

# 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

02 Aug 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](#).

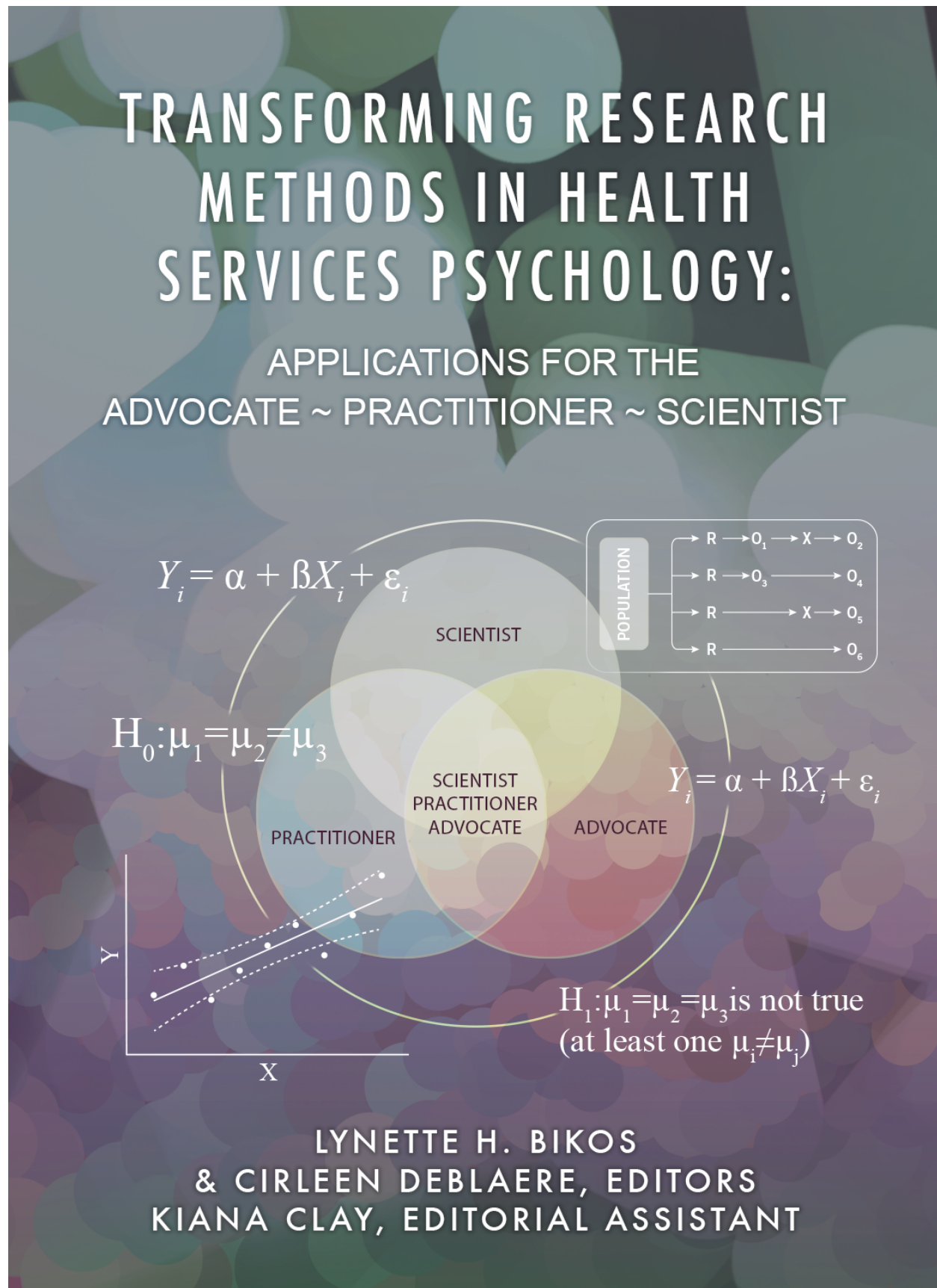


Figure 1: An image of the book cover. It includes three overlapping, pastel-colored, circles representing advocacy, practice, and science. These are surrounded by a handful of statistical symbols and formulae. methods.

# 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 02 Aug 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](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 to which 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 serves, 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 et 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” (?). Curiously, Chubin opened their article by referencing framework that is commonly used in introductory research methods textbooks (e.g., ?). That is: Merton’s norms of science.

Merton (?) argued that *communism* (i.e., common ownership 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 (?) 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 (?) 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 foreseen and unforeseen 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; ?) are foundational to science. Yet investigations across scientific disciplines have demonstrated significant failures to both. Through its large scale, collaborative effort, the Reproducibility Project (?) 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-analytic 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 (?). 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 (??).

Bias occurs at the journal level when reviewers and editors favor significant findings over non-significant ones (?). Driessen et al. (?) 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 (?), 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 (??). Stated another way, for research to be truly reproducible, the entire process must be open to scrutiny (?). 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 (?). 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 (?). 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 (??).

Scholars have noted a number of benefits to preregistration. The most formal preregistration is a registered report (?). 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 (?).

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 (?). 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 (??).

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 (?) 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 (?) 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 (?) 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 (?). The PRP-QUANT has three primary sections: an introduction, method, and analysis plan. Each section includes multiple items that are accompanied with brief instructions (?). For qualitative researchers, a Delphi method (?) provides a 13-item, preregistration template that is freely available at the [Open Science Framework](#). Both templates align with the Journal Article Reporting Standards[JARS; ?].

