

Evidence-based recommendations of reporting results from mediation analysis:

A focus on ease of interpretation and maximum accuracy

Yoosun Kim

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

June 2016

© Yoosun Kim, 2016

Abstract

Theoretical models of mediation are common in psychological research, but there is much variability in how results of mediation analyses are reported which could result in interpretational errors, misconceptions, and differences in reader perception and experience. The goal of this study is to develop evidence-based recommendations for reporting results of mediation to reduce objective interpretational errors and maximize readers' subjective experience. These recommendations would be based on results from an experiment examining the effect of the four different forms of result reporting (text only, text and table, text and path diagram, and text, table, and path diagram) on interpretational errors and readers' experience, where reader experience is composed of four constructs: perceived time, ease of understanding, satisfaction in understanding reported results, and confidence in understanding reported results. Results show that including a path diagram may benefit reducing comprehensive interpretational error and increasing positive reader experiences.

Keywords: mediation, reporting, table, graphs, interpretation, satisfaction.

TABLE OF CONTENTS

| | |
|---|-----------|
| Abstract | ii |
| Table of Contents | iii |
| List of Tables | iv |
| List of Figures | v |
| List of Equations | vi |
| Chapter One: Introduction | 1 |
| 1.1. Purpose of Present Study | 3 |
| 1.2. Simple Mediation..... | 4 |
| 1.3. Present Study | 8 |
| 1.4. Hypotheses..... | 11 |
| Chapter Two: Method | 18 |
| 2.1. Participants | 18 |
| 2.2. Measures | 18 |
| 2.3. Design..... | 24 |
| 2.4. Procedure | 26 |
| Chapter Three: Results | 28 |
| 3.1. Preliminary Analysis..... | 28 |
| 3.2. Primary Analysis | 33 |
| Chapter Four: Discussion | 50 |
| 4.1. Reading for Basic Information | 51 |
| 4.2. Reading for Detailed Information | 52 |
| 4.3. Recommendations..... | 55 |
| 4.4. Limitations..... | 56 |
| 4.5. Future Directions | 57 |
| References | 59 |
| Appendices..... | 64 |
| Appendix A: Interpretational Error Scale..... | 64 |
| Appendix B: Attitude Toward Mediation Scale | 70 |
| Appendix C: Demographics | 76 |
| Appendix D: Experimental Stimuli | 78 |
| Appendix E: Informed Consents..... | 85 |
| Appendix F: Debriefing | 88 |

LIST OF TABLES

| | |
|---|----|
| Table 1: Characteristics of Participants | 19 |
| Table 2: Correlation Between Interpretational Error and Reader Experience | 30 |
| Table 3: Predicted Means in Reading for Basic Information | 35 |
| Table 4: Predicted Means in Reading for Detailed Information | 41 |
| Table 5: Frequency of the Preference in Presentation Format | 48 |

LIST OF FIGURES

| | |
|---|----|
| Figure 1: Path Diagram for Simple Mediation | 1 |
| Figure 2: Procedure of Experiment..... | 26 |

LIST OF EQUATIONS

| | |
|--|----|
| Equation 1: Linear Regression for Path c | 4 |
| Equation 2: Linear Regression for Path a | 4 |
| Equation 3: Multiple Linear Regression for Path b and c' | 4 |
| Equation 4: Sobel Test z statistic | 6 |
| Equation 5: Sobel Test Standard Error | 6 |
| Equation 6: ANCOVA Model | 34 |

Chapter One: Introduction

The simple mediation model is highly popular in psychological research as it is often used to investigate underlying causal mechanisms. Suppose that the change in an independent variable (X) influences the change in a dependent variable (Y) as shown in path c of the path diagram in Panel A of Figure 1.

