Contrasting Contrasts: An Exploration of Methods for Comparing Indirect Effects in Mediation Models

A Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Jacob J. Coutts, B.S.

Graduate Program in Department of Psychology

The Ohio State University

2020

Thesis Committee

Dr. Andrew F. Hayes, Advisor

Dr. Jolynn Pek

Dr. Steven N. MacEachern

Abstract

Establishing a cause–effect relationship between two variables is a fundamental goal of scientific research. It is valuable to know that one variable causally influences another. Just as important, however, is establishing *how*, or through what mechanism(s), this effect operates. Mediation analysis is a popular method used to answer such questions. A simple application of the mediation model looks at how one intervening variable (a "mediator") can explain the relationship between two others. The quantification of an effect through a mediator is called an indirect effect. Real-world processes are complex, however, and effects are often transmitted by more than one mechanism. Consequently, it can be beneficial to simultaneously look at multiple mediators that could explain the connection between an antecedent variable and its consequent. As multiple mediator models continue to grow in popularity, it is theoretically and practically useful to explore whether one mechanism is “stronger” or “more important” in producing an effect than another. This can be done by comparing the relative sizes of the indirect effects. Although several methods have been proposed in the methodological literature for comparing indirect effects, little to no literature exists exploring whether one method is better than another. The goal of this thesis is to first give a background on mediation analysis and multiple mediator models. Then I discuss current approaches to comparing

indirect effects and suggest alternative ways of doing so. Next, I demonstrate these concepts by conducting analyses in a real-world example. Finally, I conduct a simulation study comparing methods for comparing indirect effects and give suggestions for future research.

Dedication

This thesis is dedicated to the wonderful persons who have supported me throughout graduate school: my parents, siblings, cohort, advisor, department faculty, and friends who have been patient with my busy schedule. It has been a roller coaster that I'm privileged to be on.

Acknowledgments

I would like to acknowledge the great feedback I received from my committee and their patience during the strangest year of my life. I would especially like to acknowledge my advisor, Dr. Andrew Hayes for helping refine my topic (and simulation code).

Vita

November 29, 1995 ……………………………………… Born - Illinois, USA

May 2018 ………………………………………………… B.S. Psychological Sciences
Northern Arizona University

August 2018 – Present …… …………………………...... Master of Applied Statistics
The Ohio State University

August 2018 – Present …… …………………………...... Ph.D. Quantitative Psychology
The Ohio State University

August 2018 – August 2019 …………………………...... University Fellow,
The Ohio State University

August 2019 – Present ………………………………....... Graduate Teaching Associate,
The Ohio State University

Publications

Hayes, A.F. & Coutts, J.J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But… *Communication Methods and Measures*. *14*(1), 1-24.

Coutts, J.J., Hayes, A.F., & Jiang, T. (2019). Easy statistical mediation analysis with distinguishable dyadic data. *Journal of Communication*, *69*(6), 612-649.

Fields of Study

Major Field: Psychology

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Psychology and other scientific disciplines have long sought to explore causal relationships between variables. Knowing that one variable causes another provides valuable information to the scientific community. For instance, if self-compassion causes well-being to improve, it can become a focal point of behavioral or therapeutic interventions to improve an individual's state of mind. But knowing *that* a relationship exists is only part of the story. Just as important is to understand *how* a relationship exists. Perhaps self-compassion increases well-being because greater self-compassion first causes greater optimism, which, in turn, causes well-being to improve. That is, perhaps self-compassion is acting indirectly *through* optimism to enhance one's well-being. Mediation analysis can be used to answer such questions. Mediation analysis is used to understand how a predictor (hereinafter, X) is related to an outcome (hereinafter, Y) through one or more mediating variables (hereinafter, M).

Mediation models are abundant in applied research. In one study, Schönfeld, Brailovskaia, Bieda, Zhang, and Margraf (2016) found that daily stress had a significant and negative effect on a person's mental health. But the researchers thought that this association was at least partially explained by another, intermediary variable (a mediator). They posited that more stress would reduce a person's level of self-efficacy, and this reduction in self-efficacy is what would cause poorer mental health. This explanation was supported by the data. Another study measured the relationship between compassion for others and happiness (Sanchez, Haynes, Parada, & Demir, 2018). Those who reported higher levels of compassion for others also reported greater levels of happiness and mental well-being. The authors proposed that being compassionate leads to prosocial friendship maintenance behavior like providing support/advice or reflecting on good times together. This friendship maintenance, they argue, is what leads to greater levels of happiness. These are merely two of many examples of mediation analysis conducted in research. Mediation models have been applied in a number of other fields including communication (Hoffman & Young, 2011), public health (Ho, Peh, & Soh, 2013), education (Jin, McDonald, & Park, 2018), nursing (Van der Heijden, Mahoney, & Xu, 2019), and business (Wieder & Ossimitz, 2015), among many other allied areas of research.

Although the studies above provide evidence for a mechanism underlying a causal relationship, it is unlikely that just *one* mediator explains the relationship between two variables linked in a causal process. Rather, it is more likely that one variable causes multiple others that in turn lead to some outcome. For instance, in the first example above, could it be that daily stress causes another variable, say rumination, and rumination *also* causes poorer mental health? Introducing more mediators into a model provides more information about the nomological network of the variables of interest. Simultaneously estimating mechanisms between X and Y is beneficial to researchers for a variety of reasons, chief among these is that researchers can test competing theories against each other by comparing mechanisms present in a model. In the example so far—if both mediators were measured in the same study—the researchers could test if there was a difference between the effect of stress through rumination on well-being and the effect of stress through optimism on well-being and see whether one was "more important" in explaining the relationship than another. A number of suggestions have been proposed in the literature for comparing such effects, but none of them have been tested to see whether one performs better than another or in what conditions one performs better (e.g., sample size or size of effects being compared).

As the use of mediation models—especially those with multiple mediators—continues to rise, it is beneficial for methodological scholars to investigate such questions so substantive researchers can carry out tests that have good statistical properties. This thesis expands on existing recommendations for comparing indirect effects by testing several contemporary approaches in a simulation study. First, I discuss the mathematics behind mediation models. Then I describe simple and parallel mediation models and how one can obtain inference about effects in these models. After, I talk about current practice and new approaches for comparing indirect effects, step through a real-world example, and conduct a simulation study comparing the performance and merits of these different approaches. Finally, I discuss the implications of this research for substantive researchers.

1.1 Data from a substantive example

I will refer to a substantive example throughout this thesis. The data I use are inspired by and simulated from a paper published by Pitts, Safer, Castro-Chapman, and Russell

(2018). The study sought to explore how veterans' combat experiences caused posttraumatic symptoms once they returned home from war. It is well-understood that combat experiences are a cause of PTSD in veterans, but combat is an unavoidable experience for those going to war. Mediation analysis can help researchers uncover what combat experiences cause that might then lead to PTSD. This information can help inform interventions and therapeutic efforts to improve the well-being of those in the military. The authors hypothesized that army medics who experienced combat during their most recent deployment would perceive more threats to their life, and this greater level of perceived threat would lead to an increase in posttraumatic stress symptoms upon return from deployment. They also hypothesized that combat experience would lead to an increase in the perceived benefits of deployment which would in turn lead to a decrease in posttraumatic stress symptoms. The researchers *X* was combat experience, the mediators, *M*s, were perceived threat and perceived benefits of deployment, and the outcome, *Y*, was posttraumatic stress symptoms.

Chapter 2: The Mediation Model

Mediation models are typically estimated in one of two ways: structural equation modeling (SEM) or through a series of ordinary least squares (OLS) regression analyses. Although some researchers have taken a stance about which approach should be used to estimate the effects in these models, research suggests that the results between observed-variable SEM and OLS regression are nearly identical (Hayes, Montoya, & Rockwood, 2017; Rijnhart, Twisk, Chinapaw, de Boer, & Heymans, 2017). There are pros and cons to conducting mediation analysis through maximum likelihood estimation in SEM and through OLS criteria in regression (detailed in the papers above), but this thesis will discuss mediation analysis in the context of a series of regression models.

Mediation models can be estimated using two equations where the paths are estimated in each equation separately (as opposed to simultaneously in SEM). If M and Y are continuous variables and X is either dichotomous or continuous, then a mediation model can be estimated by fitting the following equations to the data:

$$M = d_M + aX + \varepsilon_M \tag{2.1}$$

$$Y = d_Y + c'X + bM + \varepsilon_Y \tag{2.2}$$

Figure 2.1 shows a statistical diagram of this model. This is called the *simple mediation model* (this is not to imply that the model is conceptually simple or that there aren't complexities and controversies in interpretation—simple in this case means that it is a three-variable system with one X, M, and Y). In equation 2.1, a represents the expected difference in M with a one-unit increase in X. In equation 2.2, b represents the expected difference in Y with a one-unit increase in M holding X constant, and c' represents the expected difference in Y with a one-unit increase in X holding M constant (this is called the *direct effect* of X). The *indirect effect* of X, the effect usually of interest in mediation analysis (the effect of X on Y through M), is obtained by multiplying a and b together (i.e., ab). The *total effect* of X on Y, c, can be obtained by adding the indirect and direct effects (i.e., $c = ab + c'$). The total effect represents the expected difference in Y with a one-unit increase in X and is what would result from regressing Y on X alone.

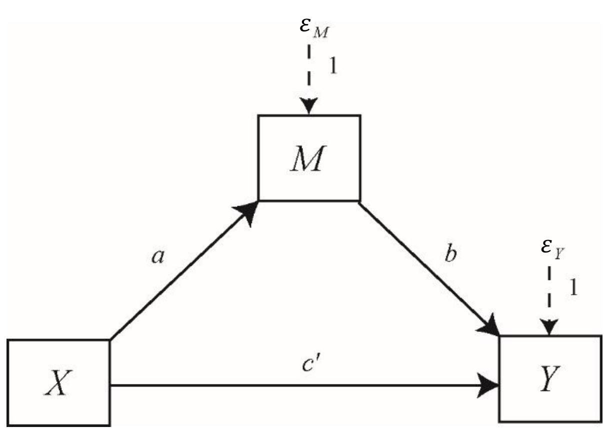

Figure 2.1: A statistical diagram for the simple mediation model. ε denotes residuals in estimation.

Inference for a, b, c, and c' can be obtained through standard regression procedures. That is, you can obtain p-values by dividing the regression weight by its standard error and comparing the resulting test statistic to a t-distribution (to test against the null that the weight is equal to zero) or obtain 95% confidence intervals by calculating, for example for a,

$$a \pm t_{.975,n-2} se_a \tag{2.3}$$

where se_a is the standard error of a and $t_{.975,n-2}$ is the value for which 97.5% of the t-distribution falls below at $n-2$ degrees of freedom. (Inference for b, c, and c' would be conducted similarly.)

Inference for the indirect effect is not as simple since it is the product of two regression coefficients. Historically, several methods have been used for inference about the indirect effect including the Sobel test (Sobel, 1982), the causal steps approach (Baron & Kenny, 1986), and the joint significance test (MacKinnon et al., 2002). These methods are no longer recommended by mediation scholars.

2.1 The Bootstrap

The bootstrap is arguably the most popular contemporary method for inference about the indirect effect. The bootstrap is a nonparametric resampling approach that has gained popularity in recent years with the increase in general computing power and computational tools available to substantive researchers to carry out such methods (e.g., PROCESS; Hayes, 2018). The appeal of the bootstrap as an inferential method is that it makes no assumptions about the shape of the sampling distribution of the indirect effect.

When bootstrapping, the original sample is treated as a population and an arbitrarily large number of *j* samples of size *n*—the size of the original sample—are drawn with replacement from the data (consequently it's important that the original sample is representative of the population of interest). In general, at least 5,000 to 10,000 resamples of the original data set are recommended. In each of these bootstrap samples, the indirect effect (*ab*) is estimated and recorded and this procedure is repeated *j* times. Typically, for inference, a confidence interval is constructed from the resulting bootstrap distribution of indirect effects. There are a variety of ways to construct this bootstrap confidence interval.

2.1.1 Percentile Bootstrap Confidence Intervals

The percentile bootstrap confidence interval is simple to construct and is one of the most used for inference about the indirect effect. A *ci*% confidence interval can be constructed by calculating the upper and lower $\frac{(100-ci)}{2}$ percentiles of the distribution of the *j* estimates of *ab* (with *ci* being the specified level of confidence). For instance, if *j* = 1,000 bootstrap samples, the 2.5 and 97.5th percentiles of the distribution would be the endpoints for a 95% *percentile bootstrap confidence interval*. If the confidence interval for *ab* does not contain zero, then *M* is deemed to mediate the relationship.

