

Invited Scholarly Essay

A Discussion of Falsifiability and Evaluating Research: Issues of Variance Accounted For and External Validity

David R. Ewoldsen

Department of Media and
Information, Michigan
State University, USA

Corresponding to
David R. Ewoldsen

Department of Media and
Information, Michigan State
University, 404 Wilson Road,
Communication Arts and
Sciences Building, Michigan
State University, East
Lansing, MI 48824, USA
Email: ewoldsen@msu.edu

Received
25 July 2022
Revised
29 July 2022
Accepted
30 July 2022

About the Author

David R. Ewoldsen is a professor in the Department of Media & Information at Michigan State University. He was the founding co-editor of *Media Psychology*, founding editor of *Communication Methods and Measures*, the first editor of the *Annals of the International Communication Association*, and he will be serving as a co-editor of the *Journal of Communication* starting in 2023. His research interests include media psychology, attitude accessibility, entertainment theory, and comprehension processes. He has published in the *Journal of Communication*, *Media Psychology*, *Human Communication Research*, *Communication Research*, *Communication Theory*, *Journal of Personality and Social Psychology*, *Journal of Experimental Social Psychology*, *Perspectives in Psychological Science* and other journals. He was named a Fellow of the International Communication Association in 2016. He earned a joint Ph.D. in psychology and communication studies at Indiana University and did a post-doctoral fellowship in cognitive sciences at Vanderbilt University. He grew up in rural Minnesota and Iowa and loves kimchi jjigae.

KEYWORDS

falsifiability, variance explained, generalizability

When I was in graduate school, the discipline of communication was going through somewhat of an identity crisis. Who were we? Did we belong among the “important” disciplines such as sociology, psychology, and so forth? In many ways, the discipline felt like a forgotten stepchild. As is common in these periods, there were a lot of discussions about the nature of science. Probably the two most popular philosophers of science within the discipline at that time were Thomas Kuhn (1970) and Sir Karl Popper (1959). Graduate students who have worked with me across the years have heard me recounted in exaggerated terms how at conferences such as ICA, you’d find groups of Kuhnian’s running around in packs all looking cool in their leisure suits. On the other hand, the Popperian’s were a stuffy lot standing around

in their tweed jackets looking down on everyone else because no one was trying to prove their ideas were wrong (e.g., falsification).

My MA theory courses had an unusually heavy focus on philosophy of science, and I started my Ph.D. program as an avowed Kuhnian. But while working on my Ph.D., I took courses in the Philosophy of Science program which resulted in my reading several volumes of Popper's work. And I became, and continue to be, an advocate of Popper's views on how science progresses. I want to be clear that I do not reject Kuhn's (1970) basic ideas (see Ewoldsen, 2017, 2020), but I do find Popper's (1965, 1979) notions of how science progresses to be more useful in guiding my own theoretical endeavors. It is my perception that there is a basic misunderstanding of Popper's philosophy within the discipline. People largely rely on Popper's (1959) classic *The Logic of Scientific Discovery* without reading his later works that clarified the ideas presented in *The Logic – most notably Conjectures and Refutations and Objective Knowledge*. In the following essay, I will present and defend my interpretation of Popper's writings, and then discuss the implications of Popper's philosophy of science for several different methodological issues, including variance accounted for and generalization.

Sir Karl Popper: A Misunderstood Philosopher

One of the foundations of Popper's work was the elaboration of a solution to the eighteenth-century philosopher David Hume's (1896) induction problem. Hume demonstrated in his classic volume, *A Treatise of Human Nature*, that nothing can be proven with certainty using induction. Recall that induction is when we draw general conclusions from specific observations. When we use experiments or ethnographies (looking at specific instances) to test theories (general conclusions) we are relying on induction. Consequently, Hume's induction problem