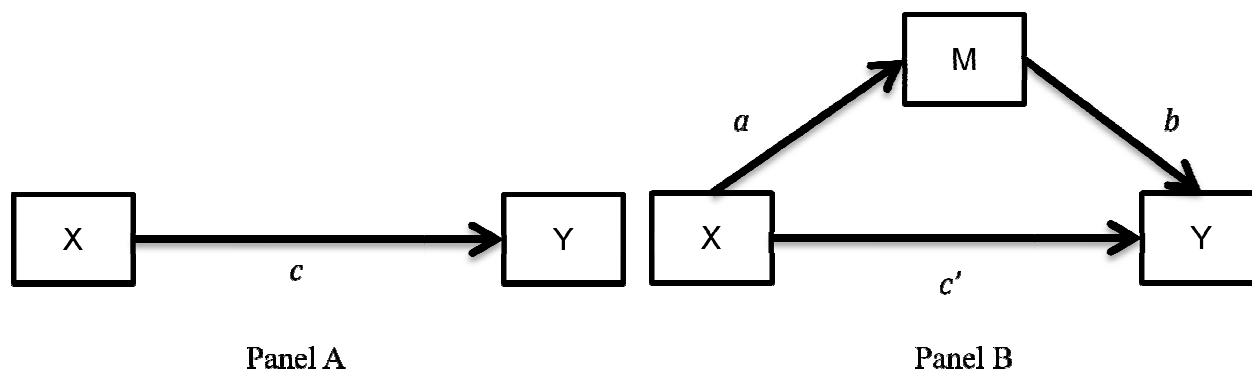


Figure 1. Path diagram for simple mediation

The effect of X on Y is said to occur through M , and is quantified by either the product of path a and b , ab , or the difference between paths c and c' , $c - c'$ (Baron & Kenny, 1986). As the mediation model was designed to identify a mediator as an underlying mechanism of the effect of X on Y , many researchers make use of simple mediation in their research. The mediation model is broadly applied in various areas of psychology, such as social, cognitive, developmental, organizational, clinical, and prevention research (MacKinnon, Fairchild, & Fritz, 2007).

At least 14 approaches for testing simple mediation have been proposed (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002), and currently popular approaches are Baron and Kenny's (1986) causal steps, the Sobel (1982, 1986) test, and resampling approaches such as the bootstrap confidence intervals (Bollen & Stine, 1990). The large number of approaches to test simple mediation adds to the variability in what and how results from simple mediation are

reported (Gelfand, Mensinger, & Tenhave, 2009). What are most commonly reported in results are either standardized or unstandardized regression coefficients or slopes. Recommendations have been made for reporting both standardized and unstandardized regression coefficients, accompanied by t test statistic values or standard errors for all predictors in applications of multiple linear regression (Kashy, Donnellan, Ackerman, & Russell, 2009). Yet, these recommendations have not been widely implemented in practice for reporting results using simple mediation. For instance, Zhang, Jemmott, and Jemmott (2015) reported unstandardized coefficients and their p -values, Stoddard, Varela, and Zimmerman (2015) reported unstandardized coefficients without their p -values, and Piscotty, Kalisch, and Gracey-Thomas (2015) reported all unstandardized coefficients and a selective subset of test statistics and p -values for some coefficients.

Beyond what different types of statistical results are reported, based on different approaches to test mediation (e.g., causal steps, Sobel test, and resampling approaches), how the results are presented for simple mediation also varies across publications. The three most commonly used forms of presenting results are text only, a table of regression coefficients, and a path diagram (e.g., Panel B in Figure 1 on page 1). For instance, Kamphoff, Gill, and Huddleston (2005) presented their results on mediation using only text. El-Sheikh and Elmore-Staton (2004) presented results from similar analyses using a table of coefficients, and Garcia, Kerekes, and Archer (2012) made use of a path diagram with standardized regression coefficient estimates. Other authors, such as Gelsema, van der Doef, Maes, Akerboom, and Verhoeven, (2005), used a combination of both a table and a path diagram to present their findings. Currently, there are no discussions about how best to present simple mediation results (cf. Kashy et al., 2009).

Such variability in reporting results of the simple mediation model potentially results in some undesirable outcomes in practice. Specifically, high variability in result reporting can potentially increase interpretational errors associated with the simple mediation model. How results are presented is also likely to influence readers' experience in understanding the study's conclusions. In related research using undergraduate students, Agus, Peró-Cebollero, Penna, and Guàrdia-Olmos (2015) showed how different formats used to present general statistical problems affected interpretational errors and undergraduate students' confidence in making correct answers. In Agus et al.'s (2015) study that examined many different outcomes, significant differences in the number of correct responses between two different presentation formats (i.e., format using numbers and format using graphs) were found in participants' understanding of association between variables. Thus, recommendations for how best to report results can potentially decrease interpretational errors and reduce variability in readers' experience.

1.1. Purpose of Present Study

The present study aims to develop evidence-based recommendations for reporting results of simple mediation so as to reduce the existing variability in reported results. It is expected that reduced variability in the way results are reported would enhance reader experience and reduce potential interpretational errors. An empirical study is designed to examine the effects of different formats of reporting results in simple mediation on two sets of outcomes. The first set of outcomes is about number of interpretational errors made by readers. Interpretational accuracy is important because the main goal of research reporting is to accurately communicate findings to the readers. In addition to interpretational errors, the second set of outcomes includes four domains of readers' experience. These domains are: (i) perceived time; (ii) ease of understanding; (iii) satisfaction in understanding; and (iv) confidence in understanding the reported results.