2.1.2 Bias-Corrected and Bias-Corrected and Accelerated Bootstrap Confidence Intervals

The percentile bootstrap approach can sometimes lead to over or underestimates of the population value (Efron, 1987; Efron & Tibshirani, 1993). Bias-corrected (BC) and bias-corrected and accelerated (BCa) confidence intervals are alternatives that attempt to correct for bias/inaccuracies observed with the percentile method (Efron & Tibshirani, 1993). These methods still rely on percentile values from the bootstrap distribution to obtain confidence intervals, but the endpoints will vary based on approximations of the bias and skewness of the bootstrap distribution. To make these corrections, two values must be estimated: z_0 (the estimate of the bias) and a_c (the acceleration constant). z_0 is the *z*-score for the proportion of bootstrap estimates that are less than or equal to the sample estimate of the effect. In this case, z_0 is

$$z_0 = \phi^{-1}\left(\frac{\sum_{j=1}^{J} ab'_j \leq ab}{J}\right) \tag{2.4}$$

where *J* is the number of bootstrap samples, ab'_j is the jth bootstrap estimate of the indirect effect, and ϕ^{-1} is the inverse of a standard normal cumulative distribution function (CDF). Since this serves as an estimate of the bias, the resulting value will be zero if there is no bias (i.e., half of the bootstrap estimates are below *ab* and half are above; $z_0 = \phi^{-1}(.5) = 0$).

The acceleration constant in BCa confidence intervals approximates the skewness of the distribution. More specifically, BCa intervals utilize jackknife resampling to estimate the rate of change of the standard error of the statistic (*ab* here) relative to the true parameter value (Efron & Tibshirani, 1993). The acceleration parameter a_c can be can calculated by

$$a_c = \frac{\left(\sum_{i=1}^{n}(\overline{ab} - ab_{-i})\right)^3}{6\left[\sum_{i=1}^{n}(\overline{ab} - ab_{-i})^2\right]^{\frac{3}{2}}} \tag{2.5}$$

where ab_{-i} is the estimate of *ab* with case *i* removed, and $\overline{ab}$ is the mean of the *n* jackknife estimates.

In order to obtain the limits for the BCa confidence intervals, calculate the upper and lower bounds

$$z_{lower} = z_0 + \frac{z_0 + z_{\frac{\alpha}{2}}}{1 - a_c\left(z_0 + z_{\frac{\alpha}{2}}\right)} \tag{2.6}$$

and

$$z_{upper} = z_0 + \frac{z_0 + z_{1-\frac{\alpha}{2}}}{1 - a_c\left(z_0 + z_{1-\frac{\alpha}{2}}\right)} \tag{2.7}$$

where α is the specified error rate, $z_{\frac{\alpha}{2}}$ is the value for the $\left(\frac{\alpha}{2}\right)100$ percentile of the *z*-distribution and $z_{1-\frac{\alpha}{2}}$ is the value for the $\left(1-\frac{\alpha}{2}\right)100$ percentile of the *z*-distribution. The percentiles for the endpoints of the confidence interval are then calculated as $\phi(z_{lower})$ and $\phi(z_{upper})$ where ϕ indicates the standard normal CDF.

As an illustration, say z_0 was estimated to be 0.45 and a_c was estimated to be 0.15. For a 95% BCa CI, the percentile for the lower endpoint would be $\phi\left(z_0 + \frac{z_0 + z_{\frac{\alpha}{2}}}{1-a_c\left(z_0 + z_{\frac{\alpha}{2}}\right)}\right) = \phi\left(0.45 + \frac{0.45+(-1.96)}{1-0.15(0.45+(-1.96))}\right) = \phi(-0.781) = .217(100) = 21.7$

and the percentile for the upper endpoint would be $\phi\left(z_0 + \frac{z_0 + z_{1-\frac{\alpha}{2}}}{1-a_c\left(z_0 + z_{1-\frac{\alpha}{2}}\right)}\right) =$

$\phi\left(0.45 + \frac{0.45+1.96}{1-0.15(0.45+1.96)}\right) = \phi(4.224) = .999(100) = 99.9$. Thus, for j = 1,000 bootstrap samples, the limits of the 95% BCa CI would be 217th and 1,000th ordered values. The BC confidence intervals are computed largely the same way with the assumption that the acceleration parameter $a_c = 0$.

If bias and acceleration are zero, this will result in the same percentiles as the percentile bootstrap CI. That is, the percentile for the lower endpoint is $\phi\left(0 + \frac{0+(-1.96)}{1-0(0+(-1.96))}\right) = \phi(-1.96) = .025(100) = 2.5$ and the percentile for the upper endpoint is $\phi\left(0 + \frac{0+1.96}{1-0(0+1.96)}\right) = \phi(1.96) = .975(100) = 97.5$.

2.2 The Parallel Mediation Model

In reality, it is likely most (if not all) effects operate through more than one mechanism. It is unlikely that something as complex and variable as human cognition/behavior can be adequately described by a three-variable system. One mediator is simply not enough to explain most real-world processes—especially social phenomena. The parallel mediation model (depicted in Figure 2.2) allows mediators to correlate with one another but does not allow them to influence each other causally. Instead of the equations in the simple mediation model, the indirect, direct, and total effects can be derived through the following equations

$$M_i = d_{M_i} + a_i X + \varepsilon_{M_i} \qquad (2.8)$$

$$Y = d_Y + c'X + \sum_{i=1}^{k} b_i M_i + \varepsilon_Y \qquad (2.9)$$

$$c = c' + \sum_{i=1}^{k} a_i b_i \qquad (2.10)$$

where *k* is the number of mediators. Indirect effects of *X* on *Y* are constructed by multiplying a_i and b_i (for all *i* = 1,…,*k*). For example, with two mediators there are two indirect effects: a_1b_1 and a_2b_2. Each of these indirect effects is called a *specific indirect effect*. The interpretation of this effect is *not* the indirect effect of *X* on *Y* through M_i. It is the indirect effect of *X* on *Y* through M_i controlling for the other M_{k-1} mediators. Thus, you're accounting for the correlation between mediators. Accounting for this shared association is why mediation scholars recommend including all mediators in one single model as opposed to *k* independent simple mediation models (even if you suspect they're independent before running analyses; VanderWeele & Vansteelandt, 2014).

Parallel mediation models are commonly estimated by substantive researchers. For example, Wiedow et al. (2013) found that learning as a team significantly improved team outcomes (e.g., individual and group performance). This relationship was mediated by both increased task knowledge and trust in other teammates. Another study found that cancer patients who were able to find benefits in their condition receive more support from friends and family, have greater acceptance of their condition, and have greater acceptance of their emotions. The increased levels of social support, acceptance of their emotions, and acceptance of their cancer resulted in lower levels of depression (Manne et al., 2018).

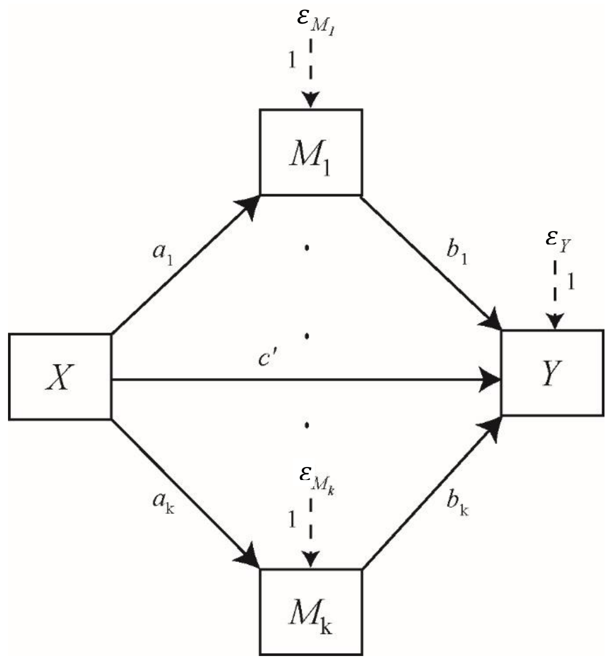

Figure 2.2: A statistical diagram for a parallel multiple mediator model with *k* mediators. *ε* denotes residuals in estimation.

When conducting mediation analysis with *k* mediators in a model, it is possible to estimate what is called the *total indirect effect*. The total indirect effect is the sum of all *k* indirect effects in the model. Although there are cases where you may be interested in testing whether a relationship operates through a set of mechanisms, the total indirect effect is generally not a useful statistic. Consider the case where the specific indirect effects are equal in size but opposite in sign. Here, the total indirect effect will be zero even though the specific indirect effects are different from zero. Just like a significant total effect is not a precondition of meditation, neither is a significant total indirect effect. Testing the total indirect effect may be useful in cases where mechanisms operating in the same direction are correlated with one another (where the size of the specific indirect effects can change when controlling for the other mediators), but this is one particular case and it is up to the researcher to decide whether this is theoretically interesting. I would argue that it is usually the specific indirect effects in the model that one is interested in testing.

Although the total indirect effect is typically not interesting, testing the difference between specific indirect effects in a model can be very interesting. In the next chapter, I will discuss why comparing indirect effects can be important and theoretically useful.

Chapter 3: Comparing Indirect Effects

Since a causal effect can operate through multiple mechanisms, a question to ask is whether a particular mediator is more influential in explaining a causal process than another. This question can be answered by comparing indirect effects. This section will discuss a contrast of indirect effects in a parallel multiple mediator model with two mediators, but the discussion generalizes to more complex models with many specific indirect effects, resulting in many possible comparisons between them.

3.1 The Importance of Contrasts

Contrasts are important for testing and building substantive theory. For an example of why this might be important, consider the case of substance abuse and intervention research. Mediation analysis is common in these areas of study because it helps identify the mechanisms that lead to behavior change after an intervention (MacKinnon, 2000). If certain mechanisms lead to a greater change in a target behavior than others, then these intervention programs can focus on the more effective pathways and eliminate the ineffective (or less effective) ones, thus improving the quality of treatment and likelihood of behavior change.

Or consider a case where two competing theories explain how optimism causes happiness. One theory might posit that optimism leads to resilience, and this increased resilience is what causes a person to be happy. A different theory might argue that optimism leads to increased self-trust, and this enhanced self-trust is what causes happiness to increase. If inference of the contrast of two indirect effects quantifying these theories suggested that the effect through resilience was larger than the effect through self-trust, the theory about resilience has more support than the theory regarding self-trust.

3.2 Using a Statistical Test

It is important to carry out a formal statistical test and not fall victim to the fallacy that a difference in the significance of two indirect effects means that they are significantly different (Gelman & Stern, 2006). In this "naïve" approach to comparing indirect effects, researchers test the significance of specific indirect effects—if indirect effect A is deemed significantly different from zero but indirect effect B is not, the researcher would conclude that indirect effect A is larger or more important in explaining the relationship than indirect effect B. It is not uncommon to find examples of this approach (Lapointe et al., 2012; Rudy, Davis, & Matthews, 2012). Some researchers have even conducted contrasts between indirect effects that were/were not significant and dismissed the usefulness of such a comparison based on this logic (Yap & Baharudin, 2016).

In a perfect world where inferential tests were always accurate in identifying zero or nonzero effects this reasoning may hold up, but in a world filled with sampling variability and other complexities, you cannot deem two effects to be different without a statistical test. This is especially important since the null hypothesis is never confirmed in methods used by researchers. It's merely that zero can't be ruled out as a possibility and consequently we cannot confirm the alternative hypothesis. Even if this weren't an issue, it's still important to test whether two indirect effects that are significantly different from zero are significantly different from each other. And indirect effects can be compared even if neither is statistically significant—testing whether a_1b_1 = 0 and a_2b_2 = 0 is not the same as testing whether $a_1b_1 - a_2b_2$ = 0. If there were a series of studies measuring indirect effects about a treatment program where neither indirect effect was different from zero but the effect through mediator A was consistently bigger than the effect through mediator B, this is still important information to know and it can change the direction and focus of future treatments.

3.3 The Raw Difference

Defining a contrast as the difference between two indirect effects seems like the most intuitive approach. Indeed, this method has been used by a great number of substantive researchers (Romero-Moreno, Losada, Márquez-González, & Mausbach, 2016; Schotanus-Dijkstra, Pieterse, Drossaert, Walburg, & Bohlmeijer, 2019; Yıldız, 2016). To carry out this approach, simply calculate the difference between two indirect effects

$$a_1b_1 - a_2b_2 \tag{3.1}$$

MacKinnon (2000) discussed inference for the difference between two indirect effects using the multivariate delta method. This method assumes the sampling distribution of the difference of two indirect effects is normally distributed. With this assumption, the estimate of the difference can be divided by its standard error and the resulting value can be compared to a standard normal distribution for inference. The formula for the standard error of a difference between indirect effects is quite complex and its accuracy as an estimator of the standard error of the difference of indirect effects requires meeting many assumptions.

But assumptions don't need to be made about the sampling distribution of the difference and it doesn't need to be difficult to compare two indirect effects. We can also conduct an inference for the difference between two indirect effects using bootstrapping. For inference, calculate the difference in each of the *j* bootstrap samples and construct a bootstrap confidence interval at a desired level of confidence in the same way we constructed a confidence interval for inference about the indirect effect (using either percentile, BC, or BCa confidence intervals). If the confidence interval does not contain zero, then the indirect effects are deemed different from each other. Testing the raw difference of two indirect effects answers a question about whether they are equal in value.