became a major issue for empirical researchers. Hume's point was that when we rely on induction, we cannot base our conclusions on all the observations of the past (because we were not there to observe them) or the future (because they haven't happened yet). Consequently, we can't know with certainty that the same event or process will occur in the future. A classic example involves swans. For a scientist operating in the eighteenth century in Europe, it is highly likely that the first swan they encountered would be white. This could lead to the tentative inference that whiteness is a defining characteristic of being a swan. As the scientist encounters more and more white swans, the inference that all swans are white would likely be held with more and more confidence. After seeing hundreds of white swans, the scientist may conclude that, indeed, all swans are white. It seems reasonable for the scientist to conclude this because all of the numerous swans that the scientist has observed were white, but of course the scientist is wrong. The scientist is unable to observe all swans and is unaware that there are black swans in other parts of the world (and, indeed, now there are black swans found in Europe).

The general approach to Hume's induction problem that had been largely adopted was the observation that induction can only result in various degrees of confidence in a theory. However, the reality is that we all use induction and often treat it as infallible (e.g., we all have confidence that the sun will appear each morning, although we have not witnessed all sunrises in the past and the future). Popper's contribution to the induction problem was the observation that although we can't prove something is true via induction, we can use induction to prove something is false with certainty. In other words, science can operate by falsifying conjectures.

Unfortunately, many people have interpreted Popper and his notion of falsification as indicating that we always should be trying to falsify our theories. There is a grain of truth in

this understanding of Popper, but his arguments regarding how science advances are much more complex than this simple approach to falsification. The principle of falsifiability – that scientific theories should be able to be falsified – was, for Popper, what demarcated science from other epistemologies such as rationalism, common sense, or faith. Science operates by making conjectures, making observations based on those conjectures, and trying to improve theory based on this process.

At the heart of Popper's solution to the induction problem is the concept of verisimilitude. Traditionally, verisimilitude involves the degree of truth or truthfulness found within a theory. The problem of induction denied that we could establish the truth-value of a theory. For Popper verisimilitude involves decreasing the level of falsity within a theory via induction. There is often a fair degree of ambiguity in the initial stages of studying a particular phenomenon. Consequently, the theories within that area of study are rather vague because the nature of what is to be explained is, as of yet, not well understood. For Popper, this lack of knowledge or vagueness indicates that there may be a lot of falseness in vague theories. The greater the vagueness, the greater the potential for falseness. How do we make precise predictions about something if we know very little about it?

For Popper, the goal of decreasing the falseness of a theory involves trying to make the theory more specific. As a theory's specificity increases, the precision of the theory's predictions increases and the theory becomes more falsifiable (e.g., there are more observations that will be inconsistent with the theory). Popper proposed that an emphasis on falsification would allow theories to evolve and become more precise and, as a consequence, our understanding of the phenomenon of interest would increase.

Many scholars have interpreted this as meaning we should always be attempting to prove our theories wrong. When the predictions of a theory

are not confirmed by an observation, the theory should be considered falsified and rejected. This is certainly a legitimate possibility, but only one possibility. As Popper (1965, 1979) noted in his later writings, there are many reasons for why an observation may "falsify" a theory. One possibility is that the theory is false and should be rejected. Other reasons include random chance, improper manipulation of the independent variables, or invalid measurements. To reject a theory because of bad methodological choices or random chance hardly serves the goal of advancing our knowledge. Indeed, if the theory is rejected because of bad observations, we have learned nothing. Instead of simply rejecting each "falsified" theory, Popper argued for an evolutionary approach to theory development. When there is an anomaly (e.g., when there are observations that are inconsistent with the theory or the theory has been falsified) and people are confident that the anomaly is accepted, then one possibility is to modify the theory to make it more specific by adding tenets to the theory that explain why this anomaly occurred or setting boundary conditions which specify when the theory works and when it is expected to not work. These modifications to the theory decrease the falseness of the theory because the anomalous result that was inconsistent with the theory is now explained. For Popper, a theory becomes less false through modifications that make the theory more specific.