Before reviewing the outcomes of interpretational errors and readers' experience in detail, the concepts behind simple mediation are reviewed first.

1.2. Simple Mediation

The three multiple linear regression (MLR) equations which are typically used by the various approaches for testing simple mediation in practice are presented below:

$$Y = i_1 + cX + e_1 \quad (1)$$

$$M = i_2 + aX + e_2 \quad (2)$$

$$Y = i_3 + c'X + bM + e_3, \quad (3)$$

where i_1 , i_2 , and i_3 are intercepts, e_1 , e_2 , and e_3 are error terms, and X , M , and Y are the independent variable, the mediator, and the dependent variable, respectively. These three regression equations map onto the path diagrams in Figure 1 on page 1. The total or marginal effect of X on Y is the coefficient or slope c (Panel A, Figure 1). The slope a quantifies the effect of X on M , the slope b quantifies the effect of M on Y , controlling for the direct effect of X on Y , and finally the slope c' represents the direct or conditional effect of X on Y (Panel B, Figure 1).

Each of the estimated coefficients from these MLR models does not directly estimate the mediated or indirect effect, which is carried by the product ab or the difference $c - c'$ (Baron & Kenny, 1986). Additional statistical procedures that make use of the estimated coefficients (\hat{a} , \hat{b} , \hat{c} , and \hat{c}') are typically conducted to evaluate the statistical significance of the mediated effect.

In order to control the variability in what is reported by presenting results of the mediation analysis, this study presents three approaches that are widely used in practice: the causal steps approach (Baron & Kenny, 1986), the Sobel (1982, 1986) test, and the percentile bootstrap confidence interval (Bollen & Stine, 1990; Shrout & Bolger, 2002).

1.2.1. Causal steps

The causal steps method assesses significance of various regression coefficients in the following steps. Step one, the total or marginal effect of X on Y (path c of Panel A in Figure 1 on page 1), is required to be significant, which indicates an effect of X on Y to be potentially accounted for by a mediator (M). Step two, the inclusion of M into the regression of Y on X should reduce the initial significant effect of X on Y (path c of Panel A in Figure 1) to nonsignificance in path c' (Panel B in Figure 1). This second step implies that M potentially explains the effect of X on Y . Step three, the effect of X on M (path a of Panel B in Figure 1) and the effect of M on Y (path b of Panel B in Figure 1) should be significant, situating M between X and Y as a mediator. All three conditions need to be met before mediation can be concluded.

Simulation studies have shown that the causal steps approach testing for a mediated effect has the lowest power compared to the Sobel test and other resampling methods (MacKinnon et al., 2002). Rucker, Preacher, Tormala, and Petty (2011) also discussed the low power in causal steps approach. Additionally, the indirect or mediated effect, quantified by ab or $c - c'$, is not directly evaluated under the causal steps approach. As a supplement to the causal steps approach, Baron and Kenny (1986) recommended use of the Sobel (1982, 1986) test.

1.2.2. Sobel test

The Sobel test assesses significance of the indirect effect, where a z -test statistic is derived for testing the null hypothesis of the indirect effect, $H_0: ab = 0$. This null hypothesis states that the indirect effect of ab in the population is zero, and rejection of this null hypothesis indicates significance of the mediated effect. The z -test statistic is defined as

$$z = \frac{\hat{a}\hat{b}}{SE_{\hat{a}\hat{b}}}, \quad (4)$$

where $\hat{a}\hat{b}$ is the estimate of the indirect effect, and $SE_{\hat{a}\hat{b}}$ is the standard error of the indirect effect estimate. The coefficients \hat{a} and \hat{b} are unstandardized estimates as defined in Equation (2) and (3) on page 4. In particular, the standard error is defined as

$$SE_{\hat{a}\hat{b}} = \sqrt{\hat{a}^2 s_b^2 + \hat{b}^2 s_a^2}, \quad (5)$$

where s_a^2 and s_b^2 are the estimated variances of the coefficients \hat{a} and \hat{b} , respectively. The standard error of the indirect effect estimate, $SE_{\hat{a}\hat{b}}$, is derived with the multivariate delta method, which is a linear approximation of the nonlinear function associated with the nonlinear transformation of ab . Note that one of the key assumptions underlying the performance of the Sobel test is normality of the distribution of $\hat{a}\hat{b}$.