Some may wonder if indirect effects can be compared if the mediators are measured on different scales. There are no problems with doing this since the scaling of the mediators falls out in the computation of specific indirect effects (MacKinnon, 2000; Preacher & Hayes, 2008). For a mathematical representation of this, consider that a regression coefficient can be calculated as a function of correlations and standard deviations of the pertinent variables (Cohen, Cohen, West, & Aiken, 2003). Preacher and Hayes (2008) derived the following formulas for a_1 and b_1 in a two-mediator model

$$a_1 = r_{XM_1}\left(\frac{SD_{M_1}}{SD_X}\right) \quad (3.2)$$

$$b_1 = \frac{r_{XY}(r_{XM_2}r_{M_1M_2} - r_{XM_1}) + r_{M_1Y}(1 - r_{XM_2}^2) + r_{M_2Y}(r_{XM_1}r_{XM_2} - r_{M_1M_2})}{1 - r_{XM_1}^2 - r_{XM_2}^2 - r_{M_1M_2}^2 + 2r_{XM_1}r_{XM_2}r_{M_1M_2}}\left(\frac{SD_Y}{SD_{M_1}}\right) \quad (3.3)$$

When you multiply a_1 and b_1, the scaling of M_1 (SD_{M_1}) cancels out. All that remains are the correlation terms and the standard deviations of X and Y $\left(\frac{SD_Y}{SD_X}\right)$. The same applies for M_2 when constructing a_2b_2. Thus, if two indirect effects share the same scaling for X and Y, they can be (statistically) meaningfully compared.

Chapter 4: Comparing Opposing Indirect Effects

A potential problem arises with using the raw difference described in equation 3.1 when the indirect effects are different in sign. Consider a situation where two indirect effects are of the same magnitude but different in sign. Recall that in section 1.1, Pitts et al. (2018) posited that combat experience would increase both perceived threats to life and perceived benefits of deployment upon return to civilian life. They also argued that perceived threats to life would *increase* posttraumatic stress symptoms and perceived benefits from deployment would *decrease* posttraumatic stress symptoms. The resulting indirect effects are of opposing signs.

Suppose the effect through perceived benefit, $a_1 b_1$, was –0.60 and the effect through perceived threat, $a_2 b_2$, was 0.60. It's not as if one of these is necessarily more important than the other in explaining the relationship between combat experiences and PTSD, they are merely operating in different directions. A test of the raw difference, however, would result in a point estimate of –1.20 by equation 3.1. If the confidence interval for this difference didn't contain zero, one would correctly conclude the indirect effects are different—they are just different in *value*. But they are equal in *strength*. To avoid this, if one wants to test whether two indirect effects are different in strength when they are of different signs, a different method must be used.

Testing whether opposing indirect effects are different in strength not only improves conceptual/theoretical understanding of the mechanisms measured in a study, it can also help researchers and practitioners improve treatment programs. Consider a program such as alcoholics anonymous (AA). An AA program can attempt to reduce alcohol consumption through a number of different mediators such as increased self-efficacy and reduced depression (O'Rourke, & MacKinnon, 2018). Sometimes intervention programs can have unintended effects on behavior, such as physical activity interventions which can increase BMI as a result of caloric intake (Cerin & MacKinnon, 2009) and drug-prevention programs that might inform participants of more reasons to use a drug (MacKinnon, 2000).

Assume self-efficacy and reasons to drink were measured as part of an evaluation of an AA program. When looking at the effect through self-efficacy, the indirect effect ($a_1 b_1$) is negative—the AA program would increase self-efficacy (a_1 > 0), and this increased self-efficacy would decrease drinking behavior (b_1 < 0). When looking at the effect through reasons to drink, the indirect effect ($a_2 b_2$) could be positive (the program may increase the reasons one may choose to drink, a_2 > 0, and this increase in reasons would increase drinking behavior, b_2 > 0). Both mechanisms affect drinking behavior, but they are operating in different directions. Comparing such indirect effects could allow practitioners to modify an AA program to make them more effective in reducing how much one drinks. We cannot draw such conclusions using the raw difference in cases like this and consequently must use an alternative approach.

The literature on comparing indirect effects of opposite signs is sparse. Other than a recommendation by Hayes (2018) suggesting one calculate the difference of the absolute values of the indirect effects when they are of opposing signs and Ter Hoeven, van Zoonen, and Fonner (2016) who added two indirect effects of opposing signs, no special measures have been taken or recommendations have been given to compare indirect effects of different signs. However, there are a number of possibilities for substantive researchers to compare the strength of indirect effects that differ in sign and comparing these approaches is the purpose of this thesis.

4.1 The Difference in the Absolute Values

One possible approach when the indirect effects are of different signs involves using the absolute values of the indirect effects. To carry out this approach, calculate the difference in the absolute values

$$|a_1 b_1| - |a_2 b_2| \tag{4.1}$$

For inference, do this for each bootstrap sample, construct a bootstrap confidence interval, and if the confidence interval for the difference does not contain zero, then the two effects are deemed different from each other. The difference in absolute values approach answers a question about the equality of *magnitude/strength* of two indirect effects as opposed to whether they are equal in value.

Hayes (2018) recommends using the raw difference approach if the indirect effects are of the same sign and using the difference of the absolute values approach if they are of different signs. This creates a “conditional” test where—based on the observed signs of the indirect effects—either the raw difference or difference of absolute values will be calculated. (This is called the "conditional difference approach" because which method you use is conditional on the signs of the observed indirect effects.)

4.2 The Sum

An interesting contrast for comparing two indirect effects that are different in sign was used by Ter Hoeven, van Zoonen, and Fonner (2016). In their model, they proposed adding two indirect effects of different signs together as a form of contrast. This approach makes sense when the indirect effects are of opposing signs because if they were of the same strength, you would expect their sum to equal zero. Thus, you can carry out this method by calculating the sum of the indirect effects

$$a_1 b_1 + a_2 b_2 \tag{4.2}$$

For inference, do this for each bootstrap sample, construct a bootstrap confidence interval, and if the resulting interval doesn’t contain zero, then the indirect effects are deemed different in strength from each other.

4.3 The Ratio

The ratio of two effects takes into account both sign and magnitude. To test whether two indirect effects are different using this approach, calculate the ratio of two indirect effects

$$\frac{a_1 b_1}{a_2 b_2} \tag{4.3}$$

For inference, do this for each bootstrap sample, construct a percentile bootstrap confidence interval, and if the resulting interval does not contain -1 or 1, this provides evidence that the two indirect effects are different in strength and value.

The biggest issue with this approach is the presence of a zero in the denominator. In equation 4.3, if $a_2 b_2$ is zero then the result is undefined and impossible to interpret. You could possibly flip the indirect effects for such cases, but this fixes nothing if *both* indirect effects are zero. And if the numerator is zero and the denominator is nonzero, the ratio is 0. This could lead to a confusing interpretation in the context of the problem if one of the indirect effects is zero. It is mathematically but not substantively meaningful.

Even if one of the indirect effects is *close* to zero, the ratio can be infinitely small or large and the same issues exist when trying to interpret the results. Calculating the ratio with small numbers can lead to numerical stability issues/strange bootstrap distributions. In the example we will use in Chapter 5, the calculated ratio of the indirect effects is –0.443. The bootstrap ratio values in 10,000 resamples of the data produces a range between [–26, 48]. (This range can get even larger when the values in the numerator and/or denominator get smaller.) Additionally, if during the bootstrap algorithm one of the ratio calculations is undefined, it could stop the computations or generate other errors if the software is not equipped to handle such values.

4.4 The Ratio of Absolute Values

The point estimate that results from the ratio can range from $(-\infty, \infty)$. A modification of equation 4.3 involves calculating the ratio of the absolute values of the indirect effects

$$\frac{|a_1 b_1|}{|a_2 b_2|} \tag{4.4}$$

For inference, do this for each bootstrap sample, construct a bootstrap confidence interval, and if the resulting interval does not contain one, the two indirect effects are deemed different in strength. The difference here is that the point estimate is now bounded from [0, ∞). The presence of a zero or near-zero value in the numerator or the denominator still poses a problem, however.

4.5 The Proportion of Absolute Values

It is possible to further bound the estimate in equation 4.4. By calculating the proportion of the absolute value of one indirect effect to the sum of the absolute values of both indirect effects, the estimate is bound between zero and one. To conduct this test, calculate

$$\frac{|a_1 b_1|}{|a_1 b_1| + |a_2 b_2|} \tag{4.5}$$

For inference, do this for each bootstrap sample, construct a bootstrap confidence interval, and if the resulting interval does not contain .5, this provides evidence that the two indirect effects are different in strength (if the interval is entirely below .5, $|a_1 b_1|$ is deemed smaller, and if the interval is entirely above .5, $|a_1 b_1|$ is deemed larger). Equation 4.5 is contingent on neither the numerator nor denominator equaling zero for the reasons listed above.

This method should produce identical inference to the ratio of absolute values since these point estimates are transformations of each other. To demonstrate why this is, say $|a_1 b_1|$ is 8 and $|a_2 b_2|$ is 2. Plugging these values into equation 4.4. produces a value of 4 whereas equation 4.5 produces a value of .8. However, if we take $\frac{.8}{(1-.8)}$ we get the same value from equation 4.4 (i.e., 4). This is true for any two values you plug into the equations. I would argue equation 4.5 produces a cleaner estimate than equation 4.4 (the values approach 0 and 1 instead of 0 and infinity), and since they provide the same information about the contrast of two indirect effects, I would recommend equation 4.5.

Chapter 5: War Veterans and Posttraumatic Stress Disorder

To illustrate the concepts in this thesis, I will step through the example with war veterans introduced in chapter one. Throughout time, the health of veterans returning from war has been of interest to researchers, citizens, and of course, veterans and active military personnel. Understanding what leads to PTSD and other negative outcomes like high suicide rates is of the utmost importance for those wishing to better the lives of veterans who have given a part of theirs to help protect the country. Mediation analysis can help reveal how combat experiences lead to posttraumatic stress symptoms. Recall that the X was combat experience (COMBAT in Figure 5.1), the Ms were perceived threats to life (PTHREAT) and perceived benefits from deployment (BENEFITS), and the Y was posttraumatic stress symptoms (PTSD). Social desirability (SOCIAL) was a covariate in all model equations. See Figure 5.1 for a conceptual diagram. The data come from a questionnaire administered to 324 Army combat medics.

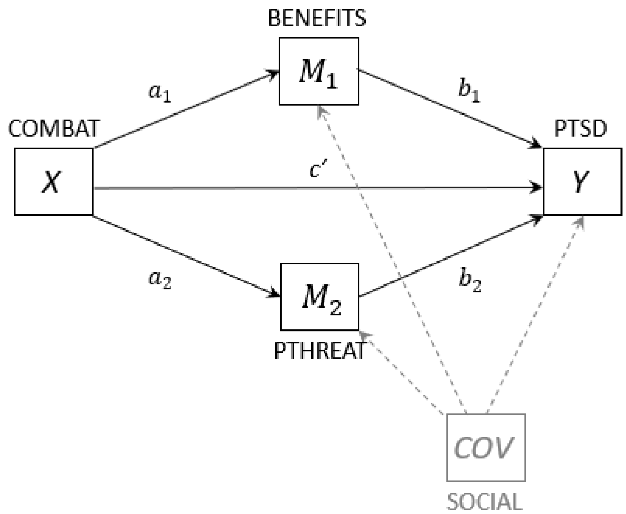

Figure 5.1 A conceptual diagram for the war veteran study.

5.1 Estimating the Specific Indirect Effects

The researchers estimated a parallel multiple mediation model controlling for social desirability. I will discuss the individual *a* and *b* paths and the specific indirect effects and their confidence intervals. See Table 5.1 for full model results. Assuming all effects are causal in nature (which is up for debate since the data are not experimental in nature), combat experiences significantly increased both the perceived benefits of the deployment and perceived threats to life (a_1 = 0.19; a_2 = 0.69). As expected, the more benefits one perceived of their deployment, the fewer PTSD symptoms they reported (b_1 = –0.38). When they reported perceiving more threats to their life, they also reported more PTSD symptoms (b_2 = 0.23). A total of 10,000 resamples were used to construct bootstrap confidence intervals for inference (only percentile bootstrap confidence intervals are reported for simplicity). The specific indirect effect through perceived benefit is negative and statistically significant ($a_1 b_1$ = 0.19(–0.38) = –0.07, 95% CI = [–0.134, –0.025]).

Thus, combat experiences *decrease* PTSD symptoms indirectly through perceived benefits of deployment. The specific indirect effect through perceived threat is significant ($a_2 b_2$ = 0.69(0.23) = 0.16, 95% CI = [0.062, 0.280]). Combat experience also *increase* PTSD symptoms indirectly through perceived threats to life. Here, combat experiences operate through both mechanisms *simultaneously* to cause a decrease (and increase) in PTSD symptoms. In this case both indirect effects are significant, but this is not necessary to go on to the next step: contrasting indirect effects.