In other words, Popper argues that scientific knowledge advances through a process of seeking falsification through more and more precise predictions. When there are "falsifying" instances, one possibility involves the rejection of the theory and, *critically*, replacing the old theory with a newer theory that can explain everything that the old theory could as well as the anomalous finding that falsified the previous theory. Consequently, the new theory is more specific. The second possibility is to modify the existing theory in order for that theory to then explain the "falsifying" instance. In this way, the old theory is refined

to make it more specific. The goal for Popper is always to attempt to increase the specificity of our theories.

As I hope I have made clear, what I have presented is my interpretation of Popper. In the next sections, I will explain what I see as the implications of Popper's views of theory for several issues that reviewers love to comment on – at least they do when reviewing my research.

How Important are the Results, Really? Variance Explained

Historically, having a statistically reliable finding (e.g., there was a less than 5% chance that the effect occurred by chance) was the foundation for whether the findings were important or not. However, during my early career reviewers would often note that while the results are statistically significant, the effect size is rather small so is this really an important finding? This is a critical question. There is a strong tendency to reject results as not meaningful because they account for only a small percentage of the variance.

A question raised by this criticism is what exactly is the level of variance that needs to be accounted? And many scholars have their own criteria for what makes a result important. I have had scholars tell me that unless my results account for at least 20% of the variance, my findings are not important. However, as the discipline has turned more attention to dynamic processes, it became clear that findings that account for a very small percent of the variance can have huge consequences (Lang & Ewoldsen, 2010). Within a dynamic system, a rather small perpetuation can result in substantial change within the system if the perpetuation falls at the bifurcation point within the dynamic system (an example of this is the so-called “butterfly effect”).

Another issue raised by this critique involves the notion of accumulated variance (Abelson, 1985). Many experiments involve a single observation of the participants. The question is whether

the impact of the experimental manipulation dissipates and is new each time a participant engages in a task (like priming) or whether the impact of the manipulation accumulates across time (e.g., increasing chronic accessibility; Ewoldsen & Rhodes, 2020). Abelson (1985) provides the example of a professional baseball player in a single at bat. If you use a professional baseball player's batting average as a proxy for skill at the plate (e.g., a player with a higher batting average is a more skilled batter), Abelson demonstrated that then skill accounts for approximately 1.3% of the variance in a single at bat. Do we really want to conclude that skill is not important because it doesn't account for much variance in an at bat? If skill (as measured by batting average – a measure of the percentage of the times a batter hits the ball out of all of their chances to hit, or “at bats”) does not account for much variance in getting a hit, why are baseball players who bat .320 paid so much more than baseball players who bat .280? The reason is that the players are paid for a season with 165 games and approximately 4 at bats per game. Across all of these at bats, the variance accounted for by skill accumulates and becomes very important. In other words, sometimes variance accumulates across time. If playing a violent video game in a cooperative way increases tit-for-tat reciprocity after 15 minutes of gameplay but only accounts for about 5% of the variance, is that important (Ewoldsen et al., 2012)? I would argue it is because most video game players do not play a violent game cooperatively for only 15 minutes. If the variance accumulates across time this could be an extremely important finding. Within a Popperian perspective, the next step should be to determine if the variance is cumulative or not rather than simply dismissing a finding out of hand because it does not account for much variance. In other words, we should be trying to increase the specificity of our understanding of the phenomenon rather than dismissing the finding out of hand.