The Sobel test is computationally simple and easy to implement, and provides a significant test of the indirect effect unlike the causal steps method. However, the sampling distribution of the indirect effect $\hat{a}\hat{b}$ is known to be nonnormal with positive skew and high kurtosis (MacKinnon, Lockwood, & Hoffman, 1998). As the Sobel test is based on the assumption of normality, results from this test tend to be erroneous in most applications of simple mediation in psychology, where sample sizes are not large enough for the central limit theorem to ensure that the distribution of $\hat{a}\hat{b}$ is normal. Additionally, simulation studies show that the Sobel test has low power rates and low Type I error rates in normally distributed data (MacKinnon et al., 2002). The Sobel test is also shown to have low Type I error rates for nonnormal data with no missing data (Biesanz, Falk, & Savalei, 2010). In this case, the Sobel test exhibited low Type I error rates which were so low that they affected the power to detect an

indirect effect. Despite these shortcomings, the Sobel test remains one of the most popular methods for testing mediation.

1.2.3. Bootstrap confidence interval

An alternative to the Sobel test for testing the significance of the mediated effect ab is to construct a bootstrap confidence interval (CI) for ab , which can also be used to test the null hypothesis of the indirect effect, $H_0: ab = 0$. The advantage of the bootstrap method over the Sobel test is that the bootstrap does not require distributional assumptions. The percentile bootstrap CI for ab (Bollen & Stine, 1990; Shrout & Bolger, 2002) is constructed by empirically simulating the sampling distribution of $\hat{a}\hat{b}$ in the following steps. First, B bootstrap replicates of size N are drawn from the sample data with replacement, where B is extremely large; typically $B=1,000$ or $B=5,000$ (Shrout & Bolger, 2002). Next, the models in Equations (2) and (3) on page 4 are fit to the B bootstrap replicates to obtain B sets of estimates. The B estimates of \hat{a} and \hat{b} are then multiplied and used to empirically construct the sampling distribution of $\hat{a}\hat{b}$. It is from this bootstrapped sampling distribution that a CI is constructed. For a 95% percentile bootstrap CI, the 2.5% and 97.5% quantiles of the empirical sampling distribution for $\hat{a}\hat{b}$ are taken to be the lower and upper bounds of the CI respectively. The percentile bootstrap CI is recommended for small to moderate sample sizes by various methodologists because of its higher power and higher Type I error rates compared to the causal steps and the Sobel test (Bollen & Stine, 1990; Shrout & Bolger, 2002; Preacher & Hayes, 2004; Biesanz et al., 2010).

In summary, the three most popular approaches to test simple mediation were reviewed. The causal steps and Sobel test are historically popular, and are continuously used in practice despite their drawbacks of low power and low Type I error rates (Rucker et al., 2011; Biesanz et

al., 2010). The causal steps method has the lowest power, and the Sobel test is associated with improper Type I error for sample sizes often used in empirical studies in Psychology (MacKinnon et al., 2002). In contrast, simulation studies show that the percentile bootstrap CI overcomes the shortcomings of the causal steps and the Sobel test because it balances power and Type I error rates under different conditions in terms of normally distributed data (MacKinnon et al., 2002; Biesanz et al., 2010).

1.3. Present Study

This study utilizes these three familiar approaches to test simple mediation in reported results. The study presents regression coefficients of the MLR models in Equations (1) to (3) on page 4, t test statistics, p -values, and R^2 . In addition, the Sobel z test statistic and the percentile bootstrap CI for ab are presented to study participants.

In an ongoing review of publications on empirical studies, which make use of mediation analysis, the popular formats of presenting results in simple mediation tend to be (i) a table with regression coefficients; (ii) a path diagram (i.e., Panel B of Figure 1 on page 1); (iii) both a table and a path diagram; or (iv) no specific form other than presenting the results with text. Reflecting current practice, four different combinations of formats are used in this study. These formats are: (i) text only; (ii) text and table of coefficients; (iii) text and path diagram; and (iv) text, table of coefficients, and path diagram. The main aim of this study is to examine whether these different forms of presenting simple mediation results possibly affect interpretational errors and readers' experiences in understanding the reported results.

1.3.1. Interpretational errors

Understanding results from a simple mediation analysis requires mastery in two domains of knowledge: (i) the concept of simple mediation; and (ii) statistical methods testing simple mediation. The concept of simple mediation is a formulation of a causal mechanism where a change of an independent variable (X) results in a change of a dependent variable (Y) through a mediator (M). Based on this concept of simple mediation, researchers then apply a statistical method to test this concept. In order to use a statistical method to test mediation, researchers have to first quantify the relationships among X , M , and Y with regression coefficients from the multiple linear regression (MLR) in Equations 1, 2, and 3 on page 4. The quantified effects from MLR are the effect of X on Y (path c of Panel A in Figure 1 on page 1), the effect of X on M (path a of Panel B in Figure 1), the effect of M on Y controlling the effect of X on Y (path b of Panel B in Figure 1), and the effect of X on Y controlling the effect of M on Y (path c' of Panel B in Figure 1). These estimates are obtained from the MLR models discussed in the previous section in Simple Mediation (refer to Equations 1, 2, and 3). Thus, researchers need to be familiar with MLR to specify the relationships in their simple mediation model.