Table 5.1 Model results from the war veteran study

Note. All values are for indirect effects of opposing signs. The 95% CI for indirect effects is a percentile bootstrap confidence interval constructed from a total of 10,000 bootstrap samples.

Outcome	Predictor	Effect	Value (SE)	p	95% CI
BENEFIT	COMBAT	a_1	0.19 (0.053)	< .001	[0.086, 0.294]
THREAT	COMBAT	a_2	0.691 (0.091)	< .001	[0.511, 0.871]
PTSD	BENEFIT	b_1	–0.378 (0.119)	.002	[–0.613, –0.144]
	THREAT	b_2	0.235 (0.069)	.001	[0.099, 0.370]
	COMBAT	c'	0.335 (0.124)	.007	[0.092, 0.579]
-	-	$a_1 b_1$	–0.072 (-)	-	[–0.134, –0.025]
-	-	$a_2 b_2$	0.162 (-)	-	[0.062, 0.280]

5.2 Obtaining Contrasts

Every contrast method discussed in this thesis was used to compare $a_1 b_1$ and $a_2 b_2$. Table 5.2 contains the full results and Figure 5.2 show the bootstrap distributions for each contrast method. The raw difference suggests that the indirect effects are different in value ($a_1b_1 - a_2b_2 =$ –0.23; 95% CI = [–0.352, –0.133]). In all other methods, there is not sufficient evidence to deem the indirect effects different in strength (the intervals either contain 0, 0.5, 1, or –1 depending on the method). This highlights the importance of using the correct method for the research question of interest. If the question is whether two indirect effects are equal in value, the raw difference would be appropriate in this situation. This is often not the question, however. Usually researchers are interested in determining whether indirect effects are different in strength. In such cases, you would use one of the alternative methods discussed in this thesis to compare these opposing indirect effects. It would be incorrect to say that a_1b_1 is less important in explaining PTSD symptoms than a_2b_2 by using a test of the raw difference because the indirect effects are of opposing signs. These results suggest that it is equally important to teach deployed soldiers to seek out personal benefits from their deployment and perceive fewer threats to their life to ameliorate PTSD upon return form war (though the latter is largely out of the soldier's control).

Figure 5.2 reveals interesting details about the contrast methods. The distributions in Panels A, B and C—the distributions for the raw difference, the difference of absolute values, and the sum—appear approximately normal. In Panels D and F, the distributions for the ratio of absolute values and the proportion of the sum of absolute values are positively skewed. And in Panel E, the distribution for the ratio is highly negatively skewed. The distributions for the ratio and the ratio of absolute values appear to suffer from the problems discussed in chapter 4 (i.e., they have wide ranges and tall peaks).

Although bootstrapping does not make assumptions about the shape of the sampling distribution, the coefficients estimated in the BC and BCa CIs may be large (this will be something to consider when interpreting the simulation results in Chapter 6).

Table 5.2 Model results from the war veteran study

Note. *The conditional contrast and difference of absolute values are the same because the observed specific indirect effects are of opposing signs. The 95% CI for contrasts is a percentile bootstrap confidence interval constructed from 10,000 bootstrap samples.

Contrast	Null (H_0)	Value	95% CI
$a_1b_1 - a_2b_2$	= 0	–0.234	[–0.352, –0.133]
$\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	= 0	–0.090	[–0.231, 0.043]
Conditional Contrast*	= 0	–0.090	[–0.231, 0.043]
$a_1b_1 + a_2b_2$	= 0	0.090	[–0.043, 0.231]
$\frac{a_1b_1}{a_2b_2}$	= ± 1	–0.443	[–1.585, –0.113]
$\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	= 1	0.443	[0.117, 1.610]
$\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	= .50	.310	[.104, .617]

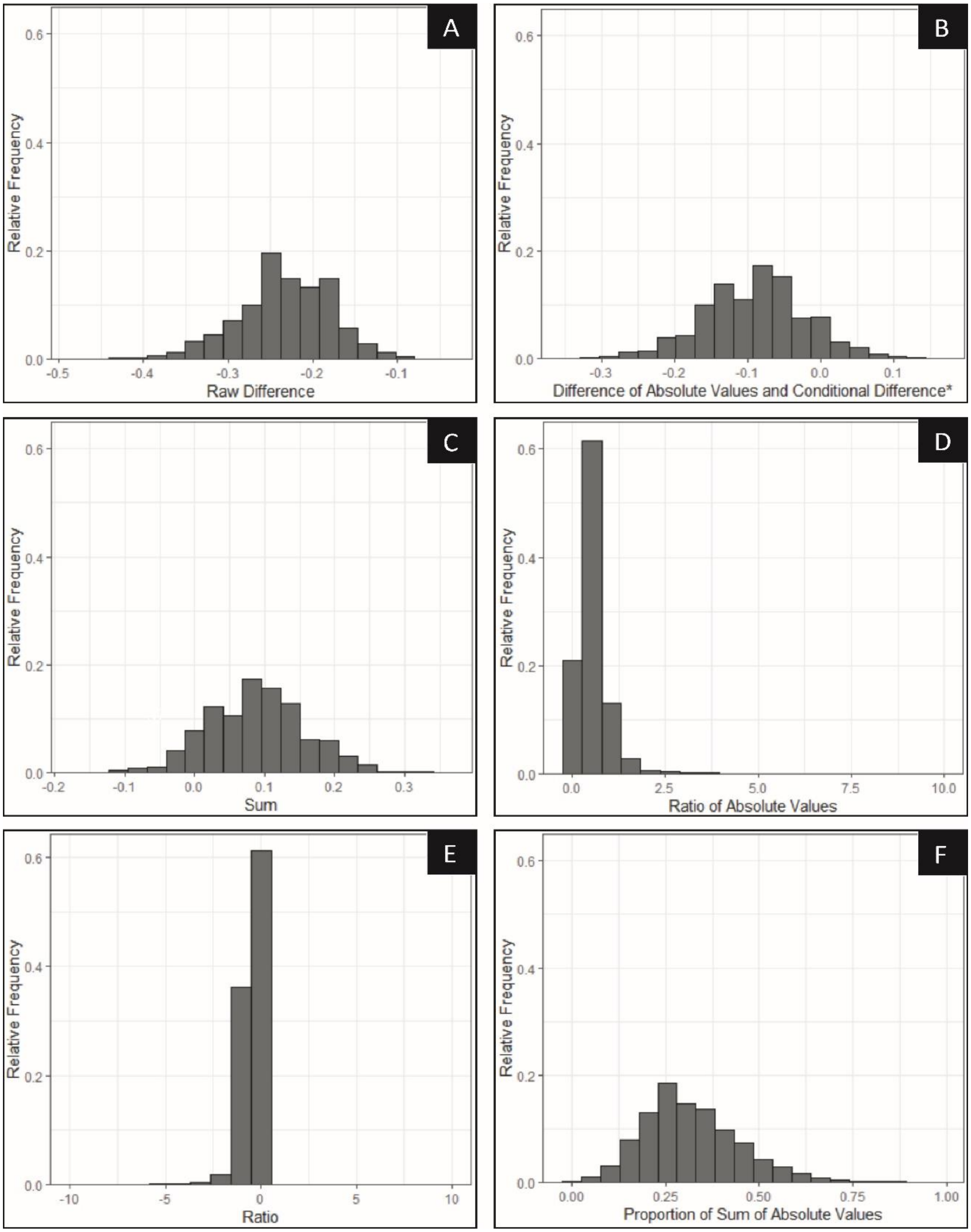

Figure 5.2: Bootstrap distributions for all contrast methods in the war veteran study. The conditional difference and difference of absolute values are the same because the indirect effects being compared are of opposing signs. The max for each histogram is the same.

Chapter 6: A Simulation Study Comparing Contrast Methods

As I've discussed, more than one mediator can (and probably should) be included in a mediation model and the resulting indirect effects can (and probably should) be compared with one another. But what method (of the ones presented in this thesis) should be used to compare indirect effects? Should any of these approaches be used at all? Is the usefulness of the approach contingent on sample size or the size of the indirect effects? All of these questions were taken into account when designing this simulation.

6.1 Setting up the Simulation

The simulation study was conducted in GAUSS 20. A parallel multiple mediator model with two mediators was the basis of the simulation. Although not representative of every possibility, the following sample sizes and parameter values were chosen to cover a wide range of scenarios and are commonly used in the mediation simulation literature (Fritz & MacKinnon, 2007; Hayes & Scharkow, 2013). Seven possible values (–0.59, –0.39, –0.14, 0, 0.14, 0.39, 0.59) were used for a_1, a_2, b_1, and b_2 to simulate a wide range of sizes and signs of indirect effects. A direct effect of zero was used in all conditions. To resemble a range of sample sizes used by substantive researchers, five different values ranging from small to large (*n* = 25, 50, 100, 200, and 500) were used. Crossing all combinations of the above values results in 12,005 different combinations/conditions (7×7×7×7×5). There were a total of 2,000 repetitions per condition.

The data for *X* were sampled from a standard normal distribution and the *M*s were generated from *X* using the corresponding a paths in the model for the *M*s (see equation 2.8; e.g., a_1 = 0.59, a_2 = 0.14, and both of these values were multiplied by *X* to generate data for M_1 and M_2). Standard normal errors were added to M_1 and M_2 to generate sampling variance (ε_{M_i} in equation 2.8). *Y* was generated from the *M*s and *X* using both bs and c' in the model for *Y* (see equation 2.9; e.g., b_1 = 0.39, b_2 = –0.14, c' = 0, and these values were multiplied by the data for M_1, M_2, and *X* respectively to generate values for *Y*). Standard normal errors were also added to *Y* in each repetition. These served as the data. In each repetition, a total of 5,000 bootstrap samples were constructed and the indirect effects a_1b_1 and a_2b_2 were calculated in each bootstrap sample and compared using the difference in absolute values, the sum, the ratio, the ratio of the absolute values, the proportion of the absolute value of a_1b_1 to the sum of the absolute values, and the conditional method proposed by Hayes (2018). This resulted in 5,000 differences in each repetition where percentile, BC, and BCa bootstrap confidence intervals were constructed for inference.

6.2 Simulation Results

In each of the conditions (e.g., a_1 = 0.59, a_2 = 0.14, b_1 = 0.39, b_2 = 0.14, *n* = 25), the number of times the 95% confidence interval covered the null difference and the number of times the 95% confidence interval covered the true value of the contrast were divided by the number of repetitions (2,000) and multiplied by 100 to calculate the percentage for each method's coverage rate for the corresponding *a*s, *b*s, and *n*. The null difference is a term used for the value that indicates the indirect effects are equal in a contrast method (e.g., 0 for the difference of absolute values, 1 for the ratio of absolute values).

6.2.1 Coverage of the true difference

First, it is important to test how often each method produces a confidence interval that contains the true value of the contrast relative to the level of confidence specified by the researcher. For 95% confidence intervals, we would expect 95% of the confidence intervals for a method to contain the true value. The percentage of times the confidence intervals contain the value of interest is called the coverage rate. Thus, for this simulation study, the coverage rate of each method (and type of confidence interval) should be approximately 95(%). If the coverage is greater than 95, the test is containing the value too often and is considered statistically *conservative* (the intervals are too wide for the confidence level specified). If the coverage is less than 95, the intervals aren't containing the true value enough and the method would be statistically *liberal* (the intervals are too narrow for the confidence level specified). Table 6.1 contains the coverage rates for each contrast method for each type of confidence interval at a variety of sample sizes.

Table 6.1 Overall coverage of true contrast value for each method

Note. BC = Bias-corrected. BCa = Bias-corrected and accelerated. CI = Confidence interval. Coverage is displayed as the % of times the method contains the true contrast value. The conditional contrast method is excluded from this table as the true contrast value varies.