Within Popper's philosophy, there is also another instance where science is advancing, but variance accounted for within a study is likely decreasing. Recall that Popper's position is that increased specificity is the hallmark of scientific advancement. As our theories become more and more specific, tests of the theory will also, by necessity, involve more and more specific predictions. Likewise, critical tests of two or more highly specific theories may involve an incredibly precise test. Consider the common ingroup identity (CII) model (Gaertner et al., 2000). The CII has developed out of a stream of theories and research originating with social identity theory (SIT) and self-categorization theory. SIT has been the most influential theory used to explain ingroup and outgroup effects within our discipline. People augment their perception of their ingroup to enhance their self-esteem (e.g., I am wonderful because I belong to a wonderful group) and they disparage outgroups, again, to enhance their self-esteem (e.g., only horrible people belong to that group, our group is so much better). Self-categorization theory was an extension of SIT that maintains that people naturally categorize themselves into one of three levels of identity: the self, the group, and the supraordinate category. The level of the category that is most salient drives the types of inferences that a person makes. When the person is focused on the individual level, characteristics of that individual predict the person's behavior such as the person's traits and personal roles. When the group level is salient, the processes outlined by SIT operate and people categorize themselves and others into groups. People in outgroups are judged negatively and stereotyped. The supraordinate category is the highest or most abstract category and involves categories such as "human" and stereotyping (putting someone in the outgroup) should decrease.

The CII builds upon refinements made by self-categorization theory to specify when people are likely to operate at the personal, group, or

supraordinate category level. Specifically, CII hypothesizes that how a group is categorized (as a mid-level group such as "white people") or as part of the supraordinate group (human) can be influenced by situational factors that make one level or the other more salient. For example, simply using the supraordinate category of "human" does not guarantee that people will respond in a non-stereotypical manner. If the individual is a white person and the situation makes white people salient, the person may represent human as meaning "white humans." In this case, stereotypes and prejudices may unfortunately still operate or even increase (Ellithorpe et al., 2018). However, if the situation makes the diversity of humans salient, then the supraordinate category "human" is more likely to include all humans instead of just white humans. In this later instance, operating at the supraordinate "human" category should decrease stereotyping.

From a Popperian perspective, this is the way science should progress because as we move from SIT to self-categorization theory to CII, the predictions made by the next generation of theory become more and more specific. The theories are becoming increasingly falsifiable because the false content of the theories is decreasing and our understanding of how these processes work is increasing. Clearly, this hardly means we know everything we need to know about in- and outgroup processes, stereotyping and so forth (cf., Holt et al., in press). But certainly, our knowledge is growing because of the increased specificity.

Returning to the issue of Popper and variance, I hope the point is clear is that as theories progress, they make more and more precise predictions. Consider a recent test of the applicability of the CII to the media. Ellithorpe et al. (2018) hypothesized that programming where the villain was a supernatural creature (e.g., a werewolf or zombie) would make the supraordinate category of human salient and consequently, people would be more likely to perceive groups based

on their humanness rather than more mid-level group categories such as “Black people” or “Asian people.” To test this, Ellithorpe and colleagues presented participants with a story that involved a human hero fighting either other humans or supernatural creatures. In addition, Ellithorpe et al. manipulated whether the hero was a White person or a Black person to make salient the diversity of humans. The prediction was that when the creatures were supernatural, that would make the supraordinate category of human salient. If the hero was a Black person, that should further make the diversity of the human category salient which should increase the salience of a diverse supraordinate category of humans and decrease prejudice toward Black people. In other words, they predicted a three-step serial mediation (salience of being human to strength of the supraordinate human category to attitudes toward Black people) moderated by the race of the hero. This is a very precise prediction and provides a very strong test of the theory. And, indeed, this is what Ellithorpe et al. found in their study. However, the R^2 for this model was approximately .09. Indeed, for some people, the critical issue would be that this very precise prediction does not account for much variance (~9%) so it can't be that important. Obviously, as a co-author on that paper I disagree. But personal interest aside, I think this is an issue that people need to think seriously about. As we know more and more within a particular domain, our theories will become more and more precise. These precise predictions can provide very elaborate tests of a theory, but often they do not account for much variance. In other words, an implication of Popper's epistemology is that often these most critical tests of a theory will not account for much variance.