After obtaining the estimates of the relationships among X , M , and Y in these regression equations, researchers often test whether the mediation is significant by using approaches such as the causal steps method, the Sobel test, and the percentile bootstrap confidence interval about the path a and path b . In addition to familiarity with MLR, researchers need to have some knowledge of the different statistical methods which are developed to test significance of simple mediation. In particular, the indirect effect is estimated from the estimates of path a and path b or path c and c' such that the significance of either ab or $c - c'$ is tested. Many requirements and complexities behind examining mediation potentially allow for interpretational errors. These

interpretational errors are a measure of participants' objective performance in terms of understanding results that are presented from simple mediation analyses. Additionally, these interpretational errors are also potentially influenced by how the results are presented.

1.3.2. Reader experience

Readers' experience is defined as the subjective appraisal of the reported results in simple mediation. In a study by Agus et al. (2015) of undergraduate students, confidence in the correctness of participants' answers was measured. Any effect of different formats of presenting statistical questions (e.g., text and numbers versus graphs) on the measured confidence in the correctness of participants' answers was also examined. The participants in Agus et al. (2015) gave a higher number of correct responses in the presentation formats using numbers and text (N format) compared to the presentation format using graphs (G format). Agus et al. (2015) also measured how much time participants spent completing the questionnaires. The participants in the N format spent less time ($M = 22:31$ minutes) solving the statistical questions compared to the participants in the G format ($M = 23:05$ minutes).

In the present study, after participants complete items evaluating their understanding of the reported results, participants' confidence and perceived time taken, as well as their satisfaction and ease in understanding of the reported results are measured.

1.3.3. Familiarity with mediation.

Participants' familiarity with mediation is a potential confounder in examining the effect of different presentation formats on interpretational errors and readers' experience. Thus, readers' familiarity with mediation is also measured in addition to interpretational errors and readers' experience.

1.4. Hypotheses

Recall that four different combinations of formats are used in this study: (i) text only; (ii) text and table of coefficients; (iii) text and path diagram; and (iv) text, table of coefficients, and path diagram.

After controlling for familiarity with mediation, it is expected that the different formats of reporting simple mediation results have an effect on participants' objective performance in terms of interpretational errors as well as on participants' subjective appraisal of the experience. As part of interpretational errors, this study examines two phases of understanding the reported results. The first phase is reading for basic information of simple mediation, reflecting the context where researchers briefly glean research papers for general results. The second phase is more in depth and detailed information focusing on the proper interpretation of the numerical values reported.

1.4.1. Reading for basic information

1.4.1.1. *Interpretational errors*

Under the phase of reading for basic information, participants' cursory understanding of simple mediation is measured. Cursory understanding is defined as having a conceptual understanding of simple mediation, which is well represented in the path diagram in Figure 1 on page 1. The path diagram uses a more realistic conceptual representation of what is tested compared to the abstract numbers in the text or table to describe the relationship between the independent variable (X) and the dependent variable (Y) through the mediator (M). Thus, it is expected that the conditions with a path diagram (i.e., text and path diagram, and text, table of

coefficients, and path diagram) will have lower cursory interpretational errors compared to the conditions with no path diagram (i.e., text only, and text and table of coefficients).

1.4.1.2. Reader experience

Perceived time and ease. In Agus et al. (2015), the actual time spent by participants in the N format was shorter compared to in the G format. In contrast, the present study measures readers' perceived time spent after answering the items measuring a cursory understanding of the results in the Interpretational Error scale because perceived time of evaluation is related to the familiarity of objects presented (Tormala, Clarkson, & Henderson, 2011). The path diagram has a more realistic conceptual representation of mediation, which is more familiar to participants, compared to the abstract numbers in the text and the table of coefficients. Thus, it is expected that the conditions with a path diagram (i.e., text and path diagram, and text, table of coefficients, and path diagram) will have shorter perceived time compared to the conditions with no path diagram (i.e., text, and text and table of coefficients).

The realistic conceptual representation of the path diagram may also increase readers' ease in having a cursory understanding of the reported results compared to only text or abstract numbers. Thus, it is expected that the conditions with a path diagram (i.e., text and path diagram, and text, table of coefficients, and path diagram) will have higher ratings of ease compared to the conditions with no path diagram (i.e., text, and text and table of coefficients).