### 2.4.2 Data sharing

Data sharing (including the releasing of raw data, measures, codebooks, and analytic scripts for data cleaning and analysis; ?) 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 (?), (b) optimizes the value of data and will accelerate scientific progress when data are exposed to secondary analysis or combined in meta-analyses (?), and (c) creates a structure for greater collaboration (?). 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 (?). This was codified in the APA ethical principles as early as 1992. The current ethical principle (?) 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 (?).

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 (?). At the time of this writing, funding entities are increasingly requiring that data be made available for use by others (?) 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 (?).

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 (?). 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 (?). 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 (?). “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 (?). 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 (?).

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 (?). 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 (?). 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 (?).

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](http://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. (?) 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 (?).

### 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* (?). With open identities, neither the reviewers nor authors are anonymized. With open reports, the review reports are published alongside the relevant article (?). 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 (?).

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 (??). *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 (??) 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; ?]. Another positive aspect of open access journals is an expedited submission-to-publication timeline (?).

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 (?). 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 (?). [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 (??). 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 (?).

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 (??). 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 (?). *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 (?). The organization provides open-source (zero-cost) and commercial software that is intended to serve in a “virtuous cycle” for mutual improvement (?).

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 (?) Monte Carlo power analysis for indirect effects ([https://schoemanna.shinyapps.io/mc\\_power\\_med/](https://schoemanna.shinyapps.io/mc_power_med/)).

More comprehensive software tools are also built with R code. For example, the program *jamovi* (?) 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 (?) 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”, ?) 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



(?).

### 2.6.2 Reference management software

[Zotero](#) (?) 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 (?).

### 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 (?). 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 (?). 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 be less helpful in co-authoring statistical code.

GitHub (?) 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 ?. 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](#) (?) 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 (?).

### 2.6.6 Open education resources

Open education resources (OERs) include any tools, materials, or techniques that are used to support access to knowledge (?). 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 (?). 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

# Chapter 3

## APA Style

*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

There are many excellent resources for learning *APA Style*. This chapter is not one of them. Rather, the intent of this chapter is to locate APA Style within our discipline. In-so-doing, we hope to convey why it is valued and how adherence to the guidelines presented in the style manual can facilitate an effective and efficient transmission of ideas, information, and scientific findings. Simultaneously, we critically examine APA Style and question whether and how unquestioned allegiance to it may contribute to exclusion and oppression. While we do not *teach* APA style, we review the journal article reporting standards (JARS) by comparing the recommendations to a recently published article. We also review some of the sometimes surprising guidance that is offered in the Style Manual.

### 3.1 Learning Objectives

Learning objectives from this lecture include the following:

- List and define three cultural characteristics or values reflected in APA style.
- Identify two ways that Thompson (2004) has suggested that APA style perpetuates Whiteness and patriarchy in the academy.
- Describe how adherence to the JARS facilitates effective and efficient transmission of information.
- Evaluate a manuscript for JARS.

### 3.2 Readings & Resources

In preparing this chapter, we 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.
  - This resource is essential for writing many types of papers within the discipline of psychology. Not only does it provide instruction on reference lists, but also on how to write effectively, how to reduce bias in language, and how to structure research papers in ways that convey information efficiently and effectively.
- 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. (?) 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 (at the time of this writing we are using the 7th Edition) Thompson (?) pointed out how aspects of APA style contribute to preserving Whiteness.
- *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).
  - Huang, Y.-T., & Chan, R. C. H. (2022). Effects of sexual orientation concealment on well-being among sexual minorities: How and when does concealment hurt? *Journal of Counseling Psychology*. <https://doi.org/10.1037/cou0000623>
  - We use this article as a tool for understanding where and how the JARS requirements are implemented into a research article.

### 3.3 Abbreviations in this Chapter

In this chapter on APA Style we use three abbreviations. Because each represents a larger, proper title, we are capitalizing them as if they were proper nouns:

- The “Style Manual” refers, generally, to the *Publication Manual of the American Psychological Association* (all editions) and the cultural expectation of using its conventions.
- The “7th Edition” refers to the most recent edition of the *Publication Manual of the American Psychological Association* (?).
- “APA Style” refers more broadly to the culture and use of the Style Manual.

### 3.4 The Culture of APA Style

The *Publication Manual of the American Psychological Association* has grown from its 7-page writers guide in the *Psychological Bulletin* (1929) to 427 in the 7th Edition (?). Many who utilize

the Style Manual are looking up guidelines for formatting parts of a paper that requires APA style, locating examples of resources to learn how to cite it properly, or consulting the journal article reporting standards (JARS) to learn what information is required in an empirical manuscript. Madigan et al. (?) has suggested that within the individual who seeks to master APA Style and across the near-century that of the Style Manual's existence, that the Style Manual has served an epistemological function. That is, through learning and practicing the guidelines, it has instilled an set of attitudes and values, many of which are unarticulated and operate outside of our explicit awareness.

### 3.4.1 APA Style as Epistemology

Madigan (?) and colleagues compared APA style to that of other disciplines (e.g., history, literary criticism) and identified characteristics that define our writing. First, we use a story schema that consists of a prescribed introduction, method, results, and discussion. On the face of it, this seems straightforward – even linear. However research is messy and this may not be reflected well (or at all) in the final paper. Within this story arc, readers and reviewers who are familiar with APA Style can expect to see a (a) review of past empirical studies, (b) collection and analysis of new data, and (c) connections between conclusions and psychological theory.