Do Your Findings Tell Us about the Real World? Issues with Generalizability

The external validity of research results is a major

issue for communication and media scholars. Our discipline cares a lot about external validity which makes sense historically because the discipline emerged from a problem focus (Krcmar et al., 2016) such as what are the effects of violent TV or how can we design better interventions to improve vaccine compliance. Given this historical focus on societal problems, it is no wonder the discipline has a high concern for the external validity of the research published in the journals. The major two issues that I often deal with involving external validity involve the research participants that were used (the “college sophomore” problem) and the reliance on single messages to manipulate an independent variable.

Probably the best article I have ever read is Mook's (1983) *In Defense of External Invalidity*. The central claim of the paper is that outstanding research is often designed with absolutely no regard for whether the research generalizes to the real world. Instead, the goal of the research is to provide a careful test of the theory. As theories become more precise, finding ways to test the theory that are externally valid can become difficult. The careful experimental control that is necessary to conduct the precise test of the theory may preclude any concern with external validity. For example, early work I was involved with on the formation of stereotypes required the creation of artificial groups so we could explore how stereotypes developed – a task that would be impossible if we used real groups where stereotypes already existed (Sherman et al., 1989).

I am particularly sensitive to the “college sophomore” problem because I conduct many laboratory-based experiments and heavily rely on college students. I cannot count the number of times reviewers have criticized my work because of my “overreliance” on college sophomores. In one of my first research methods classes, I remember the professor discussing the issue of over reliance on college sophomores for our research participants. As a college sophomore, I was somewhat insulted by the argument that my

professor was making. So exactly what was it that was the issue? How was I, as a college sophomore, different from “real” people.

In a classic paper, Sears (1986) addressed the issue of overreliance on college sophomores. As Sears noted, certainly college students are distinct from other populations in various ways. College students tend, on average, to be more intelligent, have a less well-formed attitudes, are more susceptible to social influence, and engage in less introspection. But what are the real impacts of these differences? Sears argued that there are several possibilities that can result from the use of college students. First, the results may be identical to what they would be with a representative sample. Second, the effect size may be either larger or smaller for college students than it would with a noncollege student sample. Third, there are moderators or mediators that may work with college students that do not work in a more general sample or vice versa. Finally, the use of a college student sample may result in the finding of a null effect when the effect does occur in other populations. Certainly, this last possibility is a serious outcome, particularly when the finding is used to reject what may be an underspecified theory.

To me, the mindless critique that a study relies on college students misses the point. From a Popperian perspective, the argument that a college sample may, in some way, distort findings is operating in the exact opposite way that Popper hypothesized science advances. In other words, the college student argument is essentially saying that the theory should be nonspecific and apply to all people. But for Popper, we learn more by becoming more specific. We should not reject research simply because it relies on college students, we should focus on the theory and specify what attributes of people influence how the theory operates. If I am studying dissonance theory and I am concerned that my reliance on college students may have distorted the findings because college students tend to have weakly

held attitudes, then I should conduct additional research to test whether weakly held attitudes operate as a moderator of dissonance processes. By identifying what it is we think may limit the generalizability of a finding due to the sample of participants, we have the potential to specify moderators that operate within the domain we are studying, and we have a more precise theory.

A second area where communication scholars often discuss the issue of generalizability involves the use of a single message versus the use of multiple messages within an experiment. When a single message is used, reviewers will often argue that we don't know whether the results are due to what was theoretically manipulated by the message (e.g., high vs. low fear) or are due to something else about the particular message (e.g., a third, or confounding, variable). But that is the point – even when we use multiple messages, *we don't know if there is a third variable within the message that may be causing the effect (or moderating or mediating the effect)*. Simply arguing a study should be rejected because it only uses a single message is not going to increase our understanding of the third variable. To answer the question, we need to pay more attention to theorizing about messages. For example, we could ask what else there is in the high fear appeal message that could cause the change in attitudes (or interacting with fear to change the attitude).