Satisfaction and Confidence. A study by Tasir, Abd Halim, and Harun (2012) showed that competency in information and communication technologies (ICT) is positively related to the confidence in using ICT as well as to the satisfaction on ICT training programmes in Malaysian teachers. Additionally, Agus et al. (2015) showed that participants who were

undergraduate students had more correct answers and were more confident in their answers being correct in the presentation format of tables with numbers (N format). According to Agus et al. (2015), when participants had more correct answers, they became more confident in their answers being correct. However, there are important differences between the present study and the study by Agus et al. (2015). First, the population of interest in Agus et al. (2015) is undergraduate students. In contrast, the population of interest in this study is current and potential researchers (i.e., graduate students, postdoctoral fellows, and faculty members). Second, Agus et al. (2015) examined undergraduate students' general statistical understanding, whereas, the present study examined specific concepts in simple mediation. Finally, Agus et al. (2015) used general tables with numbers and graphs while the present study uses specific presentation formats that are often employed to present results of simple mediation analyses (e.g., a table of coefficients and a path diagram). Thus, results in this study are expected to be different from the results of the study by Agus et al. (2015).

The present study hypothesizes that conditions with a path diagram will have lower cursory interpretational errors compared to the conditions with no path diagram due to a more realistic conceptual representation of mediation in a path diagram. In this vein, it is expected that the conditions with a path diagram (i.e., text and path diagram, and text, table of coefficients, and path diagram) will have higher satisfaction and higher confidence compared to the conditions with no diagram (i.e., text, and text and table of coefficients).

1.4.2. Reading for detailed information

1.4.2.1. *Interpretational errors*

In the phase where a comprehensive understanding of results is examined, participants are asked to interpret numerical values in the context of research presented in the stimuli.

Additionally, participants are also asked to identify different methods used to compute and test significance of the indirect effect presented in the experimental stimuli.

Recall that four different combinations of formats are used in this study: (i) text only; (ii) text and table of coefficients; (iii) text and path diagram; and (iv) text, table of coefficients, and path diagram. In the text and table of coefficients condition, estimates are summarized in a more ordered way compared to the text only condition. Additionally, the table of coefficients provides essential information without any unnecessary distracting features because the table only uses numbers (Gaissmaier, Wegwarth, Skopec, Muller, & Broschinski, 2012). Thus, it is expected that the text and table of coefficients condition will have less comprehensive interpretational error in understanding reported results compared to the text only condition.

Abstract numbers in a table of coefficients, however, are limited in presenting conceptual ideas about mediation unlike the visual presentation of regression models via path diagrams (e.g., Figure 1 on page 1). The strength of the graphical representation in the path diagram is providing additional information on top of the numerical information presented in the reported estimates. The path diagram itself describes relationships among X , Y , and M in a concrete conceptual map rather than abstract numbers (cf., Tukey, 1990). In an empirical study by Cosmides and Tooby (1996), graphical representation was found to enhance solving statistical problems. Thus, it is expected that the formats using a combination of numbers and text (i.e., text only, and text and

table of coefficients) will have a higher number of comprehensive interpretational errors compared to the format using graphs (i.e., text and path diagram, and text, table of coefficients, and path diagram conditions) because the path diagram uses symbols that map onto concrete concepts rather than abstract numbers.

Finally, it is expected that the condition of text, table of coefficients, and path diagram will have the lowest number of comprehensive interpretational errors because this condition provides essential numerical information as well as phenomenological information among the variables. Compared to other conditions, this condition includes all formats (i.e., text, table of coefficients, and path diagram) which satisfy individuals' different preference for processing complex information (Wilkinson & the Task Force on Statistical Inference, 1999).

1.4.2.2. Reader experience

Perceived time and ease. In general, participants' actual time spent decreases when they have less interpretational error (Agus et al., 2015), and the present study hypothesized that the text and table of coefficients condition would have less comprehensive interpretational error compared to the text only condition. In this vein, it is expected that the text and table of coefficients condition will have shorter perceived time compared to the text only condition. The representation of numerical information in a more organized way in the text and table of coefficients condition compared to the text only condition is also expected to increase readers' ease in understanding reported results. Thus, it is expected that the text and table of coefficients condition will have higher readers' ratings on ease compared to the text only condition.

