



---

# **Week 14 Lecture - Replication and Transparency**

Undergraduate Research Methods in Psychology

---

Quinton Quagliano, M.S., C.S.P

Department of Psychology

## Table of Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Chapter Overview</b>                              | <b>2</b>  |
| 1.1      | Learning Objectives . . . . .                        | 2         |
| 1.2      | Chapter Overview . . . . .                           | 2         |
| <b>2</b> | <b>Replication</b>                                   | <b>2</b>  |
| 2.1      | Overview . . . . .                                   | 2         |
| 2.2      | Types of Replication . . . . .                       | 3         |
| 2.2.1    | Direct . . . . .                                     | 3         |
| 2.2.2    | Conceptual . . . . .                                 | 4         |
| 2.2.3    | Replication and Extension . . . . .                  | 4         |
| 2.3      | Real Projects . . . . .                              | 5         |
| 2.3.1    | One Study, Many Labs . . . . .                       | 5         |
| 2.3.2    | Many Studies, Many Labs . . . . .                    | 5         |
| 2.3.3    | When It Doesn't Work . . . . .                       | 6         |
| 2.4      | Meta-analysis . . . . .                              | 6         |
| 2.4.1    | Strengths and Limitations . . . . .                  | 6         |
| 2.5      | In Popular Culture . . . . .                         | 7         |
| <b>3</b> | <b>Transparency and Credibility</b>                  | <b>7</b>  |
| 3.1      | Overview . . . . .                                   | 7         |
| 3.2      | Questionable Practices . . . . .                     | 8         |
| 3.2.1    | Under-report Null Findings . . . . .                 | 8         |
| 3.2.2    | HARKing . . . . .                                    | 8         |
| 3.2.3    | P-hacking . . . . .                                  | 9         |
| 3.3      | Transparent Practices . . . . .                      | 9         |
| 3.3.1    | Open Data and Open Materials . . . . .               | 9         |
| 3.3.2    | Preregistration . . . . .                            | 10        |
| <b>4</b> | <b>When Do We Need (Really) External Validity</b>    | <b>10</b> |
| 4.1      | Overview . . . . .                                   | 10        |
| 4.2      | Generalization to Other People . . . . .             | 10        |
| 4.3      | Generalization to Other Settings . . . . .           | 11        |
| 4.4      | Do We Always Need to Generalize? . . . . .           | 11        |
| 4.4.1    | Theory-testing Mode . . . . .                        | 11        |
| 4.4.2    | Generalization Mode . . . . .                        | 12        |
| 4.5      | Do We Always Need to be in the Real World? . . . . . | 12        |

# 1 Chapter Overview

## 1.1 Learning Objectives

- Explain why it is essential for a study to be replicated.
- Describe why transparent research practices help ensure credible science.
- Evaluate, in a nuanced way, a study's quality in the context of scientific progress and external validity
- Appreciate science as an ongoing and continuous process of improvement, guided by replication

## 1.2 Chapter Overview

- As we've learned over the course of this class, \_\_\_\_\_ of scientific work can look/work a lot of different ways.
  - This is often different based on the \_\_\_\_\_ of the study as well - experiments v. small-n v. bivariate, etc.
- In short, valid studies are those that provide enough \_\_\_\_\_ that a finding represents a real effect, change, or description - not some fluke.
  - Said fluke could come from a failure in any of our 4 \_\_\_\_\_
- One of the best ways to \_\_\_\_\_ our evidence is to demonstrate that our findings can be produced again under the same or similar conditions.

# 2 Replication

## 2.1 Overview

? In which validity have we previously described reproducibility?

- A) Statistical
- B) Internal
- C) External
- D) Construct

Explanation:


- Replication/reproducibility is where we determine whether a finding can be reproduced or if it appears to be a \_\_\_\_\_
  - It can be viewed as a general indicator of a \_\_\_\_\_ quality study
- *Recall:* Our findings are often judged against  $\alpha = 0.05$ , which means we do leave some room for “random \_\_\_\_\_.”
  - By that logic, our results could simply be \_\_\_\_\_ due to chance
- Often, we evaluate \_\_\_\_\_ as being of core importance to a study’s value - because we want to know if the results tell a consistent story
  - Remember that we look for \_\_\_\_\_ in the evidence body, not just one-off studies

## 2.2 Types of Replication

- “Replication” is a fairly loose term, and actually applies to many scientific \_\_\_\_\_
  - You have likely run across \_\_\_\_\_ studies while you looked for articles on your research proposal.
- In a replication study, our \_\_\_\_\_ is not to refute or support, our motivation is to honestly explore and assess the phenomenon
  - “Trust but verify”
- We may separate study replications into a few different types:

### 2.2.1 Direct

- This is the \_\_\_\_\_ type of replication, where we attempt to largely keep all factors, measures, and circumstances consistent with the original study.
  - This is effectively following the \_\_\_\_\_ described in a journal article *exactly* as it is written.

 Discuss: What issues could make a direct replication study harder to pull off?