CI Method	Contrast Method	Sample Size				
		25	50	100	200	500
Percentile	1. $\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	96.3	95.2	94.6	93.6	93.1
	2. $a_1b_1 + a_2b_2$	96.1	95.3	95.1	95.0	95.0
	3. $\frac{a_1b_1}{a_2b_2}$	72.1	71.9	71.4	71.0	70.8
	4. $\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	53.0	52.7	52.5	52.0	51.4
	5. $\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	53.0	52.7	52.5	52.0	51.4
BC	1. $\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	90.8	90.8	90.7	91.6	93.2
	2. $a_1b_1 + a_2b_2$	95.0	94.5	94.6	94.7	94.8
	3. $\frac{a_1b_1}{a_2b_2}$	59.7	63.1	65.7	68.4	70.2
	4. $\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	43.4	45.0	46.2	48.5	50.7
	5. $\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	43.4	45.0	46.2	48.5	50.7
BCa	1. $\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	92.9	92.0	90.9	90.6	91.5
	2. $a_1b_1 + a_2b_2$	92.9	92.6	92.6	92.7	92.7
	3. $\frac{a_1b_1}{a_2b_2}$	56.7	59.5	62.2	65.4	67.4
	4. $\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	42.7	44.2	45.3	47.4	49.6
	5. $\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	43.9	44.5	44.7	46.4	48.6

Consistent with prior research, the percentile confidence intervals tended to have better coverage than the BC and BCa confidence intervals (i.e., the percentile confidence intervals were closer to 95%; Fritz, Taylor, & MacKinnon, 2012; Hayes & Scharkow, 2013). Additionally, as sample size increased, coverage rates tended to approach 95% for each method. (Although sometimes coverage *decreased* with an increase in sample size; e.g., the difference of the absolute values for percentile CIs.) The sum and difference of absolute values performed the best of all the methods, though it should be noted that the coverage rates of the contrast methods involving division improved substantially when true indirect effects of size zero were excluded and when indirect effects were greater in magnitude (this is likely due to the problems identified in chapter 4). To get a better understanding of the performance of these methods, I will breakdown the coverage rates when there is and is not a true difference between the indirect effects.

6.2.2 Coverage of the null difference when the indirect effects are equal (same sign)

It is important to see how each method compares to the "traditional" approach for comparing indirect effects (i.e., calculating the raw difference). In order for this comparison to be meaningful, I isolated situations where the indirect effects were equal in value and sign and looked at the coverage rates for the null difference. Table 6.2 contains results for each sample size and confidence interval approach. (The sum is not in Table 6.2 since the indirect effects are of the same sign.)

Table 6.2 Correct coverage of the null difference when $a_1 b_1 = a_2 b_2$

Note. BC = Bias-corrected. BCa = Bias-corrected and accelerated. CI = Confidence interval. The raw difference is included as a reference and the sum is excluded because the indirect effects are of the same sign (and thus would not be used as a contrast).

CI Method	Contrast Method	Sample Size				
		25	50	100	200	500
Percentile	1. $a_1 b_1 - a_2 b_2$	97.2	96.1	95.5	95.3	95.2
	2. $\lvert a_1 b_1\rvert - \lvert a_2 b_2\rvert$	98.9	98.3	97.7	97.0	96.4
	3. Conditional Contrast	98.8	98.2	97.7	97.1	96.7
	4. $\frac{a_1 b_1}{a_2 b_2}$	98.3	97.7	97.0	96.3	95.9
	5. $\frac{\lvert a_1 b_1\rvert}{\lvert a_2 b_2\rvert}$	98.9	98.3	97.7	97.0	96.4
	6. $\frac{\lvert a_1 b_1\rvert}{\lvert a_1 b_1\rvert + \lvert a_2 b_2\rvert}$	98.9	98.3	97.7	97.0	96.4
BC	1. $a_1 b_1 - a_2 b_2$	95.7	94.7	94.4	94.6	94.8
	2. $\lvert a_1 b_1\rvert - \lvert a_2 b_2\rvert$	94.5	93.4	92.7	92.5	93.1
	3. Conditional Contrast	-	-	-	-	-
	4. $\frac{a_1 b_1}{a_2 b_2}$	72.2	74.0	75.3	76.9	78.1
	5. $\frac{\lvert a_1 b_1\rvert}{\lvert a_2 b_2\rvert}$	74.2	75.2	75.7	77.3	78.7
	6. $\frac{\lvert a_1 b_1\rvert}{\lvert a_1 b_1\rvert + \lvert a_2 b_2\rvert}$	74.2	75.2	75.7	77.3	78.7
BCa	1. $a_1 b_1 - a_2 b_2$	93.6	92.8	92.5	92.6	92.6
	2. $\lvert a_1 b_1\rvert - \lvert a_2 b_2\rvert$	96.6	95.5	94.3	93.3	93.1
	3. Conditional Contrast	-	-	-	-	-
	4. $\frac{a_1 b_1}{a_2 b_2}$	66.6	67.1	68.3	70.2	71.2
	5. $\frac{\lvert a_1 b_1\rvert}{\lvert a_2 b_2\rvert}$	73.5	74.2	74.8	76.3	77.6
	6. $\frac{\lvert a_1 b_1\rvert}{\lvert a_1 b_1\rvert + \lvert a_2 b_2\rvert}$	77.1	76.3	75.2	75.7	76.7

Once again, for all contrast methods, percentile CIs had the best coverage and coverage rates tended to approach 95% as sample size increased. (Subtracting the values in the table from 100 provides an estimate of the Type I error rate.) This is consistent with prior research which suggests that percentile confidence intervals have the lowest Type I error rate when compared to BC and BCa confidence intervals (Fritz, Taylor, & MacKinnon, 2012; Hayes & Scharkow, 2013). The contrast methods involving division had relatively poor coverage with BC and BCa confidence intervals. Since the bootstrap distributions for these approaches are highly skewed (see Figure 5.2), it could be that the BC and BCa intervals overcorrected for the issue (and thus had poorer performance as a result. See Section 6.2.3 for a further discussion of this.)

The raw difference appeared to perform the best here as it was the closest to 95% on average among all the contrast methods. It is worth noting that the difference of absolute values provided the same coverage as the ratio of absolute values and proportion of absolute values for the percentile CIs. This is because they provide the same information with respect to the null difference. Anytime the difference of absolute values produces a negative value (less than zero), the ratio of absolute values will be less than one (since the absolute value of a_1b_1 is less than the absolute value of a_2b_2) and the proportion of the absolute values will be less than .5. Anytime the difference of absolute values produces a positive value (greater than zero), the ratio of absolute values will be greater than one and the proportion of absolute values will be greater than .5. Thus, each bootstrap sample for the three methods will be concurrently above/below the null difference, and when we select the same percentiles for each method, the null difference will/will not equally be contained in each interval. This equality doesn't hold for the BC or BCa methods since the percentiles for the endpoints of the CI differ based on the bias and/or acceleration in the bootstrap distribution (and the distributions for the difference of absolute values and ratio/proportion are not the same). These three approaches and the ratio were close in performance to the raw difference but were more conservative for percentile CIs and more liberal for BC CIs.

6.2.3 Coverage of the null difference when the indirect effects are equal and nonzero (same sign)

The coverage rates of each method are affected when both of the indirect effects are zero. To account for this, Table 6.3 contains the results for the coverage of the null difference when a_1b_1 and a_2b_2 are equal, of the same sign, and nonzero. (The sum is still not in Table 6.3 since the indirect effects are of the same sign.)

Table 6.3 Correct coverage of the null difference when $a_1b_1 = a_2b_2 \neq 0$

Note. BC = Bias-corrected. BCa = Bias-corrected and accelerated. CI = Confidence interval. The raw difference is included as a reference and the sum is excluded because the indirect effects are of the same sign (and thus would not be used as a contrast).

CI Method	Contrast Method	Sample Size				
		25	50	100	200	500
Percentile	1. $a_1b_1 - a_2b_2$	96.4	95.4	95.1	94.9	95.0
	2. $\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	98.3	97.4	96.7	95.9	95.2
	3. Conditional Contrast	98.0	97.1	96.4	95.6	95.1
	4. $\frac{a_1b_1}{a_2b_2}$	97.9	97.5	96.9	96.4	96.4
	5. $\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	98.3	97.4	96.7	95.9	95.2
	6. $\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	98.3	97.4	96.7	95.9	95.2
BC	1. $a_1b_1 - a_2b_2$	94.8	94.1	94.2	94.5	94.7
	2. $\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	92.8	92.0	91.4	91.6	93.6
	3. Conditional Contrast	-	-	-	-	-
	4. $\frac{a_1b_1}{a_2b_2}$	80.2	85.3	89.2	93.0	95.7
	5. $\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	79.6	83.3	85.9	90.0	94.0
	6. $\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	79.6	83.3	85.9	90.0	94.0
BCa	1. $a_1b_1 - a_2b_2$	92.7	92.2	92.3	92.4	92.6
	2. $\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	95.0	93.5	92.2	91.0	91.7
	3. Conditional Contrast	-	-	-	-	-
	4. $\frac{a_1b_1}{a_2b_2}$	75.6	80.5	85.3	90.5	93.6
	5. $\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	78.3	81.7	84.4	88.3	92.0
	6. $\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	80.8	81.5	81.8	84.6	88.1

The coverage of each method was closer to 95% when both indirect effects weren't zero. This was particularly true for the methods involving division and BC/BCa confidence intervals. Recall the shape of the distributions in Figure 5.2—excluding zero indirect effects makes the average size of the indirect effects larger and the bootstrap distributions less bias/skewed. Thus, the bias and acceleration estimates would be smaller. It's possible these methods overcorrect for highly skewed distributions.

6.2.4 Coverage of the null difference when the indirect effects are equal (opposing signs)

Next, let's look at the performance of each method when the indirect effects are of opposing signs. Table 6.4 contains the coverage rates of the null difference for each contrast method for percentile, BC, and BCa confidence intervals at a variety of sample sizes. (The raw difference is not on this table since the indirect effects are of opposing signs.)

Table 6.4 Correct coverage of null difference when $-a_1 b_1 = a_2 b_2$

Note. All values are for indirect effects of opposing signs. BC = Bias-corrected. BCa = Bias-corrected and accelerated.

CI Method	Contrast Method	Sample Size				
		25	50	100	200	500
Percentile	1. $\lvert a_1 b_1 \rvert - \lvert a_2 b_2 \rvert$	98.9	98.3	97.7	97.0	96.4
	2. Conditional Contrast	98.8	98.2	97.7	97.2	96.7
	3. $a_1 b_1 + a_2 b_2$	97.0	96.1	95.6	95.3	95.2
	4. $\frac{a_1 b_1}{a_2 b_2}$	98.3	97.7	97.0	96.3	95.8
	5. $\frac{\lvert a_1 b_1 \rvert}{\lvert a_2 b_2 \rvert}$	98.9	98.3	97.7	97.0	96.4
	6. $\frac{\lvert a_1 b_1 \rvert}{\lvert a_1 b_1 \rvert + \lvert a_2 b_2 \rvert}$	98.9	98.3	97.7	97.0	96.4
BC	1. $\lvert a_1 b_1 \rvert - \lvert a_2 b_2 \rvert$	94.6	93.5	92.7	92.5	93.1
	2. Conditional Contrast	-	-	-	-	-
	3. $a_1 b_1 + a_2 b_2$	95.6	94.7	94.5	94.6	94.7
	4. $\frac{a_1 b_1}{a_2 b_2}$	72.2	73.9	75.4	77.0	77.9
	5. $\frac{\lvert a_1 b_1 \rvert}{\lvert a_2 b_2 \rvert}$	74.2	75.2	75.7	77.3	78.7
	6. $\frac{\lvert a_1 b_1 \rvert}{\lvert a_1 b_1 \rvert + \lvert a_2 b_2 \rvert}$	74.2	75.2	75.7	77.3	78.7
BCa	1. $\lvert a_1 b_1 \rvert - \lvert a_2 b_2 \rvert$	96.6	95.5	94.3	93.3	93.1
	2. Conditional Contrast	-	-	-	-	-
	3. $a_1 b_1 + a_2 b_2$	93.4	92.8	92.6	92.7	92.7
	4. $\frac{a_1 b_1}{a_2 b_2}$	66.5	67.2	68.3	70.2	71.1
	5. $\frac{\lvert a_1 b_1 \rvert}{\lvert a_2 b_2 \rvert}$	73.5	74.2	74.8	76.4	77.6
	6. $\frac{\lvert a_1 b_1 \rvert}{\lvert a_1 b_1 \rvert + \lvert a_2 b_2 \rvert}$	77.2	76.3	75.2	75.7	76.6

Percentile intervals again tended to have better coverage of the null difference (and thus a lower Type I error rate) than BC and BCa confidence intervals. BCa confidence intervals performed the worst on average across contrast method and sample size. The sum was the closest to 95% (except for BCa CIs where the difference of absolute values performed the best across sample sizes). Once again, the difference of absolute values, ratio of absolute values, and proportion contrast methods produced identical (and slightly conservative/liberal) inference.

6.2.5 Coverage of the null difference when the indirect effects are equal and nonzero (opposing signs)

The analyses in the previous section will be replicated but once again exclude indirect effects of size zero. Table 6.5 contains the coverage rates of the null difference for each contrast method for percentile, BC, and BCa confidence intervals at a variety of sample sizes. (The raw difference is still not on this table.)

Table 6.5 Correct coverage of null difference when $-a_1b_1 = a_2b_2 \neq 0$

Note. All values are for indirect effects of opposing signs. BC = Bias-corrected. BCa = Bias-corrected and accelerated.