The present study also hypothesized that the conditions with a path diagram (i.e., text and path diagram, and text, table of coefficients, and path diagram) would have lower comprehensive

interpretational errors compared to the conditions with no path diagram (i.e., text, and text and table of coefficients). In this vein, it is expected that the conditions with a path diagram (i.e., text and path diagram, and text, table of coefficients, and path diagram) will have shorter perceived time and higher readers' ratings on ease compared to the conditions with no diagram (i.e., text, and text and table of coefficients).

Finally, the present study hypothesized that the condition of text, table of coefficients, and path diagram would have the lowest number of comprehensive interpretational errors because this condition provides essential numerical information as well as conceptual representation in realistic objects, which will satisfy individuals' different ways of processing complex information. Thus, it is expected that the conditions with text, table of coefficients, and path diagram will have shorter perceived time and higher readers' ease compared to other conditions (i.e., text only, text and table of coefficients, and text and path diagram).

Satisfaction and Confidence. Similar to readers' perceived time and ease, it is expected that the text and table of coefficients condition will have higher ratings on satisfaction and confidence compared to the text only condition because the representation of numerical information with a table is more organized compared to the text only condition.

The present study also expects that the conditions with a path diagram (i.e., text and path diagram, and text, table of coefficients, and path diagram) will have higher ratings on satisfaction and confidence compared to the conditions with no diagram (i.e., text, and text and table of coefficients) because the conditions with a path diagram will have lower comprehensive interpretational errors compared to the conditions with no path diagram due to the realistic conceptual representation provided by the path diagram.

Finally, the present study expects that the conditions with text, table of coefficients, and path diagram will have higher ratings of satisfaction and confidence compared to other conditions (i.e., text only, text and table of coefficients, and text and path diagram) because this condition provides essential numerical information as well as conceptual representation in realistic objects, which will satisfy individuals' different ways of processing complex information.

1.4.3. Reader preference

Besides examining the effect of four different presentation formats (i.e., text only, text and table of coefficients, text and path diagram, and text, table of coefficients, and path diagram) on interpretational error and reader experience, reader preference for different presentation formats is examined.

Reader preference is a measure of which presentation format participants select as most useful for understanding reported results of a mediation analysis. The options are: text, a table of coefficients, and a path diagram. A path diagram is expected to be most frequently selected as a readers' preferred presentation format because the path diagram provides an accessible visualization of mediation analysis.

Chapter Two: Method

2.1. Participants

Participants of this study are recruited from a convenience sample of faculty members, postdoctoral fellows, and graduate students from the Department of Psychology at a large Canadian university. Current and future researchers in psychology are recruited because the target population for this study is researchers who use mediation. Participation was voluntary, and the participants were contacted by email via departmental listservs. Additionally, flyers soliciting participation were also posted on the walls of the building for the Department of Psychology in the university.

The majority of participants are graduate students who have some training in research (78.79%), who are female (73.44%), who are familiar with mediation (77.47%), and who have a mean age of 31.7 years ($SD = 11.21$). However, only 45.07% of participants indicate having used statistical methods to test mediation. Detailed descriptive information of the participants is presented in Table 1 on page 19.

2.2. Measures

2.2.1. Interpretational Error Scale

The Interpretational Error scale measures possible errors in interpreting the simple mediation results presented in the experimental stimuli (see Appendix A on page 64). There are two experimental phases that evaluate readers' understanding of basic information versus detailed information. Interpretational errors are measured at each phase. The first phase is involved in cursory interpretational understanding with basic information of mediation.

Table 1. Characteristics of Participants

| Characteristics | Total (<i>N</i> = 71) % | Missing |
|--|-----------------------------|---------|
| Gender | | 7 |
| Female | 73.44 | |
| Male | 15.00 | |
| Undeclared | 3.13 | |
| Age (<i>M</i> ± <i>SD</i>) ^a | 31.7±11.21 | 15 |
| Occupations | | 5 |
| Graduate student | 78.79 | |
| Postdoctoral fellow | 4.55 | |
| Faculty member | 16.67 | |
| Highest degree achieved | | 7 |
| Bachelor degrees | 21.88 | |
| Master degrees | 57.81 | |
| Doctorate degrees | 20.31 | |
| Conferred a doctorate degree | | 7 |
| No | 79.69 | |
| Yes | 20.31 | |
| Main activities at work | | 8 |
| Teaching | 6.35 | |
| Conducting research | 68.25 | |
| Providing therapy | 14.29 | |
| Others | 11.11 | |
| Primary research area | | 7 |
| Clinical | 9.52 | |
| Clinical developmental | 26.98 | |
| Cognitive / Neuroscience | 19.05 | |
| History and theory | 4.76 | |
| Quantitative methods | 3.17 | |
| Social / Personality | 28.57 | |
| Others | 4.76 | |
| Familiarity with mediation | | 0 |
| No | 22.54 | |
| Yes | 77.47 | |
| Familiarity with statistical methods to test mediation | | 0 |
| No | 54.93 | |
| Yes | 45.07 | |

Note. ^a Age is a continuous variable therefore means (standard deviations are presented).