- It is, of course, done on a new \_\_\_\_\_, but from the same theoretical population.
- Design issues are \_\_\_\_\_ forward in direct replications, as we attempt to stay largely the same (even when that may be a bad thing)
- *Example:* Imagine if I re-did the Harlow Monkey study *exactly* the same way Harlow did it.

### 2.2.2 Conceptual

- Our central hypothesis stays the same as the \_\_\_\_\_ study, but we may change minor procedures, designs, and how we operationalize our latent constructs.
  - But, in a conceptual replication, we will not \_\_\_\_\_ anything new.
- This can be a useful way to \_\_\_\_\_ some glaring validity issues in an original study and do a more refined investigation
- However, it should be viewed critically as a \_\_\_\_\_ version of the original, so corrections to the original's effects may not be as clear
- *Example:* Harlow's Monkey's again, but this time I've changed the food with the wire mother to be more appealing and I have turned up the heating element on the felt mother just a little bit more

### 2.2.3 Replication and Extension

- In this type, we add \_\_\_\_\_ conditions or variables to add more nuance and context to findings.
- Such procedures could reveal \_\_\_\_\_ and mediators, or just present other relevant outcomes for convergence
- Like any replication, this type may offer evidence to \_\_\_\_\_ the findings of the original study, and also add their own new information at the same time

 Discuss: Do you think your research proposal is one of the 3 types of replications? If so, which one?

- *Examples:* Harlow's Monkeys, but instead, I am also going to add baby gorillas in to see if their behavior differs.

## 2.3 Real Projects


- A replication effort/project can look different ways, but is almost always very time-consuming and \_\_\_\_\_.
- It is beneficial to have multiple \_\_\_\_\_ and sets of researchers replicate an effect to reduce the chance that a single experimenter's bias affects results.

### 2.3.1 One Study, Many Labs

- In this scenario, one research lab "leads" numerous other labs in applying replications to a single central \_\_\_\_\_ and research design.
- Then, all labs' \_\_\_\_\_ are looked together at once and examined for their convergence or divergence.
- Because the scope of such an effort tends to be very \_\_\_\_\_ and focused on just one original study, we are more likely to see direct and, to a lesser extent, conceptual replications.

### 2.3.2 Many Studies, Many Labs

- This is largely an \_\_\_\_\_ of the above type, where we now focus on a broader issue or topic and try different studies within that area.
- A project such as this adds value by approaching a hypothesis with a variety of conceptual and extension replications, to see how different outcomes and \_\_\_\_\_ may contribute to significant findings (or not)

 Discuss: Why not just do this type with every important study? What challenges do we run into?

### 2.3.3 When It Doesn't Work

- Some replications find \_\_\_\_\_ or even opposite results from the original study!
- One possibility is that the two studies differ in some notable, \_\_\_\_\_ way, i.e., if you change many things, it is no surprise that the results are different
- Or, either the \_\_\_\_\_ or original study is flawed or a fluke
- So what next? Replicate some more! - A single-\_\_\_\_\_ replication is likely to not be sufficient to full dissuade the scientific audience from the original finding.
  - This is the advantage of multi-\_\_\_\_\_ studies - it helps ensure that reproduction is consistent or trends in a certain direction

## 2.4 Meta-analysis

- Instead of embarking on running many replications ourselves, we may choose to \_\_\_\_\_ summarize the existing evidence.

? Meta-analyses are technically a subtype of what paper type?

- A) Empirical original journal article
- B) Scientific journalism
- C) Literature review
- D) Replication

Explanation:

- \_\_\_\_\_-analyses quantitatively averages results across many similar studies to determine an aggregate/composite average.
- The averaging can be done with any \_\_\_\_\_ of effect size, such as Cohen's d and r.

### 2.4.1 Strengths and Limitations

- Meta-analyses can be an \_\_\_\_\_ way to give a summary of certain area with a clear number, informing future researchers of the state of the science at

that time.



Discuss: Meta-analyses tend to be very popular to cite, why do you think this is?

- However, this \_\_\_\_\_ lives and dies by its attention to detail and finding all relevant studies.
- Especially, null findings may be subject to the “\_\_\_\_\_ drawer problem”, and may be obscured from being collected and aggregated.
  - Most meta-analyses do recognize the role that publication bias can play in results

## 2.5 In Popular Culture


- Rarely, does popular journalism adequately capture the \_\_\_\_\_ and changing nature of science.
- Journalism is also partially drawn to \_\_\_\_\_ and new research, but may give less emphasis to historical, but still relevant, studies.
  - Replications tend to be somewhat less interesting, but still have very important results.
  - “If it bleeds, it leads”
- In general, it is much better to start with a peer-reviewed, scientifically \_\_\_\_\_ meta-analysis or literature review when orienting to a new area.

## 3 Transparency and Credibility

### 3.1 Overview

- Even well-meaning scientists (but also malicious actors) can make \_\_\_\_\_ in truthfully reporting results



 Discuss: Why do you think researchers publish research that is faked or misleading?

