generational differences between members in one group versus the other. Random assignment is incredibly important when controlling for the effects of selection (Campbell & Stanley, 1963; Cook & Campbell, 1979; Shadish et al., 2002).

It is important to keep in mind that random assignment and random sampling are not synonymous with one another. **Random sampling** is when researchers utilize a variety of probable sampling

techniques (e.g., simple, systematic, stratified, or cluster) to recruit participants who approximate the general population. It also means that all members of a given population have an equal chance of being recruited to participate.

1.3.2.3 Matching

To further avoid the threat of selection bias, researchers may engage in the process of participant matching. This is especially helpful if a researcher cannot guarantee equivalent groups through randomization, or when sample size may be too small. Participants are matched on a variety of characteristics (e.g., cognitive or intelligence pre-test scores, gender, age, etc.), by placing members of similar characteristics in either an experimental/intervention condition or in the control condition. This helps a researcher establish some degree of equivalence between groups within a particular sample.

1.3.2.4 Blocking and Holding Variables Constant

Researchers may also choose to use some characteristic in the study's sample (e.g., cognitive or intelligence scores, ethnicity or race) as an additional independent variable. This is referred to as *blocking*. Unlike matching, researchers may employ this strategy to see if a particular characteristic of the sample has an effect on the dependent variable. Alternatively, some researchers may choose to hold a particular characteristic constant or homogenize some sample characteristic so as to eliminate any undue influence from extraneous characteristics of a sample. For instance, a researcher may choose to only recruit high school aged adolescent boys between the ages of 14 and 15 who score above a certain threshold on an anxiety and depression measure to participate in a short-term emotional regulation treatment study.

1.4 External Validity

Many researchers invest time and resources with the hopes of expanding their research findings to larger communities and contexts, especially if the research is aimed at alleviating any suffering or influence larger systemic change efforts. Thus, researchers are not only concerned with internal validity, but also the external validity of their study. **External validity** is the degree towhich research findings can be generalized to the population that approximate the original context of the study (Campbell & Stanley, 1963; Cook & Campbell, 1979; Shadish et al., 2002).

External validity is also concerned with the degree in which a study can be generalized across broader populations, treatments, settings, and conditions (Ferguson, 2004). Researchers wish to move beyond the controlled setting in which the study originally took place, and further consider ways that the findings may apply in other diverse applicational settings, time, persons, or slightly different variables or targets (Kenny, 2019; Shadish et al., 2002). For instance, a researcher who evaluated the effectiveness of a minority stress reduction intervention for transgender and gender nonbinary individuals in a controlled laboratory setting at a university, might have interest in seeing the intervention applied or replicated with transgender and gender nonbinary individuals in community based clinics, private practices, college counseling centers, and other agencies across the United States. That same researcher may have further interest in having their research used

to inform policy on affirming psychological care for transgender and gender nonbinary individuals for all mental health practitioners.

Researchers always have the hope that the finding of their research will have some degree of relevance and importance in "real-world" settings, particularly if study findings are replicated in other contexts. Replicated studies that support original study findings are considered to demonstrate high levels of external validity, as well as other forms of validity (e.g., internal validity, construct validity; Fabrigar et al., 2020). As a variation, or subcategory of external validity, **ecological validity** is concerned with whether a study's results can be applied to naturalistic or representative settings in every-day life (Andrade, 2018; Schmuckler, 2001). With this in mind, it is important to consider that there also exist threats that could interfere with a researcher's confidence in the external validity of their study's findings.

1.4.1 Threats to External Validity

Campbell and Stanley (1963) identified four particular threats, including reactive or interaction effects of testing, interaction effects of selection bias, reactive effects of arrangement, and multiple treatment interference.

1.4.1.1 Reactive or Interaction Effect of Testing

Sometimes researchers need to be cognizant that research studies, experiments, and testing procedures—in and of themselves—may be the catalyst producing some of the findings we see from research. In many "real-world" conditions, people are not tested or observed as much as they are in research. In particular, Campbell and Stanley (1963) discussed how exposure to a pre-test condition, or multiple testing conditions, may influence a study participant's degree of sensitivity to the experimental variable. Consider an example in which a researcher is interested in examining a clinical supervisor's attitudes toward racial and ethnic microaggressions in counseling. The clinical supervisor is asked to view a fictitious counseling session of a supervisee, and then asked to identify any subtle instances of discrimination from a 10 minute clip of a counseling session. Afterwards, the researcher follows up with another post-test to see if there have been any changes in attitudes toward racial and ethnic microaggressions in counseling. The potential threat to external validity in this example is that the clinical supervisors in the study have been sensitized by the pre-test condition (i.e., the fictitious counseling session), increasing their potential chances of identifying microaggressions in a counseling session. If generalized out, clinical supervisors may not respond the same way since they've not been pre-tested and sensitized.

1.4.1.2 Interaction Effects of Selection Bias

Health service psychologists pay close attention to samples, and work diligently to recruit adequate samples to participate. However, there are situations in which a particular sample in a research study would not generalize to the entire population. This could be the result of selection bias. For instance, many university researchers utilize an undergraduate psychology research pool to recruit participants for their research. But undergraduate students represent a biased sample, and do not reflect the larger population in terms of representative demographics. Thus, researchers replicating a study with a different sample may not obtain the same findings.

1.4.1.3 Reactive Effects of Experimental Arrangements

Research conducted in highly controlled settings (e.g., sterile laboratories) run the risk of not generalizing well in "real world" diverse settings or populations. This is mainly to do with the fact that research participants are willing volunteers who understand they are fully participating in experimental or study-related activities. In some instances, research participants may respond or behave a certain way because they are being observed. Frey (2018) reports that a participant may even have the desire to please a researcher by altering their performance on a particular outcome. This may sound familiar because you may understand this to be the **Hawthorne effect**, a phenomenon in which human beings change their behavior as a result of being observed.

1.4.1.4 Multiple Treatment Interference

Depending on the sequencing of a particular study, researchers may provide the same subject different treatments or interventions at different intervals. For example, a researcher may be testing multiple formats to examine the combination of psychotropic medication along with some type of psychotherapy. However, this makes it difficult for researchers to determine if the sequencing of the differing treatments played any role in any of the observed outcomes. Because of this type of sequencing, we would argue that there has been some level of treatment contamination, because it is difficult to control effects from previous treatments or studies.

1.4.2 Controlling for Threats to External Validity

It is important to note that external validity can never be assured, even when a researcher stringently addresses and controls for threats to internal validity (Ferguson, 2004). However, there are some threats to external validity that can be managed through some methodological considerations. Below we discuss only a few, including random selection, concealed research, as well as counterbalancing and strategies to control for pre-testing effects. #### Random Selection Sometimes confused with random assignment (already discussed above as a means of controlling threats to internal validity), random selection is about accessing and including a representative sample of the target population in a research study. Random sampling procedures, such as simple or stratified sampling, play a significant role when it comes to random selection. You may recall that random sampling is concerned with the notion that every individual (or observation) has an equal probability of being selected for a study. The equal probability of being selected then increases the probable chances that research findings can be generalized back to the target population (Ferguson, 2004). This form of control is particularly beneficial when considering the interaction effects of selection bias.

1.4.2.1 Single, Double, and Triple-Concealed Research

It is important to preface that the term often used in research texts is single, double, and triple "blind" research. However, we believe "blind" is often misused in a variety of contexts, and in this context we believe it perpetuates ableist ideologies. And so, we offer a slight modification by using the term "concealed." In order to reduce overt and covert forms of researcher bias in a study, researchers attempt to conceal as much as possible from participants and other members of a research team. In a single-concealed research study, only the researcher knows if participants

are in a control or experimental group. Participants do not know what condition they are in. In a double-concealed research study, neither the researcher or participant know which is the control or experimental group. In a triple-concealed research study, consistent with the double-concealed design, neither the researcher nor participant know if they are in the control or experimental group. Additionally, those responsible for analyzing or examining outcomes do not know which set of variables were the control or experimental condition. This form of control is especially helpful when addressing reactive effects of experimental arrangements.

1.4.2.2 Counterbalancing and Controlling for Pre-Test Effects

In order to address the effects of multiple treatment interference, or carryover effects, a researcher may consider the use of counterbalancing. A researcher must decide *a priori* all the possible sequences for a treatment, implement those varying permutations, and evaluate study participants in those different orders in order minimize carryover effects. If pre-testing effects is a concern, a researcher may decide not to include a pre-test at all, and compare groups at post-test only. Relatedly, a researcher may want to consider the use of the *Solomon Four-Group Design* as a means of countering the effects of a pre-test (Allen, 2017). In a Solomon Four-Group Design, a researcher will have four groups, in which some groups receive a pre-test and other groups do not.

1.5 Current Issues, Trends, and Considerations

We will briefly review current issues, trends, and consideration of internal and external validity. It is vital that health service psychology researchers attend to matters of multiculturalism and diversity, since this has direct implications on the external and ecological validity of study findings. It is also incredibly important to consider evidence based practices with culturally diverse populations. Replication of research is another pressing trend in the field of psychology that requires researchers to consider how they present and disseminate their research findings to broader communities. Relatedly, Internet-based collection procedures require new concerted efforts in our conceptualization of internal control and generalizability of results. We end with a recommendation that researchers could consider when addressing internal and external validity of research.

1.5.1 Validity Issues Related to Cultural Diversity

In a critique of psychological research, Sue (1999) wrote that researchers have overemphasized internal validity over external validity. As a result, the overemphasis on internal validity has hindered the development of research on ethnic and racial minority groups, as well as other marginalized groups. Unfortunately, this has further perpetuated psychology's own history of reinforcing oppression and inequality (Lewis, 2021). This has also led to a tension between "basic" researchers who privilege statistical power and high degrees of internal validity and "applied" researchers who privilege the nuance of intersectional research and external or ecological validity (Lewis, 2021).

As an issue of validity, psychological based research has often failed to include diverse participants, or it has failed to report on the diverse identities within samples. A 36-year review of randomized clinical trials of depression (from 1981-2016) found that less than half of the studies reported on the sample's race/ethnicity, about one in six trials had a predominantly ethnic minority sample, and one in seven studies had a predominantly low socioeconomic (SES) sample (Polo et al., 2019).

Similarly, researchers have found inconsistency in demographic data collection and reporting. Racial and ethnic minority groups are often underrepresented, along with disability, and diverse sexual orientations (Greenwell & Hough, 2008). Within research on the lesbian, gay, bisexual, transgender, queer (LGBTQ+) communities, the needs of cisgender white gay men have been privileged, and the experiences of LGBTQ+ individuals who represent women, people of color, transgender, or bisexual communities have largely been omitted or excessively medicalized (American Psychological Association, 2015; American Psychological Association, 2021; Hegarty & Rutherford, 2019).

Researchers have noted the importance of consistently collecting and reporting demographic data on race, ethnicity, sexual orientation, gender identity, SES, disability, and other identities (dickey, Hendricks, & Bockting, 2016; Greenwell & Hough, 2008; Polo et al., 2019). For instance, dickey and colleagues (2016) stressed the importance of collecting and analyzing gender identity and sexual orientation data separately in population surveys, especially since gender identity is often conflated with sexual orientation. Parent and colleagues (2013) encourage researchers to focus on the context of intersecting oppressions in addition to intersecting identities, to ensure that diverse populations are included in psychological research.