CI Method	Contrast Method	Sample Size				
		25	50	100	200	500
Percentile	1. $\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	98.3	97.4	96.7	95.9	95.2
	2. Conditional Contrast	98.1	97.2	96.6	95.8	95.1
	3. $a_1b_1 + a_2b_2$	97.0	96.1	95.6	95.3	95.2
	4. $\frac{a_1b_1}{a_2b_2}$	96.3	95.4	95.1	95.0	95.0
	5. $\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	98.3	97.4	96.7	95.9	95.2
	6. $\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	98.3	97.4	96.7	95.9	95.2
BC	1. $\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	92.9	92.1	91.4	91.7	93.6
	2. Conditional Contrast	-	-	-	-	-
	3. $a_1b_1 + a_2b_2$	94.8	94.2	94.3	94.5	94.7
	4. $\frac{a_1b_1}{a_2b_2}$	80.3	85.4	89.3	93.1	95.7
	5. $\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	79.7	83.2	85.9	90.1	94.0
	6. $\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	79.7	83.2	85.9	90.1	94.0
BCa	1. $\lvert a_1b_1\rvert - \lvert a_2b_2\rvert$	95.1	93.6	92.2	91.1	91.8
	2. Conditional Contrast	-	-	-	-	-
	3. $a_1b_1 + a_2b_2$	92.6	92.3	92.4	92.5	92.6
	4. $\frac{a_1b_1}{a_2b_2}$	75.5	80.7	85.4	90.6	93.6
	5. $\frac{\lvert a_1b_1\rvert}{\lvert a_2b_2\rvert}$	78.4	81.6	84.4	88.4	92.0
	6. $\frac{\lvert a_1b_1\rvert}{\lvert a_1b_1\rvert + \lvert a_2b_2\rvert}$	80.9	81.5	81.8	84.7	88.1

Once again, the coverage of all methods tended to improve (i.e., be closer to 95%) when excluding cases where both indirect effects were zero. This was particularly true for the methods involving division and BC/BCa CIs (for the reasons discussed earlier). Additionally, as sample size increased, coverage tended to 95% (with a couple of exceptions for BC/BCa CIs).

6.2.6 Approximate power when the true difference is small

Correctly including the null difference in a confidence interval and avoiding Type I errors is important, but just as important is ensuring a method is correctly containing the true contrast value when a true difference exists (i.e., avoiding Type II errors). To approximate power for each contrast method, I will subtract the coverage rate of the null difference when a true difference exists from one. Table 6.6 contains the coverage rates of each method for percentile, BC, and BCa confidence intervals at a variety of sample sizes. In this instance, power will be calculated for "small" effects sizes (when $|a_1b_1| - |a_2b_2|$ is nonzero and between ± 0.10). This range is not based on statistical theory, but rather contains approximately half of the cases where there is a true difference (the rest of the differences will be $\geq\ \pm.10$ and considered moderate/large).

Table 6.6 Approximate power when the true difference is nonzero and small $|a_1b_1| - |a_2b_2| \leq \pm 0.10$

Note. Values are for contrasts between indirect effects of opposing signs. BC = Bias-corrected. BCa = Bias-corrected and accelerated. CI = Confidence interval.

CI Method	Contrast Method	Sample Size				
		25	50	100	200	500
Percentile	1. $\|a_1b_1\| - \|a_2b_2\|$	1.4	3.7	8.9	19.2	36.3
	2. Conditional Contrast	1.6	3.9	9.0	19.1	36.4
	3. $a_1b_1 + a_2b_2$	6.8	15.1	25.8	40.1	58.8
	4. $\frac{a_1b_1}{a_2b_2}$	3.6	8.6	16.3	26.1	40.9
	5. $\frac{\|a_1b_1\|}{\|a_2b_2\|}$	1.4	3.7	8.9	19.2	36.3
	6. $\frac{\|a_1b_1\|}{\|a_1b_1\| + \|a_2b_2\|}$	1.4	3.7	8.9	19.2	36.3
BC	1. $\|a_1b_1\| - \|a_2b_2\|$	7.2	11.8	18.9	29.4	46.1
	2. Conditional Contrast	-	-	-	-	-
	3. $a_1b_1 + a_2b_2$	9.4	18.0	28.2	41.8	59.8
	4. $\frac{a_1b_1}{a_2b_2}$	32.9	34.7	38.3	43.6	55.0
	5. $\frac{\|a_1b_1\|}{\|a_2b_2\|}$	27.7	28.3	30.3	35.1	47.5
	6. $\frac{\|a_1b_1\|}{\|a_1b_1\| + \|a_2b_2\|}$	27.7	28.3	30.3	35.1	47.5
BCa	1. $\|a_1b_1\| - \|a_2b_2\|$	4.2	7.7	14.4	24.4	40.1
	2. Conditional Contrast	-	-	-	-	-
	3. $a_1b_1 + a_2b_2$	11.1	18.0	27.0	38.7	54.1
	4. $\frac{a_1b_1}{a_2b_2}$	41.7	45.9	40.0	54.9	63.3
	5. $\frac{\|a_1b_1\|}{\|a_2b_2\|}$	28.6	29.3	31.1	34.8	44.4
	6. $\frac{\|a_1b_1\|}{\|a_1b_1\| + \|a_2b_2\|}$	24.6	27.5	32.3	37.9	47.3

In general, power was the lowest for percentile CIs and the highest for BCa CIs, consistent with prior research (Fritz, Taylor, & MacKinnon, 2012; Hayes & Scharkow, 2013). As expected, power increased for each contrast method and confidence interval type as sample size increased. The sum tended to have the greatest power among all contrast methods, especially with percentile CIs. The contrast methods involving absolute values had the lowest power across sample sizes and CI type.

6.2.7 Approximate power when the true difference is moderate to large

Finally, let's look at cases where the true difference between indirect effects is "moderate to large." Recall that the definition of moderate to large is arbitrary and approximately the halfway point for the size of differences (i.e., $|a_1b_1| - |a_2b_2| \geq \pm 0.10$). Table 6.7 contains the approximate power for each contrast method for percentile, BC, and BCa confidence intervals at a variety of sample sizes.

Table 6.7 Approximate power when true difference is moderate to large $|a_1 b_1| - |a_2 b_2| \geq \pm 0.10$

Note. All values are for indirect effects of opposing signs. BC = Bias-corrected. BCa = Bias-corrected and accelerated.

CI Method	Contrast Method	Sample Size				
		25	50	100	200	500
Percentile	1. $\lvert a_1 b_1\rvert - \lvert a_2 b_2\rvert$	9.6	31.4	62.2	85.1	95.6
	2. Conditional Contrast	10.3	32.3	62.7	85.2	95.6
	3. $a_1 b_1 + a_2 b_2$	24.3	54.3	79.6	92.6	97.8
	4. $\frac{a_1 b_1}{a_2 b_2}$	13.6	30.6	46.7	59.6	72.0
	5. $\frac{\lvert a_1 b_1\rvert}{\lvert a_2 b_2\rvert}$	9.6	31.4	62.2	85.1	95.6
	6. $\frac{\lvert a_1 b_1\rvert}{\lvert a_1 b_1\rvert + \lvert a_2 b_2\rvert}$	9.6	31.4	62.2	85.1	95.6
BC	1. $\lvert a_1 b_1\rvert - \lvert a_2 b_2\rvert$	24.4	47.5	70.4	86.7	95.7
	2. Conditional Contrast	-	-	-	-	-
	3. $a_1 b_1 + a_2 b_2$	30.1	58.3	80.9	92.8	97.8
	4. $\frac{a_1 b_1}{a_2 b_2}$	38.2	48.6	61.3	72.1	83.1
	5. $\frac{\lvert a_1 b_1\rvert}{\lvert a_2 b_2\rvert}$	35.5	48.1	67.6	85.5	95.5
	6. $\frac{\lvert a_1 b_1\rvert}{\lvert a_1 b_1\rvert + \lvert a_2 b_2\rvert}$	35.5	48.1	67.6	85.5	95.5
BCa	1. $\lvert a_1 b_1\rvert - \lvert a_2 b_2\rvert$	15.4	37.3	63.6	82.0	93.0
	2. Conditional Contrast	-	-	-	-	-
	3. $a_1 b_1 + a_2 b_2$	27.7	50.9	74.2	.89.3	96.3
	4. $\frac{a_1 b_1}{a_2 b_2}$	48.1	57.9	70.6	84.5	94.0
	5. $\frac{\lvert a_1 b_1\rvert}{\lvert a_2 b_2\rvert}$	37.2	51.7	63.0	79.9	92.4
	6. $\frac{\lvert a_1 b_1\rvert}{\lvert a_1 b_1\rvert + \lvert a_2 b_2\rvert}$	33.7	49.7	69.4	86.6	95.9

Although BC confidence intervals performed the best on average across sample sizes, percentile intervals were nearly identical in power starting at n = 100. This also is consistent with prior research, which suggests that as effect size increases, the difference in power between percentile CIs and BC/BCa CIs decreases (Fritz, Taylor, & MacKinnon, 2012; Hayes & Scharkow, 2013). The power across all conditions for small effect sizes was relatively low (no combination of sample size, CI type, or contrast method crossed the typical power threshold of .8—the highest power was .63). In comparison, all contrast methods besides the ratio exceeded .8 at a sample size of n = 200.

6.2.8 Graphical representations

It may be useful to visually demonstrate the coverage performance for each method. Figures 6.1, 6.2, and 6.3 contain—for percentile, BC, and BCa CIs respectively—panels of boxplots showing the correct coverage of the null difference for each method at a variety of sample sizes (for varying sizes of the indirect effects that are equal in value but opposite in sign).

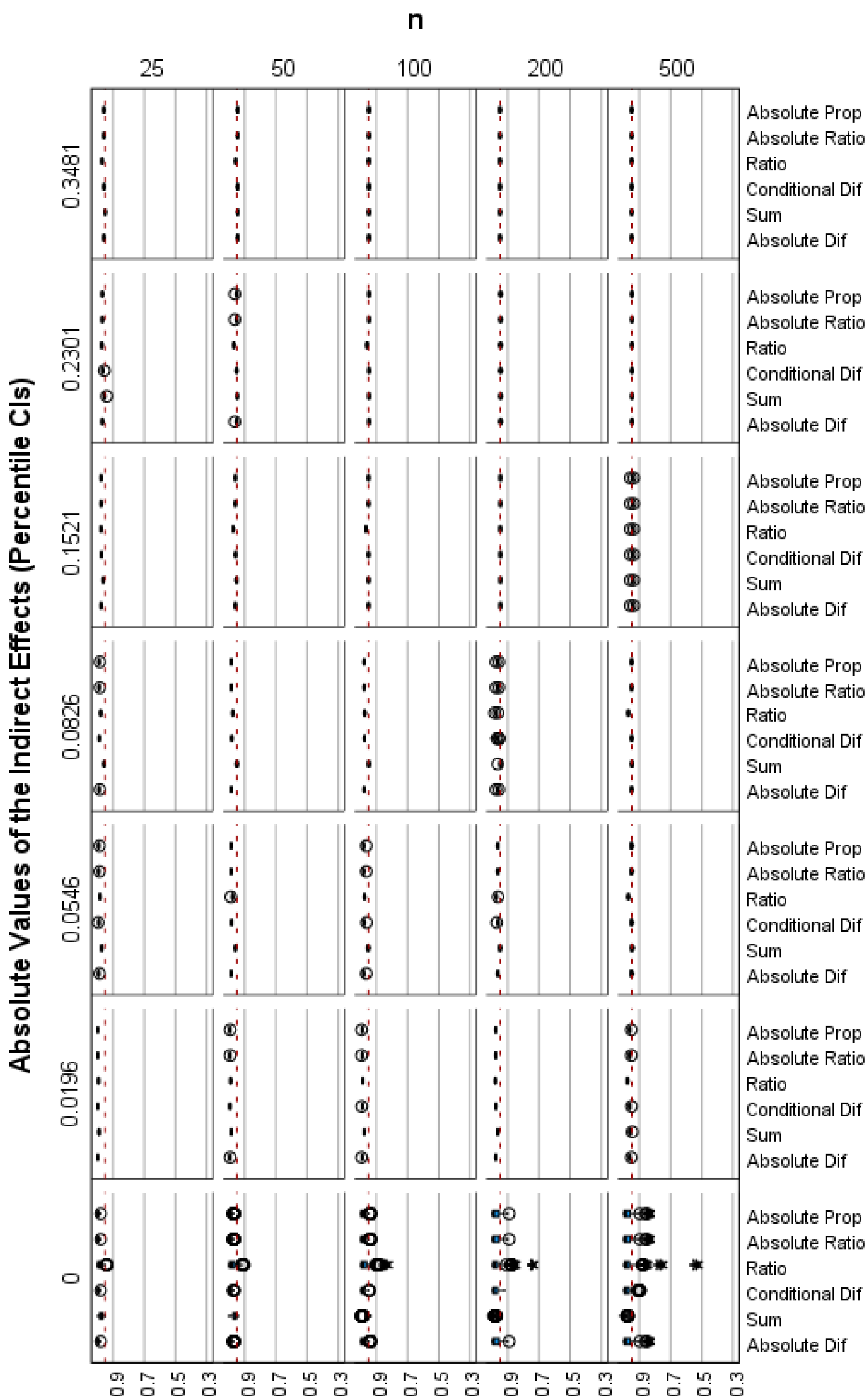

Figure 6.1 A boxplot grid for the coverage rates of the null difference for a variety of sample sizes and sizes of indirect effects. Coverage rates are for percentile bootstrap confidence intervals.