A second characteristic is the depersonalized language of disagreement that focuses on the method, analyses, and conclusions and not on the investigators as individuals (?). An example might be a statement like, “The relatively small sample likely contributed to Type I error.” Madigan and colleagues suggested that a goal of psychological science is a collaborative, cumulative endeavor based on research data that avoids animosity and confrontation and that the culture of APA style facilitates this conveyance. Related to this the use of *hedged* conclusions. Hedged words like “tend,” “suggest,” and “may” are abundant in our writing. Further, we almost always claim that “more research is needed” before we can assert firm conclusions.

APA Style involves other people (?). We are more likely to have multiple authors (than single authors) on papers. In fact the 7th Edition instructs writers to list, by name, up to 20 authors in a reference list entry (when there are 21 or more, an ellipsis follows the first 19 and then lists the last author). Further, our introduction and discussion sections include abundant citations that locate the research within a proximal body of literature. Rather than footnotes or endnotes, we use in-text references that include the last names of authors; this provides immediate information about the “who” of the research.

Madigan (?) and colleagues also documented distinct writing conventions that define the culture of our writing. For example, we tend to use headings and subheadings rather than narrative transitions; we are more likely to paraphrase than use direct quotes; and we use self-effacing and low-profile language that “does not call attention to itself.” These approaches maintain the focus on the topic of the paper and not the personhood of the author.

### 3.4.2 APA Style, Whiteness, and Patriarchy

Nearly a decade after Madigan et al.'s paper (?), Thompson (?) examined APA style (i.e., the 5th edition) through the lens of critical race feminism and Whiteness Theory and argued that the expectations regarding clarity, precision, appropriateness, sensitivity, and objectivity likely contribute to the academy's investments in Whiteness and patriarchy. Thompson's analysis of

*power and property investments* was organized into five themes that continue to be relevant for the 7th Edition.

The first theme is *property rights*. Predominantly White institutions (PWIs) have treated refereed scholarship as intellectual property, demarcated with a certain, privileged status. That is, *individuals* “own knowledge.” Thompson (?) cited evidence for this in the conventions of:

- Using last names as a shorthand reference for work (e.g., “Thompson’s [1995] critique of APA style”).
- Citing one’s own work (e.g., “Bikos [2021] has argued...”) rather than writing, “When I previously claimed...”
- Using *et al.* for in-text citations when articles have three or more authors.
- Ordering authorship on the basis of relative contribution and not professional position, power, or hierarchy (see section 1.22 in the 7th Edition). Thompson (?) also suggested that power dynamics (especially around race and gender) likely interferes with adherence to this principle.

Thompson’s (?) concern about the proprietariness of knowledge for *individuals* is that it provides no recognition for the *communities* from where the the knowledge was gathered. After all, where is stylistic guidance for citing community or institutional knowledge? For example, data can be obtained from and about the Black Church, but there is no guidance from the Style Manual on how to cite such knowledge. There is only guidance on citing others who have also written about it.

A second theme is *precedent and pedigree*. Our disciplinary commitment to the notion that knowledge is cumulative is reflected in voluminous citations of relevant earlier works. These include citations of the gurus, citations of statements that may seem obvious, and self-citations. Including these citations in the introductory sections of a research paper locates the proposed study within a body of knowledge and provides the requisite justification for its inclusion within the lineage of the research topic. This practice disadvantages novel ideas and excludes ideas and projects that would challenge existing power structures.

*Proceduralism*, the third theme, refers to the constriction of communication via the standardized format and style (?). Regarding format, Thompson suggested that the Introduction, Method, Results and Discussion sections may be insufficient to capture the lived experiences of marginalized and oppressed groups. Regarding stylistic conventions, the Style Manual’s prescriptions for *elegant* writing and *unambiguous clarity* may not work for all groups, especially when information does not fit well into clean categories. The standardization of the *et al.* convention contributes to both gender- and color-blindness. On one hand, it is “fair.” On the other hand, the use of the first author’s surname obscures their identity and the identities of coauthors. Finally, APA Style discourages footnotes. We agree with Thompson’s claim that the most juicy arguments are always in the footnotes.

Following professional *protocol* means avoiding extremes and blending in. In this fourth theme, Thompson (?) notes that when one adheres to the protocol, they signal a recognition of the prevailing relations of power, authority, and legitimacy. The 7th Edition expanded what was previously a recommendation to avoid using insensitive and pejorative language, to an entire chapter on bias-free language. Thompson has indicated that such recommendations are admirable, but they do not address unequal power relations. Further, distinctions between what is insensitive and pejorative may be undetectable to those who hold privilege.

The fifth theme was the notion of a *gentleman's agreement*; that is, a “language that conveys professionalism and formality” where “differences should be presented in a professional and non-combative manner” (7th Edition, section 4.7). Thompson (?) 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.

### 3.4.3 APA Style and White Supremacy Culture

For more than 20 years, Tema Okun (??) and colleagues have curated a list of characteristics of White supremacy culture. The Centre for Community Organizations has prepared a handout that provides [definitions](#), descriptions, examples, and antidotes for each of the White supremacy culture characteristics as they occur in organizations. With Madigan (?) and Thompson's (?) conceptualizations in the foreground, we are using this chapter to ask the question, “How might APA Style reflect and further contribute to White Supremacy Culture?”

Before we move onto an deeper investigation of some of the essentials of APA Style, we encourage you to follow the link to the [definitions](#) and review them. Then, as we present the information of APA Style, note where you notice overlap. We will return to this examination at the end of the chapter.