To redress the exclusion of marginalized identities from psychological research and scholarship, researchers recommend going beyond including marginalized individuals in samples to including them in creating the studies themselves. Participatory research is an umbrella term for research that engages those being studied in the production of knowledge to promote education and change (Cargo & Mercer, 2008). Participatory research in its many forms has been implemented with gender and sexual minority populations, refugees, individuals with disabilities, and racial and ethnic minority communities around the globe (Cargo & Mercer, 2008; Fine et al., 2021; Jacquez et al., 2021). Fine et al. (2021) argue that "the move to include and privilege those most impacted by injustice as co-researchers is not simply an act of empathy or decolonizing; it is a commitment to good science" (p.346).

1.5.2 Validity Issues Related to Replication Studies

In recent years, psychologists have argued that the discipline of psychology suffers from a replication crisis (Fabrigar et al., 2020). This calls into question the validity of psychological research and the findings that have been reported over the years. Some notable cases of outright fraud, questionable research practices (Pashler & Wagenmakers, 2012), flawed methodologies, and incorrect analysis of data (Fabrigar et al., 2020) have led to some of these replication issues. Others have argued that failures to replicate result from small sample sizes, subsequent low statistical power (Maxwell, Lau, & Howard, 2015; Schmidt & Oh, 2016), as well poor statistical conclusion validity (Fabrigar et al., 2020). Schmidt and Oh (2016) argue that "the real problem is not a lack of replication; it is the distortion of our research literatures caused by publication bias and questionable research practices" (p.32).

Addressing the replication crisis is no easy task as it requires addressing journal review processes, research practices, and reward structures in academia (Pashler & Wagenmakers, 2012). The Reproducibility Project, created by the Open Science Collaboration, aimed to address the replication crisis by replicating 100 experimental and correlational studies from key psychology journals (Open Science Collaboration, 2015). The Reproducibility Project found that while 97% of all the original studies had significant results, only 36% of replicated studies demonstrated significant results (Open Science Collaboration, 2015). Schmidt and Oh (2016) note the importance of replicating

studies with nonsignificant results and recommend meta-analysis as a solution to this issue, provided publishing bias and questionable research practices are addressed.

Some researchers argue that simply replicating studies is not enough to safeguard the validity of psychological research. In an analysis of the Reproducibility Project, Sabik and colleagues (2021) found that the studies reproduced by the Project seldom considered context and identity, even when it was central to the study's design, and that study samples were predominantly WEIRD (people from Western, educated, industrialized, rich, and democratic countries). Further, intersectionality, power, discrimination, and historical contexts were rarely considered in the Project's reports (Sabik et al., 2021). Sabik and colleagues argued that the Reproducibility Project and the discourse surrounding the replication crisis are more concerned with data transparency and methods than with the inclusion of historically oppressed and marginalized groups. To truly move the discipline forward, some argue it is necessary to set aside the emphasis on traditional research methods and reproducibility in favor of methods that center on the co-creation of knowledge (Grzanka & Cole, 2021).

1.5.3 Internet Research and the use of Crowdsourcing Platforms

Many social science researchers utilize the Internet (e.g., social media, list servs, emails) for purposes of recruitment and data collection (e.g., Survey Monkey, Qualtrics, etc.). Although this has considerable implications for internal and external validity that go beyond the scope of this chapter, it is essential that we discuss some of the implications of crowdsourcing platforms. Crowdsourcing platforms are online websites that can be used by researchers to recruit potential participants who have access to Internet and an electronic device (Peer et al., 2017). And although we cannot review all of the available crowdsourcing platforms, researchers have several platforms to choose from, including CrowdFlower and Prolific Academic (see Peer et al., 2017 for a review of the various strengths and limitations of these crowdsourcing platforms). However, Amazon's Mechanical Turk (MTurk) has garnered some of the most attention in recent years by methodologists and scholars. It is likely that many of the issues surrounding internal and external validity with MTurk are also applicable to other crowdsourcing platforms, but we will focus mostly on MTurk for this chapter.

Created in 2005 by Amazon, MTurk is an online marketplace where workers (Turkers) complete Human Intelligence Tasks (HITs) for MTurk requesters for pay. This is equivalent to a psychology undergraduate pool, but using a larger sample of the population. Typical HITs are transcribing movies, copying text from images, and participating in surveys. In addition to regular MTurk workers, there are MTurk "masters" whose accuracy are validated by previous MTurk requesters. MTurk is used extensively by businesses, academic researchers, and nonprofits (Pew Research Center, 2016). A review of key journals in psychology, psychiatry, and other social sciences found that fewer than 50 studies using MTurk data were published in 2011; in 2015, over 500 studies using MTurk data were published (Chandler & Shapiro, 2016).

Although MTurk is a cost-effective, efficient method to collect large amounts of data, there are concerns about the reliability and validity of MTurk data. For instance, several researchers have called the external validity of MTurk data into question. Turkers are relatively young and well-educated compared to national averages (Hitlin, 2016; Walters et al., 2018). Walters et al. (2018) found that MTurk workers' health status and behaviors were not comparable to a nationally representative sample. Compared to a national sample, MTurk users were over twice as likely to screen positive for depression, but they were less likely to exercise, smoke, have asthma, or have health insurance (Walters et al., 2018). Another concern regarding MTurk data is the overall decrease in data quality

resulting from an influx of computer programs or "bots" that complete HITs and individual users bypassing location restrictions using server farms ("farmers"). Chmielewski and Kucker (2019) conducted the same study over four years and found a substantial increase in low-quality data from MTurkers, including failures to replicate well-established findings, decreases in the reliability and validity of the Big Five Inventory, a widely used personality measure, and increases in participants failing response validity indicators (Chmielewski & Kucker, 2019).

MTurk remains a valuable resource for collecting data, provided the necessary steps are taken to ensure data quality. Buhrmester et al. (2018) recommend that researchers take the time to work on MTurk themselves to understand the Turker experience. Other recommendations include screening responses before approving HITs, including validity indicators, and comprehensive reporting on screening and study designs (Buhrmester et al., 2018; Cheung et al., 2017; Chmielewski & Kucker, 2019; Mason & Suri, 2011). MTurk is particularly useful for researching hard-to-reach populations such as individuals with disabilities, LGBTQ+ individuals, and those with low socioeconomic status (Smith et al., 2015). To safeguard against Turkers lying about being part of the target group, Smith et al. (2015) recommend providing monetary incentives that are not overly attractive and asking participants to self-identify prior to sharing the purpose of the research. Walters at al. (2018) suggested that researchers would benefit from using MTurk workers over masters because the two groups were comparable in demographics and health characteristics; however, workers are a larger sample and more cost-effective. With adequate measures to ensure data quality, MTurk remains an efficient and cost-effective option for researchers, especially those studying hard-to-reach populations.

1.6 A Consideration for Practice

First introduced in 1999, the RE-AIM framework is a tool that can help researchers balance internal and external validity when planning, designing, and evaluating health-related interventions (Dzewaltowski et al., 2004). Originally intended as guidelines for reporting research results, the framework is now also used to organize literature reviews and to translate research into practice. This has incredible implications for both internal and external validity, as it attempts to take a study beyond just epistemology and into direct practice with populations.

The RE-AIM framework's five dimensions include: *Reach, Efficacy/Effectiveness, Adoption, Implementation, and Maintenance* (Dzewaltowski et al., 2004; Glasgow, Vogt, & Boles, 1999). Reach and Efficacy/Effectiveness are both individual-level dimensions. Reach considers the percentage of the population of interest included in the intervention and how representative they are, whereas Efficacy/Effectiveness considers the impacts (both positive and negative) on participants (Dzewaltowski et al., 2004; Glasgow, Vogt, & Boles, 1999). Adoption and Implementation are organizational-level dimensions that consider the type and proportions of settings that will adopt the intervention and the extent to which the intervention is implemented faithfully in the real world (Dzewaltowski et al., 2004; Glasgow, Vogt, & Boles, 1999). Finally, Maintenance examines the continuity of the program over time at both the individual and the organizational levels (Glasgow, Vogt, & Boles, 1999).

RE-AIM is used in a variety of fields and settings, such as chronic illness management, mental health, smoking cessation, health policy, and diabetes prevention (Kwan et al., 2019). Although the RE-AIM framework is becoming more widely used, a systematic review noted that the framework is often used inconsistently (Gaglio, Shoup, & Glasgow, 2013; Glasgow et al., 2019). Several

adaptations and clarifications have been offered to mitigate confusion and inconsistency using the RE-AIM framework. Holtrop and colleagues (2018) offered guidance on integrating qualitative methods into the RE-AIM framework. Holtrop et al. (2021) offered clarifications on common misconceptions about the framework. The Practical, Robust, Implementation, and Sustainability (PRISM) model is an emerging complement to the RE-AIM framework that focuses on contextual factors (Glasgow et al., 2019). With the original goals of producing valid and relevant research and translating research into practice, RE-AIM and PRISM will continue to grow and evolve as researchers apply these frameworks to new populations and settings. Researchers and students can directly go to the website to learn more about how to implement these principles in their research, as well as access various resources, tools, and checklists (http://www.re-aim.org).

1.7 Activity

Take a moment to locate the latest issue of a peer-refereed journal in your respective field. Some example psychology journals published by the American Psychological Association include the *Journal of Counseling Psychology, School Psychology, Journal of Consulting and Clinical Psychology, Health Psychology, Developmental Psychology, or Professional Psychology: Research and Practice.* Once you have located a recent issue, browse through the table of contents and select a quantitative article that may be of interest to you. Read the article and then consider the following prompts:

1. Which threats of internal validity were identified and controlled for in the study? Were any explicitly identified and addressed by the authors the article? Were there any that you noticed that were not addressed or controlled for in the study?
2. How were study participants recruited or sampled for the study? In what ways were study participants diverse?
3. What limitations were mentioned in the discussion section? Were issues of internal and external validity explicitly named? If so, which ones? If not, which internal and external validity issues were implied?
4. How would you replicate the study you read? What additional validity factors would you consider to improve the new proposed study's internal and external validity?

1.8 References

Chapter 2

Open Science as a Step Toward Social Responsivity in Research

Lynette H. Bikos (she/her) & Jamie Layton (she/her)
Seattle Pacific University
Seattle Pacific University was built on the unceded ancestral lands of the Duwamish people

In 2015, the Open Science Collaboration demonstrated that psychological science suffers from a problem of replicability. This has turned the field of psychology's attention toward the benefits that open science can provide. Engaging in practices such as preregistering research studies, sharing data, open peer review, open access, and utilizing open education resources can assist with issues such as transparency and access and, in turn, could lead to greater social and cultural responsivity in research.

2.1 Learning Objectives

Learning objectives for this chapter include the following:

- Distinguish reproducibility from replicability.
- Identify mechanisms for increasing transparency in the research process.
- Describe potential benefits of preregistering research studies.
- List elements of data sharing that would improve reproducibility.
- Obtain an ORCID persistent digital identifier to help track your scholarly record.

2.2 Recommended Readings

The following served as critical references in the development of this chapter. We encourage you to review them.

- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), 943–943. https://doi.org/10.1126/science.aac4716

- – This was the study that called attention to the failures of replicability in psychological science. Given its historic and activating role in promoting open science in psychology, it is well worth reviewing the original document.

- Bosnjak, M., Fiebach, C. J., Mellor, D., Mueller, S., O'Connor, D. B., Oswald, F. L., & Sokol, R. I. (2021). A template for preregistration of quantitative research in psychology: Report of the joint psychological societies preregistration task force. *American Psychologist, 77*(4), 602-615. https://doi.org/10.1037/amp0000879

  - – This document creates a detailed argument for why preregistrations are necessary and it provides instructions for completing a preregistration for quantitative research.