I want to be clear that trying to control extraneous variables is a good research practice and the use of multiple messages is often motivated by the desire to control extraneous variables. But I think there is a danger in this practice because, at least to me, it stops us from understanding the complexity of messages. If scholars were to analyze the various messages that have been used to study fear appeals, they may discover (hypothetically) that successful fear-appeal messages are more likely to use more visual descriptions. This finding could inform our theorizing about fear appeals. We are communication scholars so you would think

we might be good at theorizing about messages, but this type of theorizing has rarely been done. It may seem like an overwhelming challenge to begin developing theories that account for all of the different components of messages. But when we first started studying memory or emotion or entertainment or human motivation, the same sense of being overwhelmed was likely present. So we start somewhere and in 50 years' time, graduate students will be learning the new theories and amazed at how we could do research 50 years ago without these new theories.

It is my view that focusing on the generalizability of a study to the "real" world is asking the wrong question. Instead, I think there are two critical questions. First, does the research generalize to the theory it is trying to test? Is it a valid test of the theory and what does the research tell us about the verisimilitude of that theory? Second, and I think this is critical, how do our theories generalize to the "real" world? Communication scholars (myself included) have rarely focused on the boundary conditions for our theories. But, from a Popperian perspective, that is critical for increasing the verisimilitude of a theory. When does the theory work and when doesn't the theory work and why? As we focus more on these questions, our answers will provide better guidance for "real" world interventions.

So What?

The implications of taking an expanded view of Popper's ideas are important for the development of knowledge of communication phenomena. To accomplish this, we need to do more theory testing, by which I mean, we need to do more to make our theories specific. There are numerous ways to do this. Something I have observed as an editor across many years is a large number of manuscripts where the research is designed to find results that are consistent with the theory but not to advance the theory. For example, many manuscripts I've reviewed or handled as an editor

argued that the finding of outgroup effects was a test of SIT. Finding outgroup effects is not a test of SIT. There are other well-supported theories that also predict in- and outgroup effects. Testing SIT requires exploring the processes specified by SIT the result in in- and outgroup effects (e.g., self-esteem or enhancement of the self).

Another way to increase precision in our theories is moving from an effects focus (what is the effect of playing violent games online in groups) to focusing more on the processes underlying the effect (How do the processes outlined by SIT explain toxic gaming environments?). As I've already alluded to, specificity also comes from stipulating when a theory operates or does not operate. For example, we have advanced the claim that theories within the Reasoned Action Approach (Fishbein & Ajzen, 2010) such as the theory of reasoned action or the theory of planned behavior only explain the attitude-behavior relationship when people are engaged in more systematic or deliberative processing (Ewoldsen et al., 2015). Similarly, testing the circumstances under which one theory operates and the circumstances under which another theory operates increases the precision of both theories. For example, in the 1960s, self-perception theory and dissonance theory were pitted against each other as incommensurate theories aimed at explaining the relationship between a person's behavior and subsequent attitude change. Both theories could not be right. Yet, later research demonstrated that each theory could predict people's behavior but that they operated in different circumstances. When a person's behavior was very discrepant from their attitude, dissonance theory did a good job of explaining people's attitude change. But when the person's behavior was not very discrepant from their attitude, self-perception theory did a good job of explaining people's attitude change (Fazio et al., 1977). Neither theory was "wrong" but they do operate in different contexts.

The discipline has come a long way from those

days when we were the forgotten stepchild. Our stature within the academy has risen dramatically. But, as a consequence, there is less focus on philosophy of science and the growth of knowledge. And I think this comes at a cost. In my own experiences as an editor and a reviewer, I feel that we do not focus enough on theoretical development and how our research practices impact the evolution of our theories. Simply doing research that is consistent with a theory does not advance the theory. Research needs to push our theories to be more specific. That is how we will expand our knowledge and continue to raise our stature as a discipline.