- 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

### 3.5 The JARS: The Core of APA Style

The JARS [Journal Article Reporting Standards; ?] were born out of a concern for transparency in psychological science. When researchers present their work using these comprehensive, uniform reporting standards, it makes it easier for readers and reviewers to work their way through individual papers, compare research, and use the results in meta-analyses (?). Thus, the JARS provides an interconnected set of tables that list information that should be reported in prescribed sections of empirical manuscripts. The JARS does not specify in what order that the information should be presented – just that it should be available to the reader at the time they need it, and written such that does not interfere with the readability of the paper.

The JARS were first introduced in a 2008 feature in the *American Psychologist* (?) and were included in the 6th Edition of the Style Manual. An updated JARS, published in 2018, expanded the types of quantitative research (JARS-Quant) as well as introduced tables for qualitative (JARS-Qual) and mixed methods studies (JARS-Mixed). Chapter 3 of the 7th Edition is devoted to the JARS. It contains numerous tables, definitions, explanations, and a flowchart.

Curiously, the second chapter of the 7th Edition provides another overview of each section of an APA Style manuscript. The information provided in the second chapter has a high degree of overlap and consistency with the JARS but focuses more on formatting conventions. We encourage authors to consult both.

A third set of guidelines for authors are found in individual journals. These guidelines may have inconsistencies with APA Style. Additionally, journals, themselves will have their own author guidelines. When submitting to a specific journal, it is that journal's guidelines that likely take precedent.

Given that the JARS and Chapter 2 recommendations, should be directly consulted when drafting an APA Style paper, we will not repeat them here. We thought, however, that it might be a useful exercise to review a recently published article, mapping if and where each of the JARS elements are included. We have chosen Yu-Te Huang and Randolph C. H. Chan's (2022), "Effects of Sexual Orientation Concealment on Well-Being Among Sexual Minorities: How and When Does Concealment Hurt?" published in the *Journal of Counseling Psychology*. We strongly encourage you to work along with us by accessing a copy of the JARS either in the 7th Edition or [online](#) and identifying (by highlighting) the elements in [Huang and Chan's paper](#)

Utilization of the JARS should begin with its [flowchart](#). Huang and Chan's (?) research is quantitative in nature. The flowchart indicates that all quantitative studies should start with JARS-Quant Table 1. Because the study did not involve an experimental manipulation, the write-up should also adhere to JARS-Quant Table 3. In the table below, we paraphrase the JARS instructions, provide a quick summary of if-and-how Huang and Chan addressed it, and then provide a "+", "-", or "+/-" designation to indicate if it was present, absent, or partially addressed.

**Table 1**

JARS Element	Location in Huang and Chan (2022)	+/-
<b>Title</b>	<b>Huang and Chan (2022)</b>	+/-



JARS Element	Location in Huang and Chan (2022)	+/-
Main variables and theoretical issues and their relationship	Primary IV (concealment) and DV (well-being) were named. Asking “how” and “when” suggests that mechanisms (mediation, moderation) will be evaluated	+
Population(s) studied	No indication that the participants were from Hong Kong	-
<b>Author Note</b>	<b>Huang and Chan (2022)</b>	+/-
Acknowledgement and explanation of special circumstances (e.g., preregistration, use of data in other publications, if primary data is used in dissertation or conference papers, sources of funding/support, potential conflicts of interest, affiliation of authors if different than author byline, contact information for corresponding author, other relevant information)	Provides ORCiDs. Identifies funding source. Addresses prior presentations, publications, conflicts of interest. Lists contact information.	+
<b>Abstract</b> (120 to 250 words)	<b>Huang and Chan (2022)</b>	+/-
Problem under investigation including primary hypotheses	Defines problem under investigation. Although not specified as such, the hypotheses are reflected in an efficient presentation of results.	+
Key participant characteristics	$N = 737$ sexual minority individuals. Lists gender and age.	+
Method: research design, sample size, materials used, outcome measures, data-gathering procedures	Participants recruited through LGBTQ networks for baseline and 1-year follow-up	+
Findings: primary statistical output	Results are narrated, but no statistical output is included	-
Conclusions: results plus implications or applications	Results pointed to the importance of family and loneliness	+
<b>Introduction</b>	<b>Huang and Chan (2022)</b>	+/-
Problem statement connecting to theoretical and/or practical implications	Opening paragraph provides a problem statement. Project is grounded in self-determination theory.	+
Succinct review of relevant scholarship – connecting to extant literature	Organized into subheadings with a distinction between identity disclosure/nondisclosure and concealment.	+
Hypotheses (primary and secondary), aims, and objectives, including theoretical derivation and additional planned analyses	Three hypotheses plus and exploratory question were listed.	+
<b>Method</b>	<b>Huang and Chan (2022)</b>	+/-
Inclusion and exclusion criteria	Listed in <i>Participants</i> section	+