2.3   Defining "Open Science"

Broadly defined, *open science* is the movement to make scientific results and processes accessible and reusable by all. As such it involves dimensions of both principles or culture and practices or technology. Although Steven Mann has been credited with coining the phrase "open science" in 1998 when he registered the domain names openscience.com and openscience.org, an earlier use the phrase was found in 1985, by Daryl E. Chubin in the article, "Open Science and Closed Science: Tradeoffs in a Democracy" [noa, 2022d]. Curiously, Chubin opened their article by referencing framework that is commonly used in introductory research methods textbooks [e.g., Krathwohl, 2009]. That is: Merton's norms of science.

Merton [1942] argued that *communism* (i.e., common owership of information), *universalism* (i.e., universal standards for claims of knowledge), *disinterestedness* (i.e., integrity in gathering and interpreting data), and *organized skepticism* (i.e., critical review of scientific reports) were necessary for producing objective knowledge. Separated by more than four decades, Merton and Chubin [1985] discussed the role of science in a democracy. While further elaboration of their arguments and conclusions extend beyond the goals of this chapter, the norms of communism and organized skepticism are foundational to the current understanding of open science.

Given the expansiveness of the notion of open science and its connectedness to other open initiatives (e.g., education, government, advocacy) numerous conceptual models may apply. Fecher and Friesike's [2013] model organizes open science into five schools of thought. The *public* school of thought advocates that science needs to be accessible to a wider audience. Themes within this school include accessibility to the production of research (e.g., the citizen scientist) and consumability/accessibility to research results. The *democratic* school of thought is concerned with access to knowledge – especially when it is publicly funded. Subthemes within the democratic school emphasize that data be open for forseen and unforseen re-use and that there be open access to research publications. This is particularly true for government-funded research (e.g., the citizen should not be twice-taxed to access research findings). The *pragmatic* school of thought is concerned with efficiency, specifically utilizing online tools to allow for greater collaboration. The *infrastructure* school is specifically concerned with the technical infrastructure including software tools, applications, and computing networks. Finally, the *measurement* school of thought considers alternative standards to evaluating scientific impact of findings. For example this school challenges the "impact factor" (i.e., a measure of the average number of citations to an article in a journal), and suggests also counting (and publishing) elements like peer review and dissemination via emerging publishing formats.

In light of these foundational inputs, we will approach open science with a simultaneous review of its principles and practices in the three broad areas of (a) transparency of the research process; (b) access to research findings; and (c) tools for open science.

2.4 Transparency of the Research Process

Replicability (re-performing an experiment and collecting new data) and *reproducibility* (re-performing the same analysis with the same code using a different analyst; Patil et al. [2016]) are foundational to science. Yet investigations across scientific disciplines have demonstrated significant failures to both. Through its large scale, collaborative effort, the Reproducibility Project [Open Science Collaboration, 2015] revealed that psychological research often fails to replicate prior research. The collaborative selected 100 studies (97% of which had significant results) from three leading journals. Research teams committed to conducting high-powered, high-fidelity replications. Results were compared to the original studies with five metrics that assessed various dimensions of the analyses. Only 36% of the replications yielded statistically significant results; similarly a subjective analysis concurred that only 38% of the effects replicated the original results. Regarding effect sizes, 47% of the original effects were in the 95% confidence interval of the replication effect size. Finally, when original and replication results were combined (e.g., a meta-anlytic combination), 70% of the effects were statistically significant.

When studies are replicated, there are many reasons that the outcomes may differ. There may have been small-to-large differences in design and methods that are impactful enough to change the result. There could be a Type I (false positive) or Type II (false negative) error. Or there could be confirmation biases at any (or multiple) stage(s) of the design [Stevens, 2017]. It is these biases that are particularly problematic. Bias at the individual-study-level occurs when the researcher reports, out of the many possible analyses, the one(s) that provide the most consistent or significant results [Hengartner, 2018, van't Veer and Giner-Sorolla, 2016].

Bias occurs at the journal level when reviewers and editors favor significant findings over non-significant ones [Stevens, 2017]. Driessen et al. [2015] has suggested that there is a 25% reduction in the estimated effect of psychotherapy because 24% of all NIH-funded trials aimed at evaluating the efficacy of psychological treatment for major depressive disorder were never published. Dubbed the "file drawer problem" by Rosenthal [1979], this problem is compounded by researchers who are skeptical about the possibility of finding null results and the tendency of journals to not publish them.

Across scientific disciplines, transparency has been suggested as a potential remedy [Open Science Collaboration, 2015, van't Veer and Giner-Sorolla, 2016]. Stated another way, for research to be truly reproducible, the entire process must be open to scrutiny [Stevens, 2017]. We review three practices that would lead to more transparent, open, science. These include preregistration, data sharing, and open peer review.

2.4.1 Preregistration

Preregistration of a study involves specifying, in advance, the research questions, variables, research design and planned analyses [Stevens, 2017]. There are a number of websites for such preregistrations; a common one for psychological scientists is the Open Science Framework. Researchers are encouraged to preregister studies that involve the testing of a priori hypotheses and models as

well as studies that are intended as exploratory [Bosnjak et al., 2021]. Although studies can be preregistered at any time, posting the preregistration in advance of seeing (or even more strictly, in advance of collecting) the data provides a mechanism for reviewers and consumers to evaluate the degree of consistency with which the research design and planned analyses were followed [Bosnjak et al., 2021, Haven et al., 2020].

Scholars have noted a number of benefits to preregistration. The most formal preregistration is a registered report [OSF, 2022]. Registered reports first appeared in 2012 in the journals *Cortex* and *Perspectives on Psychological Science.* In 2013, *Social Psychology* adopted the practice. As of 2022, 300 journals across a wide variety of disciplines invite (or require) registered reports [Chambers and Tzavella, 2022].

Whereas traditional publication in peer-reviewed journals involves submitting the paper for review after the study is completed, the registered report is submitted and reviewed – twice. During the first stage, authors submit a detailed research proposal specifying research questions, hypotheses, methods,and planned analyses. Proposals that successfully pass through the review and revision(s) phase receive an "in principal acceptance", which commits the journal to publishing the final paper regardless of whether the hypotheses are supported. Once the research is completed, the authors submit the completed manuscript. Any deviations or additions to the protocol must be clearly identified. Substantial departures from the proposed analyses may result in a rejection at this second review. Because the purpose is to prioritize sound scientific practice over significant outcomes, if the authors followed the *a* priori specified protocol, the article should be published [Chambers and Tzavella, 2022]. By making the publication decision to accept-or-reject before the results are known, preregistered reports are an attempt to reduce biased research practices on the part of the researcher (e.g., HARKing [hypothesizing after the results are known], *p*-hacking, and selective reporting) and publication bias on the part of journals and reviewers [Chambers and Tzavella, 2022, van't Veer and Giner-Sorolla, 2016].

Not all journals invite registered reports. In these cases preregistration of studies is voluntary and researchers may wonder if preregistration is worth the time and effort. van't Veer & Giner-Sorolla [2016] have suggested that the detailed mapping involved in a preregistration may improve the overall quality of the study and that this, alone, is beneficial. Further, when reviewers and readers are aware that a study was preregistered, and that the plan was followed, the credibility of the findings may be enhanced.

Not surprisingly, there is resistance to preregistration. Moore [2016] described three primary concerns. First, preregistration constrains flexibility, exploration, and serendipitous discovery. Second, the preregistration invites additional scrutiny to the research process. Third, preregistration is more work.

Moore [2016] countered these concerns by suggesting that researchers are still free to engage in exploratory work. The preregistration itself is not limited to a priori hypotheses; researchers can specify their intentions to be purely exploratory or consider post-hoc analysis that follow the preregistered ones. Further, researchers can deviate from a preregistration; researchers are just expected to narrate how and why they did so. Regarding the concern of additional scrutiny, Moore argues that a study that followed a registered report or voluntary preregistration is likely to have greater credibility in the eyes of the reader or reviewer. Finally, Moore suggests that preregistration templates, specific to psychology, can streamline the process.

As an example for quantitative empirical research in psychology, the PRP-QUANT Template was designed by the Preregistration Task Force [2021]. The PRP-QUANT has three primary sections:

an introduction, method, and analysis plan. Each section includes multiple items that are accompanied with brief instructions [Bosnjak et al., 2021]. For qualitative researchers, a Delphi method [Haven et al., 2020] provides a 13-item, pregistration template that is freely available at the Open Science Framework. Both templates align with the Journal Article Reporting Standards[JARS; American Psychological Association [2020]].

2.4.2 Data sharing

Data sharing (including the releasing of raw data, measures, codebooks, and analytic scripts for data cleaning and analysis; Alter and Gonzalez [2018]) is a second pathway to a more transparent science. Proponents of data sharing argue that it (a) increases scientific integrity through greater transparency and the increased probability of reproducibility [Martone et al., 2018], (b) optimizes the value of data and will accelerate scientific progress when data are exposed to secondary analysis or combined in meta-analyses [Ross et al., 2018], and (c) creates a structure for greater collaboration [Bezjak et al., 2018]. Along with public and private entities, the U.S. Government contributes to data sharing when its agencies such as the Bureau of Labor Statistics, Department of Education, and Census Bureau offer data freely over the internet or restricted license.

The concept of data sharing is not new. In 1983, an article in the *American Psychologist* called for a mandate for data sharing [Ceci and Walker, 1983]). This was codified in the APA ethical principles as early as 1992. The current ethical principle [American Psychological Association, 2017] states that, "psychologists should freely share published data with peers requesting access for the purpose of verification or reanalysis." The language in the ethics code implies a one-to-one (researcher-to-requester) relationship; the current calls from funders and proponents of open science are calling for broader access [Martone et al., 2018].

Whether in an institutional repository, with the journal, or in an open source, collaborative, internet platform such as the Open Science Framework, data sharing involves archiving the data (along with descriptions and codebooks) and the record of the analysis [Stevens, 2017]. At the time of this writing, funding entities are increasingly requiring that data be made available for use by others [Ross et al., 2018] and journals are, similarly, requiring or encouraging such practices.

In contrast to the language in the APA ethical principles which suggest that sharing data should be "for the purpose of verification or reanalysis," most proponents of open data encourage the researcher to license the data such that there are no restrictions on reuse or redistribution. Exceptions could be made to protect the identity of the human participants or special limitations or restrictions related to ethical concerns [Bezjak et al., 2018].

Not surprisingly, there is resistance to data sharing. Beyond ethical considerations, common concerns include (a) being critiqued for analytical errors, (b) being "scooped" (i.e., someone else analyzes and publishes first), and (c) expending significant effort that will be unrewarded [Martone et al., 2018]. Regarding the first concern of "being critiqued," a core principle of open science is that all research results are available for challenge through reexamination, reanalysis, reproducibility, and replication [Alter and Gonzalez, 2018]. So, yes; being critiqued is a very possible consequence of sharing data.

In contrast, "being scooped" is a less likely consequence. In microbiology, the original data creators tended to publish two years after the data were made available whereas other authors tended to publish five or more years after the data were made available [Martone et al., 2018]. "Being scooped" may be less of a concern if researchers limit data sharing to the variables used in their

analysis and if it is shared at the time of publication. Countering the concern about "unrewarded effort," there are initiatives underway that would recognize the contributions of those who share data and analytic code. For example, the Association for Psychological Science has adopted the use of Open Science Foundation badges in its journal, *Psychological Science*. Further, evaluation of this project has suggested that a display of badges is correlated with significant increases in data sharing [Martone et al., 2018]. Several have argued that shared data, analytic code, and preregistrations should be assigned persistent identifiers (e.g., the DOI, digital object identifier) and be treated as scholarly products. That is, they should be listed on the contributor's curriculum vita and counted in professional evaluations such as promotion and tenure applications [Alter and Gonzalez, 2018].