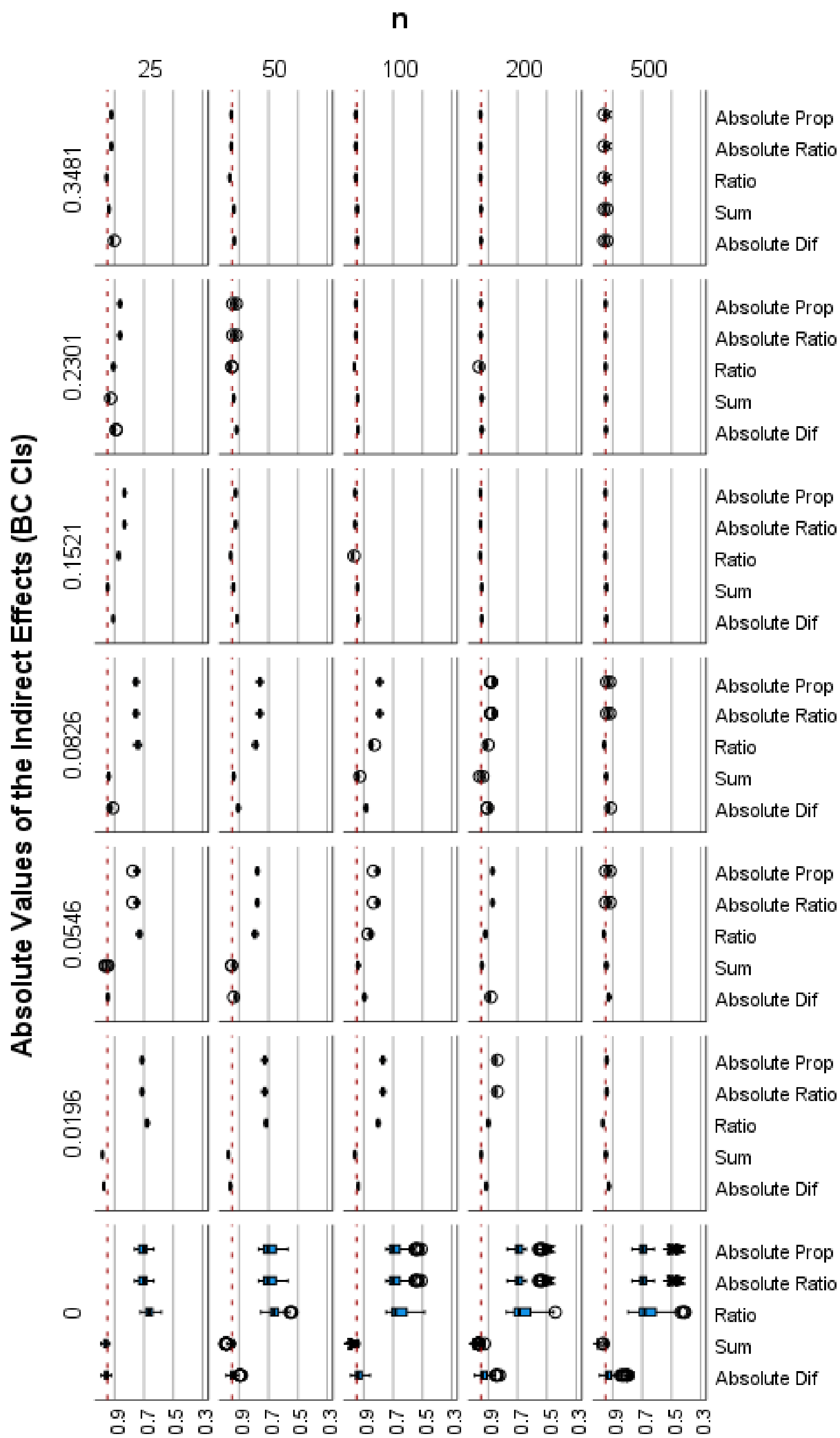

Figure 6.2 A boxplot grid for the coverage rates of the null difference for a variety of sample sizes and sizes of indirect effects. Coverage rates are for bias-corrected bootstrap confidence intervals.

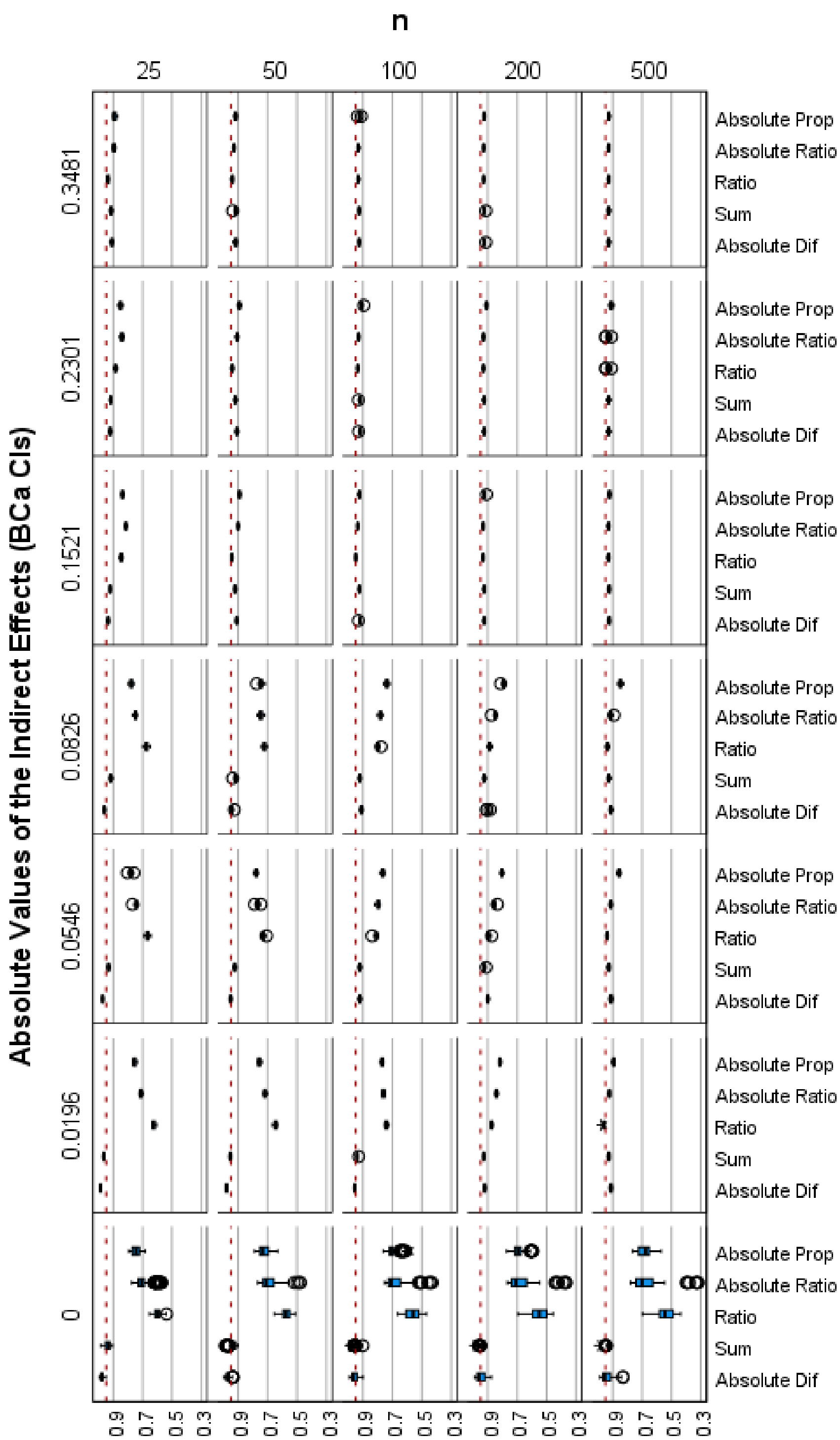

Figure 6.3 A boxplot grid for the coverage rates of the null difference for a variety of sample sizes and sizes of indirect effects. Coverage rates are for bias-corrected and accelerated bootstrap confidence intervals.

The figures support the information from Tables 6.1-6.5. Percentile CIs had less variability than BC and BCa CIs and greater coverage of the null difference, however as the size of the indirect effects and sample size increased, the differences between the CI types decreased. With large indirect effects and a sample size of 500, all contrast methods across all CI types had coverage rates centered around .95. The contrast methods involving division had far greater variability than the difference of absolute values and sum, especially with small indirect effects.

Figure 6.4 consists of boxplots that show the difference in power across percentile, BC, and BCa CIs for the contrast methods. As expected, percentile CIs have the greatest variability (with all methods having power ranging from 0 to 1), however the median power value for each contrast method is similar across the CI types (with exception to the ratio).

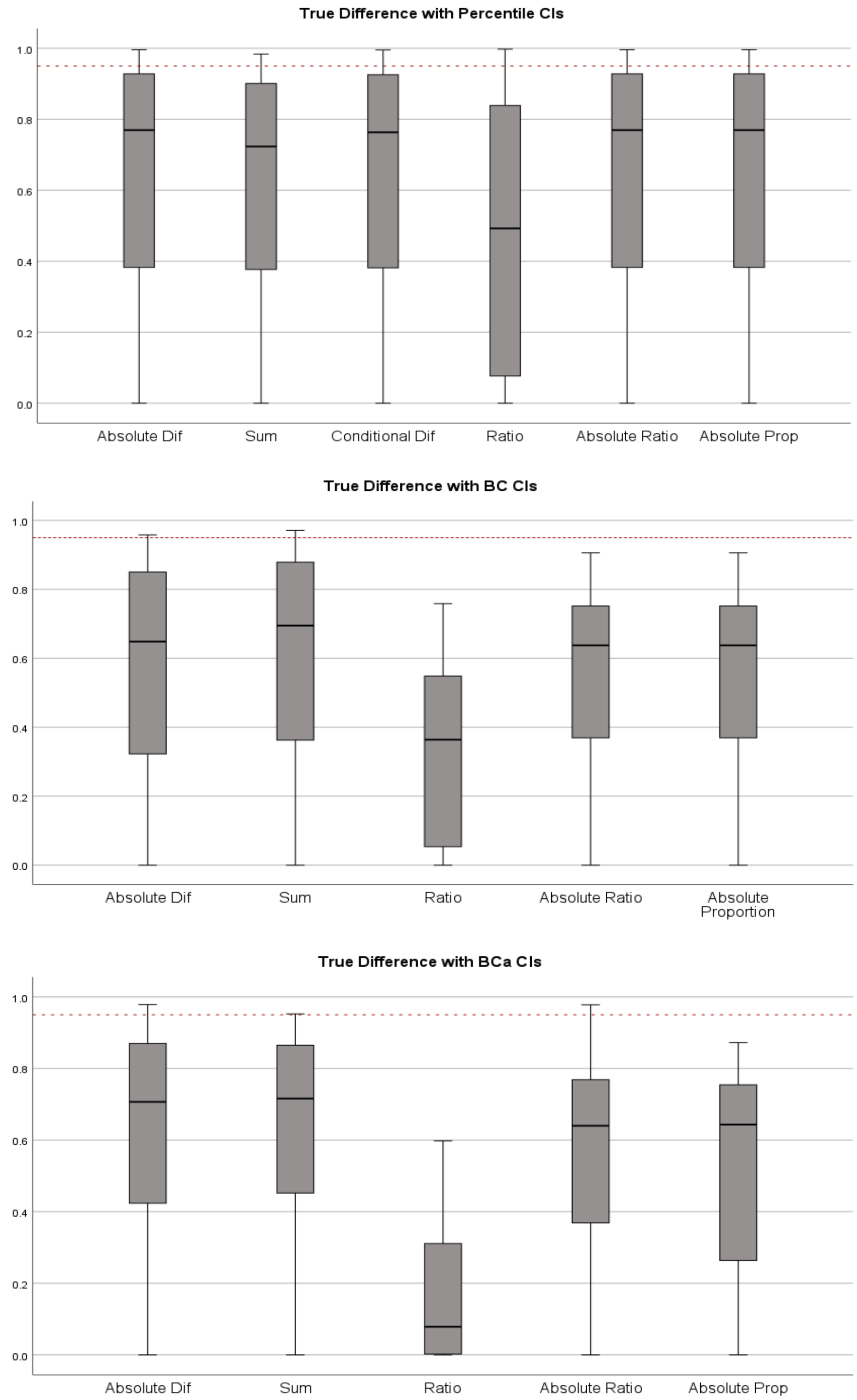

Figure 6.4 A boxplot grid for the coverage rates of the true difference (i.e., power) for various bootstrap confidence intervals. BC = Bias-corrected, BCa = Bias-corrected and accelerated.