- It should always be our goal to make \_\_\_\_\_-driven predictions and designs - not making changes just for beneficial results.
- We have a variety of questionable practices to \_\_\_\_\_ and transparent practices to try to stick to which aid in reproducibility

## 3.2 Questionable Practices

### 3.2.1 Under-report Null Findings

- “Real” research tends to involve more than one outcome variable of interest, and sometimes more than one \_\_\_\_\_
- But, some authors may downplay many null \_\_\_\_\_ in favor of focusing on the shiny significant findings.
- This creates a narrative of \_\_\_\_\_ importance and reproducibility of findings, when the majority of evidence in a study actually points to the contrary.

### 3.2.2 HARKing

- This stands for “Hypothesis After Results are \_\_\_\_\_”
  - Generally, we want to make predictions and hypotheses prior to our data \_\_\_\_\_ and analyses
  - As a general ethical rule, it is bad to say “I knew it all along!” - as it is not \_\_\_\_\_ or literature driven.
-

### 3.2.3 P-hacking

- This occurs when multiple analysis methods are used \_\_\_\_\_ a significant p-value is found, and usually only the final method is reported
- Similar to under-reporting null findings, this tends to \_\_\_\_\_ contrary evidence and overstate otherwise weak findings
- Early decisions on design, hypotheses, and analyses need to be guided first-and-foremost by \_\_\_\_\_

? Review - what is it called when someone changes data to fit their hypothesis better?

- A) Fabrication
- B) Falsification
- C) Fixing
- D) Flubbing


Explanation:

## 3.3 Transparent Practices

- To \_\_\_\_\_ the previous, questionable techniques - we may use a few strategies to increase public and scientific confidence in our results.
- Often, these “open” methods intentionally make us \_\_\_\_\_ to wide critique from the broader public - but this is good for science!

### 3.3.1 Open Data and Open Materials

- Open \_\_\_\_\_ is when we publicly share anonymized data that can be re-analyzed by anyone to ensure results were correct
  - Open materials provide all in-depth procedures, measures, and tools to be used in any \_\_\_\_\_.
  - Both of these may be posted on repositories like <https://github.com> or <https://osf.io>
-

 Discuss: What circumstances might prevent releasing data and materials to the public?

### 3.3.2 Preregistration

- As previously mentioned, this is the \_\_\_\_\_ announcement of one's hypothesis and plans prior to completing a study.
- Doing this helps \_\_\_\_\_ any desire to HARK or p-hack, as any interested party could easily point out that you changed your future analysis.

## 4 When Do We Need (Really) External Validity

### 4.1 Overview

- Replication (especially \_\_\_\_\_) can help us establish just how widespread and generalizable effects are outside just our original sample of individuals.
- As we \_\_\_\_\_ original studies, we are able to better see just what people and contexts they can apply to
- External validity is one of the most \_\_\_\_\_ and difficult to establish validities - therefore we must discern how to balance our priorities with it

### 4.2 Generalization to Other People

- Remember that our sample comes from a \_\_\_\_\_ population that we chose, not just everyone in general
  - Samples are primarily strong by virtue of their \_\_\_\_\_, not only their size!
  - Samples can be easily \_\_\_\_\_ by convenience - think carefully about the impact that it is likely to have!
-

? Review - which of the following techniques results in an unbiased sample?

- A) Quota sampling
- B) Convenience sampling
- C) Self sampling
- D) Cluster sampling

Explanation:

- Results may not generalize to \_\_\_\_\_, but we must consider whether this is vitally important to us.

### 4.3 Generalization to Other Settings

- Many \_\_\_\_\_-based studies may benefit from replication in more “real-world” settings
- On the other hand, we may be interested to see if we can isolate a \_\_\_\_\_ behavior within the confines of the lab
- More broadly, we likely also want to examine how our studies look outside of the geographic and \_\_\_\_\_ areas they are conducted.

### 4.4 Do We Always Need to Generalize?

#### 4.4.1 Theory-testing Mode

- This is the “first” stage, where we attempt to \_\_\_\_\_ and describe a theory accurately under more controlled conditions
  - We work knowing external validity is \_\_\_\_\_, but taking care to establish other strong validities.
  - In this scenario, often \_\_\_\_\_ validity is treated as most important (if we have an experiment)
  - \_\_\_\_\_ claims usually work in a “theory-testing” mode.
-

? Early in the semester, we gave types of research based on their scope, similar to the present discussion - which of the following was not one of those types?

- A) Basic
- B) Supplementary
- C) Applied
- D) Translational

Explanation:

#### 4.4.2 Generalization Mode

- \_\_\_\_\_ and association claims more often aim to be widely applicable and relevant.
- Generalization mode places a special emphasis on making sure theories have strong \_\_\_\_\_ validity, putting less focus on internal, like when theory-testing
- \_\_\_\_\_ psychology is a sub-discipline almost entirely determined to investigate cross-cultural differences and similarities, working in generalization mode.

#### 4.5 Do We Always Need to be in the Real World?

- Valuable research happens in \_\_\_\_\_ setting, both lab-controlled, and completely ecological (i.e., the “real world”)!
  - Researchers in these separate settings often \_\_\_\_\_ greatly from the mutual shared knowledge.
-