There are also significant ethical concerns related to the protections and rights of research participants. In the U.S. it has been standard practice to (a) inform participants of the restricted purpose for which their data will be used and (b) assure participants that all the information they provide during the research study will (to the extent permitted by the law) be kept confidential, only to be viewed by members of the research team. When non-anonymous data are collected, the informed consent may also indicate that data will be de-identified. These standard practices are counter to the notion that data be shared broadly with researchers-in-general and could be used for purposes other than that which was stated in the informed consent [Ross et al., 2018]. Data collected where informed consent forms contained these traditional practices should not be shared. However, for new data collection, researchers who wish to share data should inform potential participants that consent extends beyond the present study; further, the nature and intent of future uses of the data are unknown [Ross et al., 2018]. The participants should also be informed about the types of identifiable private information that will be retained and the types of researchers who may have access to that information [Alter and Gonzalez, 2018].

Conveying this information to potential participants in a manner that provides them fully informed consent while also encouraging their participation in the study can be tricky. Here is some language we use on our informed consents:

> The information in the study records will be kept confidential. We do request your e-mail address so that we can send and link the quarterly surveys. Data will be stored securely and will be made available only to persons conducting the study unless you specifically give permission in writing to do otherwise. No reference will be made in oral or written reports that could link you to the study. Your de-identified data may be used in present and future (a) research collaborations, (b) scholarly/professional publications and presentations, and (c) in classroom teaching, projects, and demonstrations.

> Consistent with both journal/guild expectations and the ethical principles of open science, a fully anonymous and non-identifiable version of the response (i.e., dataset) may be posted online (e.g., to the APA-endorsed "Open Science Framework" (www.osf.io) or to the journal, submitted with the research article). This data may be reanalyzed for purposes that we cannot anticipate. No data posted will contain any information that could identify participants in any way, either directly or indirectly. All data will be thoroughly inspected by the Principal Investigators prior to posting to confirm that no participant-provided responses could inadvertently identify or expose a participant.

> Posting data (commonly referred to as "data sharing") is necessary for reproducibility and replicability in science, allows peer reviewers and meta-analysts to check statistical

assumptions, protects the field against data fraud, and is increasingly seen as an ethical obligation within psychological science.

Even with updated and IRB-approved informed consent forms, data sharing can be problematic, however. When samples or specific cell frequencies are small, certain combinations of information (e.g., tenure status, department, gender) could render a row of data identifiable. Another risk is when the data are used in unexpected ways that result in harm to the individuals or community. Ross et al. [2018] shared the story of blood samples being collected from an Indigenous group for what was believed to be a diabetes study. Later it was discovered that the data had been shared with other researchers to study topics that brought social and psychological harms to the tribe as a whole. Further, data that is shared in a public repository will likely be available globally where laws and ethical conventions for using research data may vary.

As chapter authors, we are generally proponents of open science. However, we urge researchers to give thoughtful consideration of plans for data sharing at the beginning of the project, to imagine intended and unintended consequences, and to seek IRB consultation and review. Further, data takes many different forms (e.g., qualitative, clinical interviews, survey data, geographical identifiers) and protection of the research participants will require different types of considerations before deciding if and how it can be shared [Ross et al., 2018].

2.4.3 Open peer review

Open peer review is another avenue for increased transparency in the scientific process. Although the term is interpreted differently, there are two primary mechanisms: *open identities* and *open reports* [Ross-Hellauer, 2017]. With open identities, neither the reviewers nor authors are anonymized. With open reports, the review reports are published alongside the relevant article [Ross-Hellauer, 2017]. As journals and other outlets experiment with open peer review, there may be variations of one or more of these mechanisms. Some have argued that published open reports – which could be cited by others and counted as a scholarly product for evaluative activities such as tenure and promotion – might incentivize scholars to accept peer review assignments (which are usually completed with no compensation) and invest the time and energy necessary to provide a constructive critique of the work and formative feedback to the researchers [Bezjak et al., 2018].

At the time of this writing, there is evidence of gradual movement toward open reviews. Some journals will now ask if the peer review can be transferred to another journal (if the manuscript is rejected) and if the reviewer's name can be transferred with it. While this is neither open identity nor open reports, it is a small step in the direction of sharing the work of peer review. Additionally, many journals are now inviting peer reviewers to register with Publons, a commercial organization that provides a mechanism for collecting and summarizing scholarly impact as a peer reviewer.

In creating this textbook, we engaged in a form of open review. It was important to us that each chapter be peer reviewed, yet we desired the peer review process to be non-anonymous, constructive, and formative (as opposed to anonymized, critical, and summative/gate-keeping). Along with a common rubric, drafts and revisions of each chapter were placed in a Google Docs folder. Two reviewers and the authors could access these materials at any time. Peer reviewers were asked to comment on each element of the rubric and leave suggested edits and comments/questions directly on the chapter draft.

2.5   Access to the Scientific Literature

Another aspect of open science concerns access to the literature. There are multiple forms of open access publishing [Bezjak et al., 2018, Shah, 2017]. *Self-archiving* is the process of placing a published version of an author's article into institutional repositories or websites. Sometimes self-archived articles have an embargo period (months to years) that must elapse before the article becomes open access. *Open access* publishing is immediately, freely available, upon its publication. This level of access usually involves an article processing charge (APC), a one-time payment by the author. A third type is the *hybrid article.* This happens when a pay-walled journal offers individual open access articles. This usually requires the author to pay a fee that is higher than the APC associated with open access. If an article is accepted into a hybrid model, it means that in a single journal issue, readers will find both open access and paywalled articles. Traditional journals are often motivated to use this model because it fulfills funder policies such as requiring immediate public access to research.

Not surprisingly, there are pros and cons to open access literature. Regarding pros, with no subscription fees, fees for individual articles, nor requirements to be associated with an institution who has access to the book or journal [Bezjak et al., 2018, Shah, 2017] open access materials are free to everyone, including in international contexts (e.g., low-middle income countries) where barriers to scientific literature may be greater. This benefits the potential readership as well as the author – in that there may be greater dissemination [and, in turn, citations; Hagger [2022]]. Another positive aspect of open access journals is an expedited submission-to-publication timeline [Shah, 2017].

Alternatively, while the rapid turnaround for open access articles may be viewed as a positive attribute of open access journals, this is not always the case. Sometimes, legitimate open-access publishers feel forced into this accelerated submission-to-publication timeline in order to compete. As a result, they may weaken their peer-review process to meet this deadline [Beall, 2012]. In addition, not all open access journals are well intentioned or legitimate. Predatory publishers cleverly spam researchers with calls-for-papers and fail to mention required authors fees that range from 1200 USD to 1800 USD. An author who learns of the APC after signing the contract (which generally includes surrendering the copyright) has lost the right to withdraw and is faced with paying the fee and losing the possibility to publish elsewhere – therefore, essentially losing their work.

Another common concern of open access research is that while these journals and articles are free to readers, this doesn't mean that readers are always able to locate these publication. A crucial piece in making articles accessible to readers is getting the journal indexed in as many relevant databases as possible [Fortney and Murphy, 2016]. APA PSYCH Info an example of a commonly used database in psychology.

A well-indexed article is more likely to be discovered and read regardless of whether or not the reader is familiar with the journal itself. Regrettably, for those hoping to join the open science movement, getting journals indexed can be incredibly difficult. In some cases, it will take multiple attempts over several years before a journal is accepted into an index or database [Fortney and Murphy, 2016, Shah, 2017]. While indexing journals is historically challenging, PubMed and Wellcome Trust are among some of the databases working with APA to increase accessibility of open access psychology research [Martone et al., 2018].

Finally, we think it is critical to raise equity concerns about the APC. Much of academia involves pressure to "publish or perish." Open access affords those who can afford APC a wider choice of

outlets for dissemination. Those without similar financial resources may be unable to logistically access these journals as an author and may need to rely on institutional support, grant support or the backlog of traditional journals. Thus, the APC associated with open access journals may further widen the equity gaps in tenure and promotion, making it easier for those with greater financial resources to accrue publications more quickly.

2.6 Tools for an Open Science

Tools that support open science abound and are constantly evolving and emerging. We review six types of tools including: statistical software, reference management systems, persistent identifiers, data repositories, collaborative platforms, and open educational resources (OERs). Behind many of these tools is the notion of open source. Open source tools are freely shared and the code (or platform) that powers them can be modified and redistributed.

2.6.1 Statistical Software

The R statistical software environment is an open source tool that includes features for commenting on code and enabling reproducible data analysis [Alter and Gonzalez, 2018, Bezjak et al., 2018]. Using R requires statistical training as well as fluency with the integration of base R, R Studio, and R packages. Because each of these elements are continuously updated, the R user must always adapt to changes in the underlying sourcecode that powers the analyses.

The birth of R might be associated with the 1997 launch of the CRAN (Comprehensive R Archive Network). The CRAN hosts R's executable files, source code, and packages contributed by the users [noa, 2022e]. *Mirrors* are the network of ftp (file-transfer-protocol) and web servers around the world that store identical (hence, "mirror") materials. There were only three mirrors in 1997; currently there are more than 100. As can be seen on the global list, they represent a variety of organizational types ranging from higher education to corporations to governments and nonprofits. Users are encouraged to select the CRAN that is geographically closest.

R Studio is a public benefit corporation, that is, a statutory or government owned corporation whose mission is to provide free or subsidized services to the public [noa, 2022f]. The organization provides open-source (zero-cost) and commercial software that is intended to serve in a "virtuous cycle" for mutual improvement [noa, 2022f].

Some are wary of open source statistical software, voicing concerns of trustworthiness. It may be reassuring to learn that the most common way to install an R package is through the CRAN. Packages available on the CRAN must adhere to its repository policy and are vetted prior to posting. Further, those engaged in psychological science who may scour the internet for "how-to" tutorials and blogs will quickly learn that there are a number of reliable, commonly used R packages to conduct the analyses that are frequently discussed and critiqued. Further, resources such as the peer-reviewed Journal of Statistical Software provide in-depth coverage of many R packages.

Because writing R script can be daunting, GUI (graphical user interface, point-and-click) alternatives are emerging that are open-sourced and, perhaps, easier to use. The R package, *shiny*, allows developers to build interactive tools known as "shiny apps." These apps tend to perform limited functions. An example is Shoemann et al.'s [2017] Monte Carlo power analysis for indirect effects (https://schoemanna.shinyapps.io/mc_power_med/).

More comprehensive software tools are also built with R code. For example, the program *jamovi* [jamovi project, 2021] was designed as an alternative to fee-for-use programs (e.g., SPSS) for the social sciences. Although the user points-and-clicks, a syntax mode allows the production of R syntax for inputting directly to R or for retrieval as a completely reproducible project. Although three individuals are credited as co-founding jamovi and their website lists additional team members [jamovi project, 2021] there is not information about its organizational status. The jamovi website seeks financial contributions and volunteers for advocacy, content creation, and module development.