JARS Element	Location in Huang and Chan (2022)	+/-
Participant characteristics: standard demographic and topic-specific	<i>Participants</i> section includes summary of general (e.g., age, education) and topic-specific variables.	+
Sampling procedures addressing the following (if used): sampling plan, response rate, self-selection, settings and dates, participant agreements/compensation, and IRB and safety monitoring.	<i>Procedures</i> section describes recruitment platforms, voluntary nature, consent process. <i>Participants</i> section describes attrition across the two waves.	+
Sample size, power, and precision: intended and achieved sample size, power analysis, use of interim analyses or stopping rules	<i>Participants</i> section identifies how the minimum sample size was determined.	+
Measures and covariates: all primary and secondary measures and covariates (including those not utilized in report)	<i>Measures</i> section lists the variables and tools used to measure them. <i>Preliminary analysis</i> lists covariates (demographic variables); the tools for assessing them were not described. There is no mention of additional, unused data.	+/-
Data collection: methods used to collect data	<i>Procedures</i> indicates that an online survey platform was used and describe the invitation, consent, follow-up protocol.	+
Quality of measurements: Methods used to enhance quality of measurements including training and reliability of data collectors and use of multiple observations	NA since surveys collected online.	NA
Instrumentation: review of psychometric and biometric properties	<i>Measures</i> provides descriptions of the instruments used, including citations related to their psychometric properties, and study-specific internal consistency coefficients for four of the five instruments.	+
Masking: In experimental manipulations, describe and evaluate presence of masking techniques (if used)	NA	NA
Psychometrics: reliability and validity estimates for instruments within the reported study and related studies or test manuals	<i>Measures</i> provides descriptions of the instruments used, including citations related to their psychometric properties, and study-specific internal consistency coefficients for four of the five instruments.	+
Conditions and design: since this is non-experimental, it directs to JARS Table 3	NA; a quick review of Table suggests that meeting the criteria has already been addressed	NA

JARS Element	Location in Huang and Chan (2022)	+/-
Data diagnostics: planned data diagnostics including criteria for post-data-collection exclusion, description of managing missing data, definition and processing of statistical outliers, analyses of distributional characteristics, data transformations (if used)	<i>Results/Preliminary Analysis</i> reviews the screening for normality and provides descriptive and correlational results. Independent <i>t</i> tests and one-way ANOVAS were used to identify differences in the primary variables and in missingness as a function of demographic variables. MNAR (missing not at random) was assumed and FIML (full information maximum likelihood) was used in primary analyses.	+
Analytic strategy: for primary, secondary, and exploratory hypothesis; include plans to protect against experiment-wise error	There was not a separate section in the <i>Method</i> to first outline the analytic strategy. Rather, detailed descriptions of the analytic process were presented in the <i>Results</i>	+/-
<b>Results</b>	<b>Huang and Chan (2022)</b>	+/-
Participant flow: Include <i>n</i> of each group at each stage of study; consider including a figure	<i>Method/Participants</i> details the <i>N</i> from initial interest through the second wave.	+
Recruitment: Dates defining periods of recruitment and follow-up	A timeline regarding Time1 and Time 2 was provided; dates of collection were not provided.	+/-
Statistics and data analysis: the JARS provides detailed information about what to include regarding missing data and analysis, descriptive data, inferential statistics, and complex statistics; additionally, problems with statistical assumptions, data distributions, convergence are to be reported	Subsections in the <i>Results</i> organize the primary and exploratory hypotheses. While they did not review plans to protect against experiment-wise error, they reported <i>B</i> weights, confidence intervals, and fit statistics.	+/-
<b>Discussion</b>	<b>Huang and Chan (2022)</b>	+/-
Support (or nonsupport) of original hypotheses: distinguish between primary, secondary, exploratory; consider implication of exploratory as it relates to error rates	Although the opening paragraphs of the <i>Discussion</i> summarize the findings, they do not identify them as hypotheses or exploration.	+/-
Similarity of results to findings and work of others	The mid-section of the <i>Discussion</i> compares findings to prior literature.	+
Interpretation of results, considering validity threats, sample size, measurement protocols	<i>Study Limitations</i> considers threats to internal validity and the measures that were used.	+
Generalizability taking into consideration target population and contextual issues	<i>Study Limitations</i> considers threats to external validity, particularly around self-selection into the study.	+

JARS Element	Location in Huang and Chan (2022)	+/-
Implications for future research, program, policy	<i>Practical Implications</i> provide guidance to counseling psychologists when working with sexual minority clients.	+

## 3.6 What the Style Manual (?) Has to Say About...

It can be surprising to learn that the Style Manual addresses issues beyond the order, content, and conventions of a manuscript. With explicit reference to the APA Ethical Principles of Psychologists and Code of Conduct [APA Ethics Code; ?], the very first chapter of the 7th Edition includes a review ethical, legal, and professional standards in publishing. Chapter 1 in the 7th Edition reviews the importance of ensuring the accuracy of scientific findings, protecting the rights and welfare of research participants, and protecting intellectual property rights. Given that most of these topics are covered elsewhere in this OER, we wanted to review only one issue, *protecting intellectual property rights*, because it addresses a sometimes sticky issue: publication credit and authorship order.

### 3.6.1 Authorship

Both the 7th Edition (section 1.21) and the APA Ethics Code (standard 8.12a) state that authorship is reserved for those who make a “substantial contribution.” The list of contributions that are *substantial* include formulating the research question, designing the research project, organizing and conducting the analysis, or interpreting the findings and result. Lesser contributions (e.g., entering data, recruiting subjects) on their own, do not garner authorship, but might warrant authorship if the potential author was involved in a number of these tasks. Once the list of authors is determined, the next task is to determine their order.

The 7th Edition (section 1.22) states that the general rule for determining author order is that the principal contributor appears first, with subsequent names appearing in order of decreasing contribution. There are some traditions where the principle investigator appears last. When authors play equal roles, this can be stated in the author note.

The Style Manual acknowledges that hierarchy has the potential to complicate the determination of authorship and its order. Relative status (e.g., department leader, tenure, promotion, student) should not determine the order of authorship. Further, given that a student’s dissertation is considered to be an independent project, it should be rare that a faculty member holds the position of first author. This can happen, though, when the dissertation is part of a collection of studies that are published together, the student uses data from an existing research project sponsored by the faculty member, and/or the faculty member (with the students’ permission) engages in post-dissertation reanalysis/rewriting to ensure that the work is published.