Chapter 7: Discussion

Comparing indirect effects allows researchers to further expand their understanding of how causal processes operate. It is important to know that an effect is transmitted through multiple mechanisms, but it is also critical to know how these effects compare to one another in size and value. The information contrasts provide can not only save resources for and streamline the focus of behavioral rehabilitation programs, it can allow researchers to build substantive theory and compare the merits of multiple theories attempting to explain the same relationship. Researchers can answer more complex and nuanced questions by utilizing contrasts.

I started the thesis with an overview of simple and parallel mediation models. I then talked about the most common approach to comparing indirect effects and three different types of bootstrap confidence intervals that can be used to obtain inference for the difference between indirect effects. Identifying a problem of comparing indirect effects of opposing signs, I proposed several methods for comparing such effects. I then illustrated these concepts with a real-world example involving war veterans and PTSD. Finally, I conducted a simulation study to compare the coverage rates of these methods in a variety of settings based on effect and sample size.

7.1 Summary and Recommendations

Percentile CIs provided a good balance between controlling for Type I error and power. When looking at power at small effect sizes, percentile CIs did not perform as well as BC and BCa CIs. However, as sample and effect size increased, the differences in power between the types of confidence intervals decreased. When looking at the coverage of the true contrast value regardless of its size, percentile confidence intervals tended to perform better across sample sizes, effect sizes, and contrast method. Percentile bootstrap confidence intervals also had the lowest Type I error rate on average. The BC and BCa intervals had particularly poor coverage for contrast methods involving division (especially when one or both of the indirect effects were exactly or close to equal to zero).

Bigger sample sizes unsurprisingly resulted in better coverage of the true contrast value, smaller Type I error, and greater power. Although there is no formal recommendation for sample size as a result of this thesis, the bigger the sample size, the better. For example, it wasn't until n = 200 that most of the contrast methods started to approach 95% correct coverage of the null difference when the indirect effects were the same. The coverage rate of some of the contrast methods was still hovering around a conservative 97% with a sample size of 500.

Substantive researchers interested in comparing the strength of indirect effects can apply the results of this thesis to their research. The conditional contrast method did not perform noticeably better than the difference of absolute values or the raw difference condition. Since this practice is not theoretically sound (as you would have to look at the data to decide what test you were going to use), it is not recommended. Although the raw difference had slightly better coverage of the true difference when the indirect effects were of the same sign, the difference of absolute values could be applied as a general method of contrast (with percentile bootstrap CIs used for inference). Since the true size of the indirect effects or the difference(s) between them is unknown, it could be reasonable to use this combination of contrast method and type of confidence interval that performed generally well across sample and effect sizes. Alternatively, if the researchers had a sense of the direction of the indirect effects before the study began, they could decide *a priori* to compare the indirect effects using the raw difference and difference of absolute values/sum (though this assumes the researcher's hypothesis of the sign of the indirect effect is correct). Methods involving division have volatile coverage probabilities (as observed in the graphs and tables) and thus are not advised

7.2 Limitations and Future Directions

There are several ways future researchers can expand on this work. First, the simulation only added one source of error (the standard normal random error added to the *M*s and *Y* to reflect departures between the model estimates of *M* and *Y* and their actual values). The simulation design assumed no random measurement error. A more sophisticated simulation would consider both sources of error (since both exist in the real world). Additionally, the simulation only considered a parallel mediation model with two mediators. What of other types of data? For example, how would contrasts perform in dyadic data with varying degrees of correlation between the *X*s and *Y*s?

Since this thesis was a simulation study and there were no proofs or derivations, the results only generalize to the conditions detailed in the simulation. It could be that a different pattern of results emerges at a larger sample or effect size. Future research could also consider looking at other measures when considering the quality of a contrast method. It could be interesting to look at the bias of a contrast method at different effect/sample sizes. A measure of spread such as RMSE could be calculated along with bias.

7.3 Concluding Remarks

It is important to consider the mutual symbiotic relationship between substantive theory and methodological research. As theory becomes more complex, so must the statistical methods used to help answer such questions. However, the same is true for methodological research (Greenwald, 2012). As methods evolve and scholars make conducting certain analyses easier, the questions that substantive researchers ask can become more intricate and representative of the real world. It is the same with advancement of contrasts in mediation models.

This thesis is the first study of its kind exploring the performance of methods that can be used to compare indirect effects. Contrasts are not exceedingly common in the substantive literature, but studies like this help researchers become more aware that contrasts are possible, relatively simple to conduct, and backed by statistical research. The results of this research could be applied to software tools that aid researchers conducting mediation analysis in their own work (e.g., PROCESS, Hayes, 2018; MEDAYD, Coutts, Hayes, & Jiang, 2019). Whether assessing the efficacy of different mechanisms of behavioral change in a rehabilitation program, comparing substantive theories, or looking at the effects of war on military veterans, contrasts in mediation models enrich researchers’ understanding of the causal processes that underlie a relationship. Now, research exists exploring the best way to do that.

References

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality & Social Psychology*, *51*, 1173-1182.

Cerin, E., & MacKinnon, D. P. (2009). A commentary on current practice in mediating variable analyses in behavioural nutrition and physical activity. *Public health nutrition*, *12*(8), 1182-1188.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates Publishers.

Coutts, J. J., Hayes, A. F., & Jiang, T. (2019). Easy statistical mediation analysis with distinguishable dyadic data. *Journal of Communication*, *69*(6), 612-649.

Ding, Z., Ng, F., & Li, J. (2014). A parallel multiple mediator model of knowledge sharing in architectural design project teams. *International Journal of Project Management*, *32*(1), 54-65.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American statistical Association*, *82*(397), 171-185.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman and Hall.

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological science*, *18*(3), 233-239.

Fritz, M. S., Taylor, A. B., & MacKinnon, D. P. (2012). Explanation of two anomalous results in statistical mediation analysis. *Multivariate Behavioral Research*, *47*(1), 61-87. DOI: 10.1080/00273171.2012.640596

Gelman, A., & Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *The American Statistician*, *60*(4), 328-331.

Greenwald, A. G. (2012). There is nothing so theoretical as a good method. *Perspectives on Psychological Science*, *7*, 99-108.Hart, P. S. (2011). One or many? The influence of episodic and thematic climate change frames on policy preferences and individual change behavior. *Science Communication, 33*, 28-51.

Grund, A., & Fries, S. (2014). Study and leisure interference as mediators between students' self-control capacities and their domain-specific functioning and general well-being. *Learning and Instruction*, *31*, 23-32.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd Ed.). New York: The Guilford Press.

Hayes, A. F., & Cai, L. (2007). Using heteroscedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. Behavior Research Methods, 39, 709–722. doi:10.3758/BF03192961

Hayes, A. F., Montoya, A. K., & Rockwood, N. J. (2017). The analysis of mechanisms and their contingencies: PROCESS versus structural equation modeling. *Australasian Marketing Journal*, *25*, 76-81.

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter?. *Psychological science*, *24*(10), 1918-1927.

Ho, S. S., Peh, X., & Soh, V. W. (2013). The cognitive mediation model: Factors influencing public knowledge of the H1N1 pandemic and intention to take precautionary behaviors. *Journal of Health Communication*, *18*(7), 773-794.

Hoffman, L. H., & Young, D. G. (2011). Satire, punch lines, and the nightly news: Untangling media effects on political participation. *Communication Research Reports*, *28*(2), 159-168.

Jin, M. H., McDonald, B., & Park, J. (2018). Does public service motivation matter in public higher education? Testing the theories of person–organization fit and organizational commitment through a serial multiple mediation model. *The American Review of Public Administration*, *48*(1), 82-97.

Lapointe, É., Morin, A. J., Courcy, F., Boilard, A., & Payette, D. (2012). Workplace affective commitment, emotional labor and burnout: A multiple mediator model. *International Journal of Business and Management*, *7*(1), 3.

Levant, R. F., & Wimer, D. J. (2014). The Relationship Between Conformity to Masculine Norms and Men's Health Behaviors: Testing a Multiple Mediator Model. *International Journal of Men's Health*, *13*(1).

Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity-consistent standard errors in the linear regression model. American Statistician, 54, 217–224.

MacKinnon, D. P. (2000). Contrasts in multiple mediator models. In J. Rose, L. Chassin, C. C. Presson, & S. J. Sherman (Eds), *Multivariate applications in substantive use research: New methods for new questions* (pp. 141-160). Mahwah, NJ: Lawrence Erlbaum Associates.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention science*, *1*(4), 173-181.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological methods*, *7*(1), 83.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate behavioral research*, *39*(1), 99-128.

Manne, S. L., Kashy, D. A., Virtue, S., Criswell, K. R., Kissane, D. W., Ozga, M., ... & Rodriguez, L. (2018). Acceptance, social support, benefit-finding, and depression in women with gynecological cancer. *Quality of Life Research*, *27*(11), 2991-3002.

Nitzl, C., Roldan, J. L., & Cepeda, G. (2016). Mediation analysis in partial least squares path modeling. *Industrial management & data systems*.

O'Rourke, H. P., & MacKinnon, D. P. (2018). Reasons for testing mediation in the absence of an intervention effect: A research imperative in prevention and intervention research. *Journal of studies on alcohol and drugs*, *79*(2), 171-181.

Pitts, B. L., Safer, M. A., Castro-Chapman, P. L., & Russell, D. W. (2018). Retrospective appraisals of threat and benefit mediate the effects of combat experiences on mental health outcomes in army medics. *Military Behavioral Health*, *6*(3), 226-233.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40,* 879-891

Rijnhart, J. J., Twisk, J. W., Chinapaw, M. J., de Boer, M. R., & Heymans, M. W. (2017). Comparison of methods for the analysis of relatively simple mediation models. *Contemporary clinical trials communications*, *7*, 130-135.

Romero-Moreno, R., Losada, A., Márquez-González, M., & Mausbach, B. T. (2016). Stressors and anxiety in dementia caregiving: multiple mediation analysis of rumination, experiential avoidance, and leisure. *International psychogeriatrics*, *28*(11), 1835-1844.

Rudy, B. M., Davis, T. E., & Matthews, R. A. (2012). The relationship among self-efficacy, negative self-referent cognitions, and social anxiety in children: A multiple mediator model. *Behavior Therapy*, *43*(3), 619-628.

Sanchez, M., Haynes, A., Parada, J. C., & Demir, M. (2018). Friendship maintenance mediates the relationship between compassion for others and happiness. *Current Psychology*, 1-12.

Schönfeld, P., Brailovskaia, J., Bieda, A., Zhang, X. C., & Margraf, J. (2016). The effects of daily stress on positive and negative mental health: Mediation through self-efficacy. *International Journal of Clinical and Health Psychology*, *16*(1), 1-10.

Schotanus-Dijkstra, M., Pieterse, M. E., Drossaert, C. H., Walburg, J. A., & Bohlmeijer, E. T. (2019). Possible mechanisms in a multicomponent email guided positive psychology intervention to improve mental well-being, anxiety and depression: A multiple mediation model. *The Journal of Positive Psychology*, *14*(2), 141-155.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychological methods*, *7*(4), 422.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equations models. In S. Leinhart (Ed.), *Sociological methodology* (pp. 290-312). San Francisco: Jossey-Bass.

Stone, C. A., & Sobel, M. E. (1990). The robustness of estimates of total indirect effects in covariance structure models estimated by maximum. *Psychometrika*, *55*(2), 337-352.

Teng, F., Gao, W., Huang, X., & Poon, K. T. (2019). Body surveillance predicts men's and women's perceived loneliness: A serial mediation model. *Sex Roles*, *81*(1-2), 97-108.

Ter Hoeven, C. L., van Zoonen, W., & Fonner, K. L. (2016). The practical paradox of technology: The influence of communication technology use on employee burnout and engagement. *Communication monographs*, *83*(2), 239-263.

Van der Heijden, B., Mahoney, C.B., & Xu, Y. (2019). Impact of job demands and resources on Nurses' burnout and occupational turnover intention towards an age-moderated mediation model for the Nursing profession. *International journal of environmental research and public health*, *16*(11), 2011.

VanderWeele, T., & Vansteelandt, S. (2014). Mediation analysis with multiple mediators. *Epidemiologic methods*, *2*(1), 95-115.

Wieder, B., & Ossimitz, M. L. (2015). The impact of Business Intelligence on the quality of decision making–a mediation model. *Procedia Computer Science*, *64*, 1163-1171.

Yap, S. T., & Baharudin, R. (2016). The relationship between adolescents' perceived parental involvement, self-efficacy beliefs, and subjective well-being: A multiple mediator model. *Social Indicators Research*, *126*(1), 257-278.

Yıldız, M. A. (2016). Multiple Mediation of Emotion Regulation Strategies in the Relationship between Loneliness and Positivity in Adolescents. *Education & Science*, *41*(186).