JASP ["just another statistics program", Team, 2022] is another GUI that is commonly used in psychological science. Sponsored by the University of Amsterdam, JASP's features include both frequentist and Bayesian analyses. Further, JASP produces APA-formatted tables that can be copy-pasted into word processing documents. Although it is a future goal of the developers, at this time, it is not easy to retrieve R code from JASP for reproducible archiving. Users of open-source software should include both text and reference list citations. Citing the software (a) provides a complete description of the method (contributing to reproducibility), (b) documents the usage and development of that software in the developer's field, and (c) credits (in the form of a scholarly citation) the developer [Smith et al., 2016].

2.6.2   Reference management software

Zotero [for Digital Scholarship, 2022] is one example of an open source reference management program operated by the non-profit group, Corporation for Digital Scholarship. Zotero stores, manages, and cites bibliographic references. Zotero allows (a) cloud and locally held storage, (b) collaboration, (c) and full integration with word processing documents and R markdown files. Zotero's open source nature makes it immediately responsive to change; merely three months after the introduction of the 7th edition of the APA style manual, Zotero upgraded the default style to match. Although a non-profit, Zotero does operate on a freemium model where basic services are free, but cloud storage (allowing synchronization across devices and collaborators) requires a subscription fee.

Another popular tool that offers similar functionality is Mendeley. In 2007, Mendeley was founded in the U.K. by three doctoral students from Germany. In 2018 it was purchased by the academic publisher, Elsevier [noa, 2022c].

2.6.3   Persistent identifiers

You may have noticed that most items in our reference lists have DOI (digital object identifiers) numbers. The DOI is an example of a persistent identifier (PID). PIDs are long-lasting digital references to objects, people, or organizations that serve to provide a reliable link from citations to the publication (or its source). There are two parts to PIDs. First, is the identifying alphanumeric string, itself. Second, is the organization or agency that commits to providing an infrastructure to ensure that a URL will map to the correct location of the object.

Owing to name, organizational, and geographic changes, researchers can also become disconnected from their work. The ORCID (Open Researcher and Contributor ID) was created as an independent nonprofit organization to provide a unique, persistent identifier to researchers. The sponsoring organization has also created an infrastructure to ensure the reliability of these connections. When

scholarly products include the authors' ORCID, there can be a permanent and clear record of research activities. Many journals request ORCID for authors and co-authors at the time of submission; ORCIDs are also commonly requested of peer reviewers.

2.6.4 Data repositories

Data repositories collect, maintain, and disseminate data over time [Alter and Gonzalez, 2018]. This is accomplished, in part, by providing a public facing citation and assigning persistent identifiers (e.g., DOI, digital object identifiers). Data repositories exist across a number of institutions and platforms. For example colleges and universities may archive theses and dissertations. Other data repositories are more disciplinary-specific.

2.6.5 Collaborative platforms

Collaborative platforms are online services that provide a virtual environment where multiple people can connect and work on the same task [Bezjak et al., 2018]. If you have used Google Docs, Dropbox, or the Microsoft packages such as OneDrive or Sharepoint, you have used a collaborative platform. While these are terrific tools for word processing, spreadsheets, and slide presentations, they may less helpful in co-authoring statistical code.

GitHub [noa, 2022a] is an example of a cloud-based, collaborative platform, that specializes in hosting code (e.g., R code) where multiple people can contribute, track, and control changes to the code. GitHub is also host to numerous open source projects. For projects that are openly licensed, others can "fork" (i.e., copy the project for yourself and make changes to it without altering the original) the project. Additionally, so long as the secondary user credits the author/developer, they can use, revise, remix, and further distribute its contents. Initially a start-up business, Microsoft purchased GitHub in 2018 noa [2022b]. GitHub's basic features are free for individuals and organizations; there are more advanced tools for a fee. Our OER is produced and hosted on GitHub and GitHub pages for zero cost.

The Open Science Framework [OSF, 2022] was created by the not-for-profit Center for Open Science and is a collaborative infrastructure to support the entire research cycle. Across multiple disciplines, including psychology, researchers can preregister studies and use the same project for sharing data and analytic code (as well as literature, IRB materials, experimental materials, presentations, preprints). OSF also offers the capacity to connect with other systems such as GitHub, Google Docs, and Dropbox. The GitHub and OSF are only two examples. Other commonly used collaborative platforms include Zenodo and Figshare [Martone et al., 2018].

2.6.6 Open education resources

Open education resources (OERs) include any tools, materials, or techniques that are used to support access to knowledge [Bezjak et al., 2018]. While not specific to open science, there is considerable overlap in the values and resources that contribute to them and emerge from them. This very textbook was created using the open source software R and R Studio. The primary packages used to format the content are R Markdown and Bookdown. The book is hosted in GitHub and is rendered to the internet via GitHub Pages. As described in the preface, the OER holds a CC BY-SA 4.0 license allowing the user to share (copy and redistribute the material in any

medium or format) and adapt (remix, transform, and build upon the material for any purpose). The license requires that proper attribution is made (e.g., appropriate credit, a link to the license, and indication of changes were made) and that your redistribution must use the same CC BY-SA 4.0 license.

2.6.7   Bearing the Costs of "Open"

Our attention to "who owns" open tools may be a surprising inclusion in a chapter on open science. We were intentional in providing this information because the who of ownership and the type of business model may have bearing on the trustworthiness with which researchers' tools, materials, and data may be protected and maintained. Further, as the popularity of open tools increase (and decrease) there is the hope that tools will improve, the risk that the tool will be ignored or deprecated, and the possibility that use fees will be added or increased. Our only recommendation is that potential users take some time to become familiar with open tools before committing to use them.

2.7   Summary, Conclusions, and/or Recommendations

Researchers have demonstrated that psychological science is threatened by problems of reproducibility and replicability [Open Science Collaboration, 2015]. Through preregistration, data sharing, and open peer review, proponents of open science have identified pathways to increase the transparency of the scientific process. Ensuring open access to research findings remains problematic. Many findings are behind pay walls (restricting their access) and open access journals often charge fees that are prohibitive for new scholars and those groups who experience marginalization. More work will be needed to solve this access. In the meantime, tools to facilitate open science are abundant, emerging, and constantly improving. We encourage emerging researchers to participate in open science. Further, as it continues to develop, we encourage all to be vigilant so that the new practices are not co-opted in ways that maintain historic and inequitable power structures and privileges.

2.8   Suggestions for Practice, Further Learning, and/or Conversation

1. Obtain an ORCID iD for yourself
2. Find a published study that had been preregistered. Trace it to its preregistered location. Is there also access to data, analytic code, and other materials? How easy was it to locate? How easy would it be to reproduce the study?
3. Download one of the preregistration templates (either quantitative or qualitative) and preregister a study of your own.

2.9   References

WRITING EMPIRICAL MANUSCRIPTS

Chapter 3

The APA Style Manuscript

Lynette H. Bikos (she/her) & Kiana Clay (she/her)
Seattle Pacific University
Seattle Pacific University was built on the unceded ancestral lands of the Duwamish people
Chapter Status: In Progress/Placeholder

This is sample text as we learn how to use the Github. Ridiculous sentence on line 14. Another ridiculous line to volley it back.

This lesson focuses on APA style. While a significant portion of the material focuses on the mechanics of APA style, I also consider:

- APA style as epistemology – and its relationship with power, privilege, and racism.
- JARS: Journal Article Reporting Standards
- Section by section
- Consideration of writing as a developmental process that occurs throughout grad school and into the profession
- Stylistic issues
- Hallmarks of APA style
- Reducing bias

3.1 Navigating this Lesson

There is about 1 hour and 45 minutes of lecture.

The R markdown file used to create this lecture as well as all of the figures are available at the OER's GitHub site.

3.1.1 Learning Objectives

Learning objectives from this lecture include the following:

- Identify the cultural characteristics of APA style.

- Identify two ways that Thompson (2004) has suggested that APA style perpetuates Whiteness and patriarchy in the academy.
- List the components of an abstract.
- Describe why JARS matters.
- Begin memorizing the minutia of APA Style for writing empirical manuscripts/journal articles (no particular order).

3.1.2   Readings & Resources

In preparing this chapter, I drew heavily from the following resource(s). Other resources are cited (when possible, linked) in the text with complete citations in the reference list.

- American Psychological Association. (2020). *Publication manual of the American Psychological Association: The official guide to APA style. (Seventh edition.).* American Psychological Association.

    - I'm guessing you'll use this more days than not, for the rest of your education.

- Tran, A. G. T. T., & Lee, R. M. (2014). You speak English well! Asian Americans' reactions to an exceptionalizing stereotype. *Journal of Counseling Psychology, 61*(3), 484–490. https://doi.org/10.1037/cou0000034

    - I use this article in several analyses in the ANOVA series as well as in this lesson when I compare/contrast it to the requirements of APA Style. This article was pubbed in 2014; but I will compare it to the 7th edition (2019) standards.

- Cooper, H. (2020). *Reporting quantitative research in psychology: How to meet APA style journal article reporting standards (Second edition, revised.).* American Psychological Association. https://alliance-primo.hosted.exlibrisgroup.com/permalink/f/rpqmv/CP71332049420001451

    - The e-text version of this may be available at your library. This resource offers section-by-section instruction of the reporting standards and provides numerous examples of writing APA empirical manuscripts.

- Madigan, R., Johnson, S., & Linton, P. (1995). The language of psychology: APA style as epistemology. American Psychologist, 50(6), 428–436. https://doi.org/10.1037/0003-066X.50.6.428

    - Madigan et al. (1995) argued that as we learn APA style we are inculcating the professional values of our discipline (and we do this without awareness).

- Thompson, A. (2004). Gentlemanly Orthodoxy: Critical Race Feminism, Whiteness Theory, and the APA Manual. *Educational Theory, 54*(1), 27–57. https://doi.org/10.1111/j.0013-2004.2004.00002.x
- Critiquing the 5th edition of the style manual (we're now on the 7th) Thompson (2004) pointed out how aspects of APA style contribute to preserving Whiteness.
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191

- – Yet another copy of the most current JARS. You can also find them in the style manual and on their own website: https://apastyle.apa.org/jars/glossary

- *White Supremacy Culture.* (n.d.). DRworksBook. Retrieved August 8, 2020, from https://www.dismantlingracism.org/white-supremacy-culture.html

  - – Identifies characteristics of White Supremacy Culture in organizations (often used to describe academia).

3.2 APA Style as Epistemology (or Worse)

A 1995 article [Madigan et al., 1995] in the American Psychologist compared APA style to that of other disciplines (history, literary criticism) and argued that APA Style is its own writing genre characterized by (among other things):

- A story schema: introduction, method, results, discussion – Seems so linear; but this is rarely the case (research is messy). This schema serves as an implicit formula for addressing psychological questions:

1) Review of past empirical studies
2) Collection and analysis of new data
3) Relation of conclusions to current psychological theories

- A depersonalized language of disagreement that focuses on the empirical process and away from investigators as individuals (e.g., "The primary criticism is that the threshold-setting procedures used in previous experiments are not adequate to ensure that…"). The goal is a collaborative, cumulative endeavor based on research data that avoids animosity and confrontation.
- Hedged conclusions: balancing a need to have substantive conclusions that do not extend beyond the data. We use the words, "tend," "suggest," and "may" (see Table 1) to allow authors to relate findings, while still recognizing the need for additional research before asserting firm conclusions. **NOTE: INCLUDE TABLE 1!!!
- A system of headings/subheadings (rather than narrative transitions)
- Paraphrasing rather than directly quoting other works. – Giving authors more flexibility in representing others' perspectives.
- Multiple authors (perhaps contributing to "more subdued prose").
- Heavy use of citations in the introduction and discussion sections so that the research is placed within an ongoing stream of empirical studies. The use of in-text references provides vital background for the study and helps the author establish credibility as an expert on the subject they are researching.
- Language that "does not call attention to itself." It can be described as: self-effacing and low-profile. Colorful language and attention-grabbing metaphors are rarely used; the focus is meant to be on the phenomena being studied.