Academia can be fiercely competitive and hurt feelings (in fact, broken relationships) over issues like authorship are common. In the preface of their book on longitudinal modeling, a beautiful story is told by Judith Singer and John Willett (?). The authors met at Harvard in 1985 when they had competed for a single position but an unexpected vacancy led to them both being hired. Their colleagues expected them to be competitive. Early in their careers they decided to be intentionally

collaborative. This involved never divulging who wrote what part of a paper, inviting the other when one was invited to speak/present, and never competing for opportunities. Part of their strategy was to include this statement in every jointly published work: “The order of the authors has been determined by randomization.”

Given that the collaboration evidenced by Singer and Willett (?) is unique, other strategies may be used. The Style Manual recommends that authors talk as soon as practical to establish the *process and criteria* for determining the author order. We have linked a [rubrics](#) that is sometimes used in research teams to help make these decisions.

### 3.6.2 Writing Effectively

There are a number of popular resources for writing scholarly, scientific, and empirical papers. The content of the fourth chapter of the 7th Edition provides guidelines for achieving the hallmarks of APA Style: continuity, flow, conciseness, and clarity. The topics within Chapter 4 represent suggestions for transitioning more casual writing (e.g., wordiness, anthropomorphisms, subject-verb agreement) to be more clear, precise, and efficient.

The last section of the 7th Edition’s fourth chapter reviews broader strategies to improve writing. We highlight two of these examples. Section 4.25 recommends the vicarious strategy of *reading*. We find it extremely valuable to find one or more articles, from a high quality journal, that is similar or parallel to what is being written or created. After reviewing the paper for content, examine it from a *meta* perspective, asking the questions that are important to you. For example:

- At what breadth or specificity does the paper open?
- What heading/subheadings are used? How are transitions made?
- How are the hypotheses identified – numbered? narrated?
- Is there a format for presenting statistical output?
- What is presented in the text versus a table?
- How does the paper end (e.g., summary vs. call-to-action)?

The 7th Edition’s Chapter 4 also recommends creating an outline. A well-structured outline should guide a reader/reviewer through a paper, even if the content is missing. To demonstrate this point, we present the headings of our example article by Huang and Chan (?). As you examine it, note how the content of the headings and subheadings in the introduction map onto the method and results.

#### **Effects of Sexual Orientation Concealment on Well-Being Among Sexual Minorities: How and When Does Concealment Hurt?**

- The Mechanisms of Sexual Orientation Concealment
  - LGB-Specific Authenticity
  - Loneliness
- The Moderating Effect of Inclusive Environments
- The Present Study
- Method

- Participants
- Procedure
- Measures
  - \* Sexual Orientation Concealment
  - \* Perceived LGBT-Friendliness
  - \* LGB-Specific Authenticity
  - \* Loneliness
  - \* Well-Being
- Results
  - Preliminary Analysis
  - The Effect of Sexual Orientation Concealment on Well-Being
  - Mediating Effects of LGB-Specific Authenticity and Loneliness
  - Moderating Effects of Perceived LGBT-Friendliness
  - Moderating Effects of Perceived LGBT-Friendliness of the Family (Model 2)
  - Moderating Effects of Perceived LGBT-Friendliness of the Peer Networks (Model 3)
  - Moderating Effects of Perceived LGBT-Friendliness of the Workplace/School (Model 4)
  - Moderating Effects of Perceived LGBT-Friendliness of the Wider Society (Model 5)
- Discussion
  - Practical Implications
  - Study Limitations
  - Conclusions
- References

In many research methods classes, tutorials, and books on writing, trainees are instructed to write an outline. In our own program our graduate students are encouraged to write a “mangy cat outline,” where “some part of the outline will have more fur (details) than others.” Writers can expect to revise, update, and restructure the outline as they work through all the steps of the research process.

Elsewhere, this OER’s chapter on [open science](#) describes the value that preregistration of a research project can have on transparency of science. We also suggest that in lieu of (or in addition to) outlining the paper, completing a preregistration may also serve as a tool for collecting and organizing ideas that can be a starting place for drafting the journal article.

### 3.6.3 Reducing Bias

New to the 7th Edition is the content of the fifth chapter, “Bias-Free Language Guidelines.” Accompanying this chapter is a strong message that the guidelines for bias-free *language* (note the use of *language* and not just *writing*) will be forever-evolving. We highlight a few general guidelines:

- Call people what they call themselves (especially in qualitative or participant-informed research designs).

- Simultaneously, recognize that no group is monolithic; remain attentive to within-group differences. For example, 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.
- Accept that language changes with time; remain attentive and flexible.
- “White,” “heteronormative,” and “U.S.” is o
  - When 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.
- In describing any population, provide a summary of *relevant* characteristics. It is unnecessary to collect and report all personal characteristics (particularly if they are distal to the focal topic).
- Describe actual differences between the target and general populations clearly and professionally. Do not be euphemistic.
- There are specific guidelines for level of specificity in labels/language for various populations (e.g., disability, age, gender identity, race/ethnicity/nationality, sexual orientation, SES). When writing about a topic or population, take time to look up the current guidelines, again.

The chapter had extensive and specific guidelines for reducing bias by age, disability, gender, research participants, racial and ethnic identity, sexual orientation, socioeconomic status, and the recognition of intersectionality. In keeping with the opening theme, the guidelines continue to evolve.