REFERENCES

- Abelson, R. P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin*, 97(1), 129–133.
<https://doi.org/10.1037/0033-2909.97.1.129>
- Ellothorpe, M. E., Ewoldsen, D. R., & Porecca, K. (2018). Die, foul creature! How the supernatural genre affects attitudes toward outgroups through strength of human identity. *Communication Research*, 45(4), 502–524.
<https://doi.org/10.1177/0093650215609674>
- Ewoldsen, D. R. (2017). Normal science and paradigm shift. In J. Matthes, C. Davis, & R. Potter (Eds.), *The international encyclopedia of communication research methods* (pp. 1–17). Wiley.
- Ewoldsen, D. R. (2020). Verification and falsification. In J. Van den Bulck, D. Ewoldsen, M.-L. Maries, & E. Sharrer. (Eds.), *The international encyclopedia of media psychology* (pp. 1–5). Wiley.
- Ewoldsen, D. R., Eno, C. A., Okdie, B. M., Velez, J. A., Guadagno, R. E., & DeCoster, J. (2012). Effect of playing violent video games cooperatively or competitively on subsequent cooperative behavior. *Cyberpsychology, Behavior, and Social Networking*, 15(5), 277–280.
<https://doi.org/10.1089/cyber.2011.0308>
- Ewoldsen, D. R., & Rhodes, N. (2020). Priming and accessibility. In M. B. Oliver, A. A. Raney, & J. Bryant (Eds.), *Media effect* (pp. 83–99). Routledge.
- Ewoldsen, D. R., Rhodes, N., & Fazio, R. H. (2015). The MODE model and its implications for studying the media. *Media Psychology*, 18(3), 312–337.
<https://doi.org/10.1080/15213269.2014.937440>
- Fazio, R. H., Zanna, M. P., & Cooper, J. (1977). Dissonance and self-perception: An integrative view of each theory's proper domain of application. *Journal of Experimental Social Psychology*, 13(5), 464–479.
[https://doi.org/10.1016/0022-1031\(77\)90031-2](https://doi.org/10.1016/0022-1031(77)90031-2)
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. Psychology Press.
- Gaertner, S. L., Dovidio, J. F., Nier, J. A., Banker, B. S., Ward, C. M., Houlette, M., & Loux, S. (2000). The common ingroup identity model for reducing intergroup bias: Progress and challenges. In D. Capozza & R. Brown (Eds.), *Social identity processes: Trends in theory and research* (pp. 133–148). Sage.
- Holt, L. F., Ellithorpe, M. E., Ewoldsen, D. R., & Velez, J. (in press). Helping and hurting on the TV screen: Bounded generalized reciprocity and interracial group expectations. *Media Psychology*.
<https://doi.org/10.1080/15213269.2022.2026228>
- Hume, D. (1896). *A treatise of human nature*. Clarendon Press.
- Krcmar, M., Ewoldsen, D. R., & Koerner, A. (2016). *Communication science, theory, and research: An advanced introduction*. Routledge.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press.

- Lang, A., & Ewoldsen, D. (2010). Beyond effects: Conceptualizing communication as dynamic, complex, nonlinear, and fundamental. In S. Allan (Ed.), *Rethinking communication* (pp. 109–120). Hampton Press.
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 38(4), 379–387.
<https://doi.org/10.1037/0003-066X.38.4.379>
- Popper, K. R. (1959). *The logic of scientific discovery*. Harper Torchbooks.
- Popper, K. R. (1965). *Conjectures and refutations: The growth of scientific knowledge*. Harper Torchbooks.
- Popper, K. R. (1979). *Objective knowledge: An evolutionary approach* (revised edition). Claredon Press.
- Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3), 515–530.
<https://doi.org/10.1037/0022-3514.51.3.515>
- Sherman, S. J., Hamilton, D. L., & Roskos-Ewoldsen, D. R. (1989). Attenuation of illusory correlation. *Personality and Social Psychology Bulletin*, 15(4), 559–571.
<https://doi.org/10.1177/0146167289154009>