The style manual has grown from its 7-page writers guide in the Psychological Bulletin (1929) to 427 in the 7th edition (2019), today. Madigan et al. [1995] suggests that the process of mastering APA

style, one is enculturated into psychology. That is, we learn APA style – we inculcate empiricist values (i.e., "unarticulated practices that reflect fundamental attitudes and values of psychologists" (p. 428). As such, APA style is epistemological.

Nearly a decade later, Thompson [2004] examined APA style through the lens of critical race feminism and Whiteness Theory. Thompson argued that the expectations regarding clarity precision, appropriateness, sensitivity, and objectivity likely contribute to the academy's investments in Whiteness and patriarchy. While her focus was on APA Style, she suspected that her critique would generalize to other style guides (e.g., Chicago, MLA).

Thompson's [2004] article focused around the gender/sexuality and race values codified by the APA manual (at the time of her article it was the 5th edition). She conducted an analysis of its *power and property investments* and organized her arguments in five themes:

Property Rights: PWIs have treated refereed scholarship (but not indigenous or community knowledge) as intellectual property, demarcated with a certain class. That is, individuals "own knowledge." Evidence includes:

- Using last names as a shorthand reference for work (e.g., "Dik & Duffy's 2012 CVQ").
- Tendency to cite own work rather than say, "When I previously said…"
- The style manual indicates that principal authorship and order of authorship credit should reflect the relative contributions of persons involved (7th ed, APA 1.22) and that relative status should not determine authorship order. Thompson argued that power dynamics (especially around race and gender) likely interfere with this principle.
- The 7th edition has reduced citations of articles with three and more articles to "First author et al., YEAR" for all citations.

- Regarding, ownership of knowledge – does it belong to the researcher or the community from which it came? Community knowledge can be studied but not cited. And what about institutions? What if knowledge came from the Black Church? How does one cite that?

Precedent and pedigree: In the social sciences, we are expected to cite and give credit to relevant earlier works. Knowledge is seen as cumulative. Citations are essential.

- No "unhedged statement" (p. 45) is made without a citation; and in introductory sections, citations are often included with claims that seem obvious! This establishes the requisite lineage.
- Self-citations, expected citations (of the gurus), citations of colleagues exist IN TENSION with a citation economy (page limits, after all).
- The requirement that scholars locate their project in the context of the existing peer-reviewed literature serves to keep out ideas/projects that would be challenging to the existing power structure.

Proceduralism: APA authors learn to address an audience unmarked (as if they were white, male, heterosexual) by gender or race as a "sign of professionalism" (p. 48).

- Standardized Format: a four section structure: Introduction, Method, Results Discussion; it could be that the experiences of marginalized and oppressed groups cannot be captured by this scientific narrative structure.

- Standardized Style: prizing scientific appearance and elegance. Thompson (2004) follows the guidelines regarding the use of the "slash" (/) and how it is impossible to standardize to the degree that it works for all groups.

    - The "First author et al., (year)" citation contributes to both gender- and color-blindness. On the one hand is it "fair." On the other hand, it obscures the person of the author.
    - APA's prohibition against footnotes minimizes "the good stuff" (the most juicy arguments are always in the footnotes).

Protocol: Propriety or protocol represents conduct that signifies one's understanding of prevailing relations of power, authority and legitimacy.

- The style manual (7th ed, Chapter 5, e.g., 5.2) acknowledges the importance of sensitivity and the avoidance of pejoratives to reference groups of people. Thompson indicates that this is "admirable" it fails to address unequal power relations. Further, distinctions between what is insensitive and pejorative may be invisible to those in positions of privilege.

The gentleman's agreement: APA style is characterized "language that conveys professionalism and formality" (7th ed, 4.7, 4.8) and "differences should be presented in a professional and noncombative manner" (7th ed, 4.7).

- Thompson [2004] is concerned that while this is offered with the hope of pluralism and the creation of safe spaces, it causes controversies to be ignored or dismissed and may bolster complicity in racism.

3.3 As We Dive into the Specifics

Please keep the perspectives of these authors in mind.

Let's also look at the characteristics of White Supremacy Culture in organizations [noa]. As we tour through the components of APA style, what resonates with this list? Refer to the .pdf handout or website for more definitions, descriptions, examples, and antidotes.

- Perfectionism

    - Perfectionistic culture
    - Worship of the written word
    - Only one right way
    - Either/or thinking

- Concentration of power

    - Power hoarding

- Paternalism

    - Defensiveness

- Right to comfort

  – Fear of open conflict

• Individualism

  – I'm the only one

• Progress is more/bigger

  – Objectivity
  – Quantity over quality
  – Sense of urgency

3.4   The JARS: The Core of APA Style

The JARS [Journal Article Reporting Standards; Appelbaum et al. [2018]] were first introduced in a 2008 feature in the American Psychologist [noa, 2008] and were included in the 6th edition of the style manual. The updated JARS, published in 2018, expanded the types of quantitative research (JARS-Quant) and included standards for reporting qualitative (JARS-Qual) and mixed methods (JARS-Mixed). Chapter 3 of the 7th edition is devoted to JARS. It contains numerous tables, definitions/explanations, and a flowchart.

The JARS are an attempt to represent what is common across approaches. There is a recognition that specialties/sub-specialties use different terminology. The terms, therefore, should be treated as placeholders and be updated to reflect the various research traditions.

In each of the sections below, I list both the APA style recommendations and JARS elements – which are somewhat annoyingly separate, adjacent, and overlapping.

In the next section I used text-citations to also refer to the section numbers of the 7th edition of the APA style manual. Additionally, because this OER is publicly available (i.e., not just used in my classroom), I have not copied the JARS elements into this document. They are freely available here.

3.4.1   Title, Authorship, Author Note (APA 2.3)

***ARE WE KEEPING THE QUIZLETS? QUIZLET: A manuscript title should

  a. Use abbreviations whenever possible
  b. Contain at least 30 words
  c. Be fully explanatory when standing alone
  d. Begin with the words, A Study of

The title page of a manuscript includes the

  a. Author's name
  b. Author's institutional affiliation'
  c. Running head
  d. Short title
  e. All of the above

3.4.1.1 Title (APA 2.4)

- Concise statement of main topic of the research,
- Identify the variables or theoretical issues under investigation (and the relationship between t Addhem),
- Focused and succinct (no prescribed word or character limit per the style manual; but a journal may have one).
- Avoid empty words/phrases like "A study of," "An investigation of"

The lecture further reviews JARS elements.

3.4.1.2 Authorship/Byline (APA 2.5) & Affiliation (APA 2.6)

- Primary contributors;
- Institutional affiliation AT THE TIME of the study, no more than 2 institutional affiliations (and they both need to have contributed equally)
- Very specific style guidelines (with superscript notations) for connecting authors and their affiliations
- In the order of contribution – lots of ethical, practical, and sensitive considerations about who gets to be an author and in what order.

Sticky Issues about Authorship

- Faculty as first authors
- Ongoing projects with years of investment sponsored by faculty
- Who gets "the call"? (is "the call": who gets to decide who is first author?)
- Order between equal contributors (e.g., Singer & Willett)
- In equal contributions it is okay (but not ideal) to mention in the author note
- GOAL: name on a project, less concerned about author order
- Generally, graduate projects include the faculty sponsor as an author
- Author-order rubrics can be useful to guide the decision

3.4.1.3 Author Note (APA 2.7)

Place the author note at the bottom half of title page. There are more specific instructions in style manual. Generally author note has four paragraphs:

- ORCID iDs https://orcid.org
- Changes in affiliation subsequent to the time of the study, "[Author's name] is now at [new affiliation]." Can also be used if an author is deceased.
- Disclosures and acknowledgements (i.e., study registration, open practices and data sharing, related reports, conflicts of interest, financial or other assistance).
- Contact information for corresponding author. Requires full name, complete mailing address, and e-mail. Prescribed format: Correspondence concerning this article should be addressed to [author's name'], [complete mailing address]. Email: author@institution.edu

The lecture further reviews JARS elements.

3.4.1.4 Running Head (APA 2.8)

- Abbreviated title printed at the top of the pages of a ms or pubbed article to identify it for the readers
- Max of 50 characters (counting letters, punctuation, spaces between words). If the title is already 50 characters or fewer, the full title can be used. Avoid abbreviations in the running head, although the ampersand can be swapped (& for "and").
- Appears flush left in all uppercase letters at the top of the title and subsequent pages

3.4.1.5 Manuscript Page Headers/Page Numbers (APA 8.03)

- Number consecutively, beginning with the title
- Identify each page with the running head along with the page number
- To facilitate a blinded review, do not include your name in the page header
- Use the automatic functions of the word-processer to generate page headers and page numbers

3.4.2 Abstract (APA 2.8, 3.3)

QUIZLET: An abstract should

a. Appear on the same page above the title and introduction
b. Be single-spaced and set with larger margins
c. Begin on page 2
d. Be no longer than 3% of the text

Abstract lengths vary from journal to journal; the typical range is from _____ to_____ words. (250 to 500???)

A good abstract is:

- Concise and specific
- Nonevaluative
- Coherent and readable
- Loaded with keywords
- May be a single paragraph (no indentation of first line) or structured (still no indentation, but labels inserted for the prescribed sections).

3.4.2.1 Recipe for an Abstract

- Identify the problem under investigation (1 sentence)
- Identify the participants and salient characteristics
- Identify the experimental method, including the apparatus, data gathering procedures, complete test names, etc.
- Identify findings (including statistical significance levels)
- Identify conclusions and implications/applications.

- Avoid *boilerplate* sentences. – At the end of an abstract, I often read, "Findings and implications will be discussed." This is boilerplate because "everyone says it." It's also empty and a waste of words because it tells us NOTHING about the study.
- The style manual includes outlines for a variety of types of manuscripts.

The lecture further reviews JARS elements.

3.4.2.2 Keywords

Keywords include single words, phrases, or acronyms that describe the most important aspects of the paper. The purpose is indexing in databases (it's what we search on in database like PsychInfo).

3.4.3 Introduction (APA 2.11)

QUIZLET: The introduction section of a research report should:

- a. Include a thorough historical review of the literature
- b. Define all of the terms that would be unintelligible to a reader with no previous exposure to the field
- c. Present the specific problem to be explored and described in the research strategy
- d. Be clearly labeled.

What question should the introduction section of a research report attempt to answer? a. What are the theoretical implications of the current research? b. What is the point of the study? c. What is the logical link between the problem and the research strategy? d. All of the above are correct. e. Only a and c of the above are correct.

Before you write, consider this:

- Why is this problem important?
- How does this study link to previous work?
- What are the primary/secondary hypotheses/objectives? What are their links to theory?
- How do the hypotheses and research design relate to each other?
- What are the theoretical and practical implications?

Three parts to the Introduction

- Brief introduction, 1-2 paragraphs
- Review of relevant scholarship, 1-2 pages
- Purpose, rational, hypotheses, 1-2 paragraph's at the section's close.

3.4.3.1   The Brief Introduction

A good introduction paragraph answers the following questions in 1-2 paragraphs

- What is the point of the study?
- How do the hypothesis and experimental design relate to the problem?
- What are the theoretical implications of the study and how does the study relate to previous work in the area?
- What are the theoretical propositions tested and how were they derived?