### 3.6.4 The Rules of Style

The sixth chapter of the 7th Edition provides the guidelines that bring uniformity to papers. Among so many other things, we learn to:

- use the Oxford comma,
- use single quotation marks only within double quotation marks,
- use lowercase for diseases, theories, therapies, or statistical procedures, and
- space once after punctuation.

The chapter also includes a very helpful table that provides properly formatted statistical abbreviations and symbols. The seventh chapter contains detailed instruction for creating tables and figures.

### 3.6.5 References

Chapters 8 through 11 of the 7th Edition address in-text citations and the reference list. These are the chapters are are probably most consulted in the Style Manual. What may be surprising is the broader information provided about (a) the degree to which material should be cited, (b) plagiarism, (c) self-plagiarism, and (d) paraphrases versus direct quotations. In the era of open science, there is also a section addressing “which” version of any document should be cited when there are multiple versions (e.g., pre-print, online first, print version)available. *The answer to this question is: cite the version used.*

Authors of the 7th Edition have continued to provide numerous examples and variations for users to emulate. They have also invested more time in creating a reference list entry when no match is perfect. The 7th Edition indicates that the general format of the reference is: Author (Date). Title. Source.

Each element of the reference entry is associated with a question:

- Author: Who is responsible for the work?
- Date: When was the work published?
- Title: What is the work called?
- Source: Where can I retrieve the work?

Table 9.1 in the 7th Edition provides a chart for creating a reference when information is missing.

The text-citations and reference lists must be 100% consistent. If you have consistently used a reference management system (e.g., Zotero, Mendeley) and each reference entry is accurate and complete, this will happen rather magically. Even so, one of the last things we recommend is a text-cite/ref-list cross check. We open two copies of the manuscript and put them side-by-side on a computer monitor. Starting with the text (using track changes), each time we come to a reference, we put an X in front of it the reference list entry. We go through the entire document. If there are items in the reference list missing, we note and add them. If there are additional references in the reference list that we have not checked, we do a quick search to see if they are really missing, and if so, we delete them. In addition to ensuring text-cite/ref-list consistency, this type of review often reveals other issues that should be corrected prior to submission or dissemination.

### 3.6.6 Publication

The 12th chapter of the 7th Edition provides an overview of the publication process including preparing publication, understanding the editorial publication process, preparing the manuscript, and copyright and permission guidelines.

Especially helpful in the 12th chapter is the overview of the publication process and the descriptions of the roles of the decision-makers: the editor-in-chief, associate editors, consulting or advisory editors, and ad hoc reviewers. Also helpful is the description and options associated with the editorial adjudications: acceptance, revise-and-resubmit, and rejection.

## 3.7 Revisiting the Culture of APA Style

As a reader, you may correctly presume that if we have taken the time to review core elements of APA Style, that we must value it. We do. A manuscript or article that is organized predictably and contains the expected information in the assigned places makes it much more efficient for me to evaluate and utilize. As the first author (Bikos), I have spent decades learning APA Style, and as Madigan (?) predicted, I have instilled many of its values. I especially love writing that is clear, precise, and efficient.

Critical consciousness requires the identification and interrogation of all the systems in which we operate. Especially and including those things we love. Earlier in this chapter we encouraged the reader to review Okun's (?) characteristics of White Supremacy Culture and to make note of places



where APA Style reflects White Supremacy Culture. Below, we have italicized the elements that noticed in APA Style. We have added a few notes describing our observations.

Below we have again listed the characteristics of White Supremacy Culture. This time we have added brief descriptions when we have observed overlap.

- Perfectionism
  - *Perfectionistic culture*
    - \* The level of negotiable detail surrounding research designs, choice of statistics and interpretation, and writing up the results – combined with the culture of peer review – means that the perfect study is never achieved. While this cautiousness offers strength in the form of humility, it may also contribute to a self-deprecation and doubt that stalls works from being completed and prevents authors from moving their scientific findings into advocacy and practice.
  - *Worship of the written word*
    - \* The entire Style Manual is about “the written word.”
    - \* Of the 7th Edition’s 427 pages, 310 (73%) are devoted to citing other work.
    - \* Abundant citation of other works is expected, especially in the Introduction, Method, and Results sections.
  - *Only one right way*
    - \* The Style Manual provides explicit instructions for everything from the formatting of statistical symbols to the use of commas. Many of these conventions are either “correct” or “incorrect.”
  - *Either/or thinking*
    - \* We are not thinking of examples of this in APA Style. With the recognition of philosophies of science including constructivist-interpretivist and critical-ideological and attention to the way that confounds and error can be introduced into a study, we do not see this as characteristic of the profession or APA Style.
- Concentration of power
  - *Power hoarding*
    - \* As described in the 7th Edition’s 12th chapter, the publication process is quite hierarchical, with ultimate decisions relying on layers of peer reviewers and decision-makers. Additionally, power is earned and retained through the requirement to locate new studies in the context of the prior literature through citations. Reviewers expect that the established citations will be present. In turn, citations are directly related to impact factors and, in turn, tenure, promotion, prestige, and status.
- Paternalism
  - *Defensiveness*
- Right to comfort
  - *Fear of open conflict*
    - \* As we noted earlier, Madigan (?) described APA Style has one that was tentative and cautious. That is, authors *hedge* with hesitant language and statements that

“more confirmation is needed.” Through the use of techniques like using first person (section 4.16) and the promotion of an active voice (section 4.13), the 7th Edition is promoting a more direct style of writing.

- \* In the publication process, peer reviewers provide their impressions to the editor, the editor summarizes the feedback and renders a decision to the author(s). The correspondence between all parties tends to be very professional and perhaps a bit ingratiating. Combative tones are discouraged.
- Individualism
  - *I’m the only one*
    - \* Although not necessarily communicated in the Style Manual, *replication* studies have been avoided because they fail to make a unique contribution to the literature. As described in the chapter on *open science*, this has likely contributed to the replicability/reproducibility crisis, in that if replications failures had been observed sooner, then researchers could begin closer investigations into why.
- Progress is more/bigger
  - *Objectivity*
    - \* The *precision*, *clarity*, and *efficiency* goals of APA Style are intended to provide the assurance of objectivity. Even in qualitative studies, researchers provide disclosure/transparency statements to assure the reviewers and readers that they have approached their topic in a manner that is unbiased and neutral.
  - *Quantity over quality*
    - \* Common among the curriculum vita of academics is the numbering of scholarly products so that they can be listed in applications for promotion and tenure. Additionally, many institutions require additional data about the impact that faculty work has had; the most common being “times cited” as well as the impact factor of the journal. This latter criteria, though, is a reflection of quality.
  - *Sense of urgency*
    - \* Historically, the publication process has not been fast. With competition, though, from open access journals (where the time from submission to publication can be a matter of days or weeks), we have observed top tier journals shortening the time-frame that volunteer peer reviewers are given to review and submit their reviews.

Our review suggests that APA Style includes some characteristics of White Supremacy Culture. Subsequently, we might ask, “How do these serve to maintain the power structures within psychology (or academia, or scholarship)?” Having been involved in international psychology and on a number of journal review boards, I (Bikos), have peer-reviewed a number of manuscripts. International scholars, for whom English is a foreign language, are one group that face significant challenges in publishing their research (?) – even when the science is meritorious.

*International Perspectives in Psychology* is one journal that prioritized disseminating work from global scholars. One response to that was an *editorial mentoring* program. Unfortunately, only three projects were completed through that program before it was deemed unsustainable and discontinued (?). The multi-layered requirements of APA Style were some of the significant challenges

### 3.8 Suggestions for Practice, Further Learning, and/or Conversation

1. Using a journal article that is of interest to you, create an outline from the headings and subheadings. Consider how well these guide the reader/reviewer through the research project.
2. Audit the same journal article using the empty table provided below. Make notes of if, how, and where the article followed the JARS guidelines.

#### 3.8.1 Empty Table for JARS Audit Activity

**Table 1**

Element	Your Article	+/-
<b>Title</b>	<b>Your Article</b>	+/-
Main variables and theoretical issues and their relationship		
Population(s) studied		
<b>Author Note</b>	<b>Your Article</b>	+/-
Acknowledgement and explanation of special circumstances (e.g., preregistration, use of data in other publications, if primary data is used in dissertation or conference papers, sources of funding/support, potential conflicts of interest, affiliation of authors if different than author byline, contact information for corresponding author, other relevant information)		
<b>Abstract</b> (120 to 250 words)	<b>Your Article</b>	+/-
Problem under investigation including primary hypotheses		
Key participant characteristics		
Method: including research design, sample size, materials used, outcome measures, data-gathering procedures		
Findings: primary statistical output		
Conclusions: results plus implications or applications		
<b>Introduction</b>	<b>Your Article</b>	+/-
Problem statement, including theoretical and/or practical implications		
Succinct review of relevant scholarship – connecting to extant literature		

Element	Your Article	+/-
Hypotheses (primary and secondary), aims, and objectives, including theoretical derivation and additional planned analyses		
<b>Method</b>	<b>Your Article</b>	<b>+/-</b>
Inclusion and exclusion criteria		+
Participant characteristics: standard demographic and topic-specific		
Sampling procedures addressing the following (if used): sampling plan, response rate, self-selection, settings and dates, participant agreements/compensation, and IRB and safety monitoring.		
Sample size, power, and precision: intended and achieved sample size, power analysis, use of interim analyses or stopping rules		
Measures and covariates: all primary and secondary measures and covariates (including those not utilized in report)		
Data collection: methods used to collect data		
Quality of measurements: Methods used to enhance quality of measurements including training and reliability of data collectors and use of multiple observations		
Instrumentation: review of psychometric and biometric properties		
Masking: In experimental manipulations, describe and evaluate presence of masking techniques (if used)		
Psychometrics: reliability and validity estimates for instruments within the reported study and related studies or test manuals		
Conditions and design: May need to consult supplemental tables		
Data diagnostics: planned data diagnostics including criteria for post-data-collection exclusion, description of managing missing data, definition and processing of statistical outliers, analyses of distributional characteristics, data transformations (if used)		
Analytic strategy: for primary, secondary, and exploratory hypothesis; include plans to protect against experiment-wise error		
<b>Results</b>	<b>Your Article</b>	<b>+/-</b>

Element	Your Article	+/-
Participant flow: Include $n$ of each group at each stage of study; consider including a figure		
Recruitment: Dates defining periods of recruitment and follow-up		
Statistics and data analysis: the JARS provides detailed information about what to include about missing data and analysis, descriptive data, inferential statistics, and complex statistics; problems with statistical assumptions, data distributions, convergence are to be reported		
<b>Discussion</b>	<b>Your Article</b>	<b>+/-</b>
Support (or nonsupport) of original hypotheses: distinguish between primary, secondary, exploratory; consider implication of exploratory as it relates to error rates		
Similarity of results to findings and work of others	T	
Interpretation of results, considering validity threats, sample size, measurement protocols		
Generalizability taking into consideration target population and contextual issues		
Implications for future research, program, policy		

### 3.9 References