In the middle...above ALL

- Demonstrate that the problem is IMPORTANT
- Narrative thread.
- Always refer to the THEORETICAL BACKGROUND
- Narrative thread.
- Closing paragraph a statement of purpose and rationale of project.
- Narrative thread.

The closing paragraph(s) of the Introduction

- What variables did I plan to manipulate?
- What results did I expect and why did I expect them? (the logic of this should also be made clear)
- What is my rationale for each hypothesis?

The lecture further reviews JARS elements.

3.4.3.2   Bikos' Developmental Perspective on Learning to Write the Empirical Manuscript

Over the years I have been learning writing skills of my own and, in my roles as instructor and peer reviewer, providing feedback to others. Below is an oversimplified model of what I think I'm observing as students and scholars learn to write empirical manuscripts (in the specific APA style).

Figure 1. The first stage of writing.

Figure 2. The second stage of writing.

Figure 3. The third stage of writing.

Figure 4. The fourth stage of writing.

3.4.3.3   Perennial Notes from Research-Project Faculty Graders

- Start with your problem statement – What is your DV and why does the reader care?
- Theory must drive your model/manuscript. – "A list of variables does not a theory-driven model make."
- Active headers are your friends – Not helpful: "Gender & RSB" – More helpful: "RSB vary across Genders"
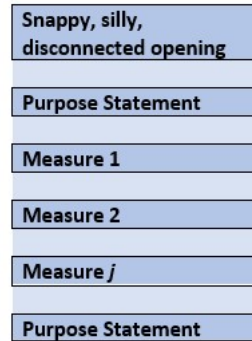
Figure 3.1: Image of the first stage of writing

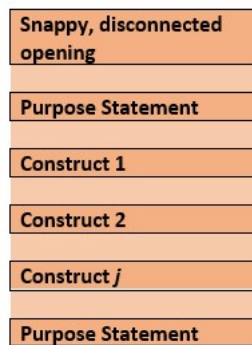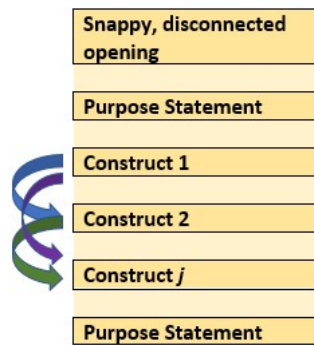Figure 3.2: Image of the second stage of writing
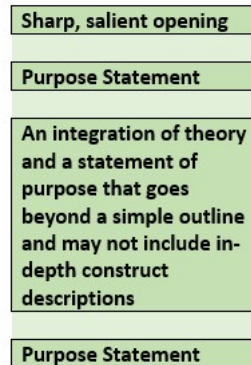

Figure 3.3: Image of the third stage of writing

Figure 3.4: Image of the fourth stage of writing

3.4.4   Method (APA 2.06; Chapter 3)

- Describes in detail how the study was conducted.
- Includes conceptual and operational definitions of variables.
- How much info is enough? Readers must be able to… – Evaluate the appropriateness of the method – Judge the reliability and validity of your results – Replicate the study

Many of the JARS elements have even more detail in the text that follows the table. In your first few papers, make sure to review and imitate.

The lecture further reviews JARS elements.

Regarding these tables – flowchart and tables in style manual, Appelbaum (2018) article, or website.

3.4.5   Results & Discussion (Chapter 3)

The lecture references the JARS elements in detail.

Section 6.43 of the APA style manual details the basic form for statistical output ("statistical strings"); 6.44 provides statistical abbreviations and symbols.

3.4.6   Headings (APA 2.7)

The 7th edition of the style manual updated the levels of headings. The change is to the L3 heading – making it a little easier for Word processing programs.

There are now 5 levels of headings. A useful table that illustrates them can be found here.

Headings follow a top-down progression.

- Each section (i.e., Title [on first page], Method, Results, Discussion) starts with an L2 heading.
- Do not include an "Introduction" heading – the first paragraphs of a paper are automatically expected to be introductory in nature. Here's the heading structure in the paper (a "brief report") we are investigating:

You Speak English Well! Asian American's Reactions to an Exceptionalizing Stereotype Microaggressions Against Asian Americans (L1, centered, bold, title case)

Method (L1, centered, bold, title case)

Sample (L2, flush left, bold, title case) **Procedure and Materials** (L2, flush left, bold, title case) **Analysis and Manipulation Check** (L2, flush left, bold, title case)

Results (L1, centered, bold, title case)

Discussion (L1, centered, bold, title case)

Conclusion (L1, centered, bold, title case)

3.4.7   Reference List

The reference list should start on a new page with "References" in bold and centered (an L1 heading – like the title).

The general format of the reference is: Author (Date). Title. Source. The 7th edition has made significant improvements in making this format consistent throughout the many types of references. Plus, Chapter 9 in the 7th edition has a number of tools for problematic reference entries.

One substantial change in the 7th edition is the immediate use of "et al." after the first author's name when there are three or more authors. In contrast, up to 20 authors can be included in the entry in the reference list.

Use a hanging indent and format this with the paragraph-formatting function of your word processor's ruler. This means that the first line of each reference is flush left and subsequent lines are indented by 0.5 inches.

The text-citations and reference lists must be 100% consistent. If you have consistently (and properly) used your reference management system, this will happen by magic. Even so, one of the last things I do is a text-ref/ref-list cross check. I open 2 copies of the manuscript and put them side-by-side on the monitor. Starting with the text (and in track changes), each time I come to a reference, I put an X in front of it on the reference list. I go through the entire document. If there are items in the reference list missing, I note it, and then add it. If there are additional references in the reference list that I have not checked, I do a quick search to see if they are really missing, and if so, I delete them. ****SHOULD THIS PART BE KEPT? For dissertation proposals and defenses, we ask first and second year students to do this for us – it familiarizes them with RVT topics, the dissertation process, and the literature.

Because of Zotero, Mendeley, RefWorks, EndNote, and other reference management systems (as well as the good tools in the style manual), I won't say more about text-citations and reference lists.

3.4.8   Stylistic Issues (APA Chapter 4 on *Writing Style and Grammar*)

APA (4.0) is committed to a literary style characterized by continuity (logical consistency of expression throughout a written work), flow (smooth cadence of words and sentences), conciseness (say only what you need to say), and clarity. I encourage the authors to read Chapter 4 of the style manual for examples of how to increase the effectiveness with which they write.

- Transitional words between sentences and phrases (or sentences) between paragraphs and sections assist with continuity and flow – Time: then, next, after, while, since – Cause-effect: therefore, consequently, as a result – Addition links: in addition, moreover, furthermore, similarly – Contrast links: but, conversely, nevertheless, however, although – Think twice about using adverbs (certainly, consequently, importantly, interestingly) as transition words.
- Active headings/subheadings contribute to continuity and flow.
- Don't add any more words than necessary. Say *only* what needs to be said: – No "fluff" (avoid: very, really) – Maintain a professional tone (avoid cliché's, rhymes, alliteration, metaphors, slang – remember a global audience)
- Clarity is improved with a varied sentence length; long sentences can be confusing.
- APA style permits both active (subject, verb, object: "students completed surveys) and passive (object of verb is presented first: "surveys were completed by students") voice. The active voice contributes to direct, clear, and concise statements.

More style-congruent advice: * Avoid abstract stacking. * Avoid organizing by an author(s)' name(s). – Organize by ideas – Unless it is important, citations should be at the end of a sentence. Ask yourself why you are starting a sentence with an author's name – Is he or she the point of your sentence? Could be but most likely not. * Use quotations sparingly: short ones, infrequently. * Precision, continuity, clarity, efficiency (stop with the long sentences) * Active voice ("It" is banned as first word in paragraphs and sentences). * Manuscript must be written in first person. – Avoid this: "Hypotheses for the present study included…" – Do this instead: "In our study, we hypothesized…" * Past tense or present perfect tense for the literature review, method, and results sections; present tense for the discussion section.

3.4.9 Reducing Bias (APA Chapter 5)

Chapter 5, "Bias-Free Language Guidelines" is new to the 7th edition. A big message is that this is-and-will-continue-to-be-evolving.

Here are some of the highlights

- Describe relevant characteristics; it is likely unnecessary to collect and/or report on all personal characteristics for the topic being explored.
- Do not shy away/hide/gloss over actual differences in the target population from the general population. Describe them clearly and professionally.
- There are specific guidelines for level of specificity in labels/language for various topics (e.g., disability, age, gender identity, race/ethnicity/nationality, sexual orientation, SES). Look them up.
- A few guiding principles (but there are so many more…read the chapter): – Call people what they call themselves (especially in qualitative or participant-informed research designs). Also recognize that no group is monolithic so there may be within-group differences about this. – Accept that language changes with time – stay open minded. – There is debate about person-first language ("a child with autism") and identity-first language ("an autistic child"). Get consultation and listen to the research participants. – "White," "cis-gendered," "heterosexual," and "U.S." is often the standard against which others are judged. This centers Whiteness. Avoid these false hierarchies. When these comparisons are made, placing the socially dominant group on the left side of a graph or at the top of a table may imply these groups are the universal standard.

A (very) few specific details: * On identification of race: – "African American" is discouraged as an umbrella term for people of African ancestry worldwide because it obscures other ethnicities or national origins (e.g, Jamaica). In these cases use Black (with a capital "B"). – "Caucasian" is discouraged as an umbrella term for White or European because it originated as a way of classifying White people as a superior race. Use White (with a capital "W"). – "Native American" and "Native North American" seem to be preferred in North America. – "Latin@" is now widely accepted for Latino and Latina; as is "Latino/a." "Latinx" can also be used as a gender-neutral or nonbinary term inclusive of all genders.

* Avoid writing about "minorities." Terms like "people of color" or "underrepresented groups" are preferred. Do not assume that members of any of these groups are "underprivileged." Terms like "economically marginalized" or "economically exploited" may be preferred. * Use "sexual orientation" not "sexual preference," "sexual identity," or "sexual orientation identity." The term "sexual and gender minorities" refers to multiple sexual and/or gender minority groups or write about "sexual orientation and gender diversity."

– LGBT is considered outdated, but LGBTQ, LGBTQ+, LGBTQIA, AND LGBTQIA+ may be used to refer to multiple groups (but be careful and specific; if you are researching transgender people...then talk about transgender people). * "They" is now acceptable to use in the singular. There are a number of circumstances when this is recommended. A few include: – When it is the person's preferred pronoun. – When the pronoun is unknown. – In place of "he or she," "s/he," and "(s)he."

3.4.10   Closing Thoughts

1. Fluency in APA style is a long journey. One of my own strategies in an APA Cheat Sheet (presently 38 pages long).

a. I wrote it to help with reviewing/editing/grading (and also for my own writing), so I can search on words/phrases that I am likely to say.
b. I can add/update it.
c. Each element in my cheatsheet has the section number (for papers I review as well as my own self-doubt when I wonder, later).

2. Having now toured the 7th ed, what are your thoughts about its relation to power, privilege, racism/anti-racism, and White supremacy culture?

Repeated are the characteristics of White supremacy culture. I have italized those that strike me as being reflected in APA Style.

- Perfectionism

  – *Perfectionistic culture*
  – *Worship of the written word*
  – *Only one right way*
  – Either/or thinking

- Concentration of power

  – *Power hoarding*

- Paternalism
  - Defensiveness
- Right to comfort
  - *Fear of open conflict*
- Individualism
  - I'm the only one
- Progress is more/bigger
  - *Objectivity*
  - Quantity over quality
  - Sense of urgency

Bibliography

White Supremacy Culture. URL https://www.dismantlingracism.org/white-supremacy-culture.html.

Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist*, 63(9):839–851, December 2008. ISSN 0003-066X. doi: 10.1037/0003-066X.63.9.839. URL https://ezproxy.spu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=psyh&AN=2008-17400-001&site=ehost-live. Publisher: American Psychological Association.

GitHub, 2022a. URL https://github.com.

GitHub, July 2022b. URL https://en.wikipedia.org/w/index.php?title=GitHub&oldid=1097996266. Page Version ID: 1097996266.

Mendeley, June 2022c. URL https://en.wikipedia.org/w/index.php?title=Mendeley&oldid=1095492319. Page Version ID: 1095492319.

Open science, May 2022d. URL https://en.wikipedia.org/w/index.php?title=Open_science&oldid=1087027147. Page Version ID: 1087027147.

R (programming language), July 2022e. URL https://en.wikipedia.org/w/index.php?title=R_(programming_language)&oldid=1098779048. Page Version ID: 1098779048.

RStudio, July 2022f. URL https://en.wikipedia.org/w/index.php?title=RStudio&oldid=1099226518. Page Version ID: 1099226518.

George Alter and Richard Gonzalez. Responsible practices for data sharing. *American Psychologist*, 73(2):146–156, February 2018. ISSN 0003-066X. doi: 10.1037/amp0000258. URL https://ezproxy.spu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=psyh&AN=2018-06770-004&site=ehost-live. Place: US Publisher: American Psychological Association.

American Psychological Association. Ethical principles of psychologists and code of conduct, 2017. URL https://www.apa.org/ethics/code.

American Psychological Association. Publication manual of the American Psychological Association: the official guide to APA style., 2020. Book Title: Publication manual of the American Psychological Association : the official guide to APA style. Edition: Seventh edition. ISBN: 9781433832154 Place: Washington, DC.

Mark Appelbaum, Harris Cooper, Rex B. Kline, Evan Mayo-Wilson, Arthur M. Nezu, and Stephen M. Rao. Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1):3–25, January 2018. ISSN 0003-066X. doi: 10.1037/amp0000191. URL https://ezproxy.spu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=psyh&AN=2018-00750-002&site=ehost-live. Publisher: American Psychological Association.

Jeffrey Beall. Predatory publishers are corrupting open access. *Nature*, 489(7415):179–179, September 2012. ISSN 1476-4687. doi: 10.1038/489179a. URL https://www.nature.com/articles/489179a. Number: 7415 Publisher: Nature Publishing Group.

Sonja Bezjak, April Clyburne-Sherin, Philipp Conzett, Pedro Fernandes, Edit Görögh, Kerstin Helbig, Bianca Kramer, Ignasi Labastida, Kyle Niemeyer, Fotis Psomopoulos, Tony Ross-Hellauer, René Schneider, Jon Tennant, Ellen Verbakel, Helene Brinken, and Lambert Heller. *Open Science Training Handbook*. Zenodo, April 2018. doi: 10.5281/ZENODO.1212496. URL https://zenodo.org/record/1212496.

Michael Bosnjak, Christian J. Fiebach, David Mellor, Stefanie Mueller, Daryl B. O'Connor, Frederick L. Oswald, and Rosemarie I. Sokol. A template for preregistration of quantitative research in psychology: Report of the joint psychological societies preregistration task force. *American Psychologist*, November 2021. ISSN 0003-066X. doi: 10.1037/amp0000879. URL https://ezproxy.spu.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=psyh&AN=2022-06923-001&site=ehost-live. Publisher: American Psychological Association.

Stephen J. Ceci and Elaine Walker. Private archives and public needs. *American Psychologist*, 38(4):414–423, April 1983. ISSN 0003-066X. doi: 10.1037/0003-066X.38.4.414. URL https://ezproxy.spu.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=psyh&AN=1984-02320-001&site=ehost-live. Publisher: American Psychological Association.

Christopher D. Chambers and Loukia Tzavella. The past, present and future of Registered Reports. *Nature Human Behaviour*, 6(1):29–42, January 2022. ISSN 2397-3374. doi: 10.1038/s41562-021-01193-7. URL https://www.nature.com/articles/s41562-021-01193-7. Number: 1 Publisher: Nature Publishing Group.

Daryl E. Chubin. Open Science and Closed Science: Tradeoffs in a Democracy. *Science, Technology, & Human Values*, 10(2):73–81, 1985. ISSN 0162-2439. URL http://www.jstor.org/stable/689511. Publisher: Sage Publications, Inc.

Ellen Driessen, Steven D. Hollon, Claudi L. H. Bockting, Pim Cuijpers, and Erick H. Turner. Does Publication Bias Inflate the Apparent Efficacy of Psychological Treatment for Major Depressive Disorder? A Systematic Review and Meta-Analysis of US National Institutes of Health-Funded Trials. *PLOS ONE*, 10(9):e0137864, September 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0137864. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0137864. Publisher: Public Library of Science.

Benedikt Fecher and Sascha Friesike. Open Science: One Term, Five Schools of Thought. *RatSWD Working Paper Series*, page 14, 2013.

Corporation for Digital Scholarship. Zotero | Your personal research assistant. 2022. URL https://www.zotero.org/.

Preregistration Task Force. Preregistration Standards for Psychology - the Psychological Research Preregistration-Quantitative (aka PRP-QUANT) Template. February 2021. URL https://www.psycharchives.org/en/item/088c79cb-237c-4545-a9e2-3616d6cc8453.

Katie Fortney and Linda Suk-Ling Murphy. Getting Found: Indexing and the Independent Open Access Journal. *Western Journal of Emergency Medicine*, 17(5):508–510, September 2016. ISSN 1936-900X. doi: 10.5811/westjem.2016.6.30836. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5017831/.

Martin S. Hagger. Developing an open science 'mindset'. *Health Psychology and Behavioral Medicine*, 10(1):1–21, December 2022. ISSN null. doi: 10.1080/21642850.2021.2012474. URL https://doi.org/10.1080/21642850.2021.2012474. Publisher: Routledge _eprint: https://doi.org/10.1080/21642850.2021.2012474.

Tamarinde Haven, Timothy M. Errington, Kristian Gleditsch, Leonie van Grootel, Alan M. Jacobs, Florian Kern, Rafael Piñeiro, Fernando Rosenblatt, and Lidwine Mokkink. Preregistering Qualitative Research: A Delphi Study. Technical report, SocArXiv, July 2020. URL https://osf.io/preprints/socarxiv/pz9jr/. type: article.

Michael P. Hengartner. Raising Awareness for the Replication Crisis in Clinical Psychology by Focusing on Inconsistencies in Psychotherapy Research: How Much Can We Rely on Published Findings from Efficacy Trials? *Frontiers in Psychology*, 9, February 2018. ISSN 1664-1078. doi: 10.3389/fpsyg.2018.00256. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5835722/.

The jamovi project. about - jamovi, 2021. URL https://www.jamovi.org.

David R. Krathwohl. *Methods of Educational and Social Science Research: The Logic of Methods*. Waveland Pr Inc, Long Grove, IL, 3 edition edition, March 2009. ISBN 978-1-57766-576-2.

Robert Madigan, Susan Johnson, and Patricia Linton. The language of psychology: APA style as epistemology. *American Psychologist*, 50(6):428–436, June 1995. ISSN 0003-066X. doi: 10.1037/0003-066X.50.6.428. URL https://ezproxy.spu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=psyh&AN=1995-31466-001&site=ehost-live. Publisher: American Psychological Association.

Maryann E. Martone, Alexander Garcia-Castro, and Gary R. VandenBos. Data sharing in psychology. *The American Psychologist*, 73(2):111–125, March 2018. ISSN 1935-990X. doi: 10.1037/amp0000242.

R. K. Merton. Science and technology in a democratic order. *Journal of Legal and Political Sociology*, 1:115–126, 1942.

Don A. Moore. Preregister if you want to. *American Psychologist*, 71(3):238–239, April 2016. ISSN 0003-066X. doi: 10.1037/a0040195. URL https://ezproxy.spu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=psyh&AN=2016-15883-005&site=ehost-live. Publisher: American Psychological Association.

Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*, 349 (6251):943–943, August 2015. ISSN 00368075. doi: 10.1126/science.aac4716.

OSF. Open Science Framework, 2022. URL https://osf.io/.

Prasad Patil, Roger D. Peng, and Jeffrey T. Leek. A statistical definition for reproducibility and replicability. Technical report, bioRxiv, July 2016. URL https://www.biorxiv.org/content/10.1101/066803v1. Section: New Results Type: article.

Robert Rosenthal. The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3):638–641, May 1979. ISSN 0033-2909. doi: 10.1037/0033-2909.86.3.638. URL https://ezproxy.spu.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=psyh&AN=1979-27602-001&site=ehost-live. Publisher: American Psychological Association.

Michael W. Ross, Martin Y. Iguchi, and Sangeeta Panicker. Ethical aspects of data sharing and research participant protections. *American Psychologist*, 73(2):138–145, February 2018. ISSN 0003-066X. doi: 10.1037/amp0000240. URL https://ezproxy.spu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip&db=psyh&AN=2018-06770-003&site=ehost-live. Place: US Publisher: American Psychological Association.

Tony Ross-Hellauer. What is open peer review? A systematic review. Technical Report 6:588, F1000Research, August 2017. URL https://f1000research.com/articles/6-588. Type: article.

Darshana Shah. Open Access Publishing: Pros, Cons, and Current Threats. *Marshall Journal of Medicine*, 3(3):1, July 2017. ISSN 2379-9536. doi: http://dx.doi.org/10.18590/mjm.2017.vol3.iss3.1. URL https://mds.marshall.edu/mjm/vol3/iss3/1.

Alex M. Shoemann, A. J. Boulton, and S. D. Short. Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science*, 8:379–386, 2017. URL https://schoemanna.shinyapps.io/mc_power_med/.

Arfon M. Smith, Daniel S. Katz, and Kyle E. Niemeyer. Software citation principles. *PeerJ Computer Science*, 2:e86, September 2016. ISSN 2376-5992. doi: 10.7717/peerj-cs.86. URL https://peerj.com/articles/cs-86. Publisher: PeerJ Inc.

Jeffrey R. Stevens. Replicability and Reproducibility in Comparative Psychology. *Frontiers in Psychology*, 8, 2017. ISSN 1664-1078. doi: 10.3389/fpsyg.2017.00862. URL https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00862/full. Publisher: Frontiers.

JASP Team. JASP (Version 0.16.3), 2022. URL https://jasp-stats.org/faq/.

Audrey Thompson. Gentlemanly Orthodoxy: Critical Race Feminism, Whiteness Theory, and the APA Manual. *Educational Theory*, 54(1):27–57, 2004. ISSN 1741-5446. doi: 10.1111/j.0013-2004.2004.00002.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0013-2004.2004.00002.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.0013-2004.2004.00002.x.

Anna Elisabeth van't Veer and Roger Giner-Sorolla. Pre-registration in social psychology—A discussion and suggested template. *Journal of Experimental Social Psychology*, 67:2–12, November 2016. ISSN 0022-1031. doi: 10.1016/j.jesp.2016.03.004. URL http://www.sciencedirect.com/science/article/pii/S0022103116301925.