The second phase involves comprehensive interpretational understanding with detailed information of reported results from mediation analyses. Cronbach's alpha for the Interpretational Error scale including both phases for the sample of the present study is .63.

2.2.1.1. Reading for basic information

The first two items of the Interpretational Error scale measured a cursory understanding of simple mediation results to reflect the context where researchers briefly glean research papers for general results. The items are "The effect of Self-Directedness on Positive Affect is mediated by Persistence." and "Self-Directedness significantly mediated the relationship between Persistence and Positive Affect." The response options for these items are either "True" or "False", which will be scored either 0 for selecting a correct response or 1 for not selecting a correct response. The sum of scores from these two items is used in the data analysis. Higher scores reflect a larger count of interpretational errors, and the total number of possible incorrect responses is two. Cronbach's alpha for items measuring cursory understanding in the Interpretational Error scale for the sample of the present study is .20. The low reliability of items measuring cursory understanding is possibly due to the skewedness of responses in the present study's sample.

2.2.1.2. Reading for detailed information

The next set of eleven items in the Interpretational Error scale examines a comprehensive understanding of presented results. These eleven items have varying response scales (see Appendix A on page 64). An example item with one correct response is, "What is the estimated effect of Self-Directedness (M) on Positive Affect (Y), controlling for the effect of Persistence (X)

on Positive Affect (Y)?” The response options for this example item are 0.23, 0.36, 0.26, 0.34, and 0.08, and the correct response is 0.36.

An example item with multiple correct responses is “From written results, what does $(.23 \times .36 = .08)$ represent? Please select all that apply.” The response options for this item are “The effect of Persistence (X) on Positive Affect (Y) through Self-Directedness (M)”, “The indirect effect of Persistence (X) on Positive Affect (Y)”, “The expected change in Positive Affect (Y) due to a one unit change in Persistence (X), controlling for Self-Directedness (M)”, “The estimated difference between the total and direct effect of Persistence (X) and Positive Affect (Y)”, and “The product of the effect of Persistence (X) on Self-Directedness (M), and the effect of Self-Directedness (M) on Positive Affect (Y), controlling for Persistence (X)”. All the response options presented are correct responses.

Responses for the 11 items are scored as either 0 for selecting a correct response or 1 for not selecting a correct response. Incorrect responses regardless of being selected or not in an item with multiple correct responses are given 0. Higher scores reflect a larger count of interpretational errors, and the total number of possible interpretational error is 27. Cronbach’s alpha for the items measuring comprehensive understanding in the Interpretational Error scale for the sample of the present study is .65.

2.2.2. Reader Experience

Five items were used to measure participants’ experience from interpreting the results presented in the experimental stimuli. Cronbach’s alpha for the items measuring reader experience in the reading for basic information and in the reading for detailed information with

the present study's sample is .89. However, each item measuring each construct of reader experience is analyzed separately.

2.2.2.1. *Perceived time and ease*

Three items were used to measure perceived time spent and perceived ease in understanding of the reported results. The item measuring perceived time is, "How the results were presented took more time than usual to understand them." The two items measuring readers' ease are, "How the results were presented made it difficult to understand them." and "How the results were presented made it easy to understand them." Items measuring readers' perceived time and ease are scored on a 5-point scale ranging from 1 to 5. (e.g., 1 as *Strongly Agree* versus 5 as *Strongly Disagree*). The present study used the opposite words of "difficult" and "easy" to measure the readers' ease in order to sample highly overlapping domains of the same construct of readers' ease. These two items measuring the readers' ease are expected to produce similar results.

2.2.2.2. *Satisfaction and confidence*

Two items were used to measure perceived satisfaction and perceived confidence separately in understanding of the reported results. The items are, "I am satisfied with my understanding of these results." and "I am confident in my understanding of these results." These items are scored on a 5-point scale ranging from 1 to 5 (e.g., 1 as *Strongly Agree* versus 5 as *Strongly Disagree*).

2.2.3. Reader Preference

Participants' preferred presentation format is measured with an item, "Which device is the most useful for understanding reported results of mediation analysis?" The response options for this item are "Text", "Table of coefficients", and "Path diagram".

2.2.4. Attitudes Toward Mediation Scale.

Questions were developed to measure general attitudes toward mediation. This scale consists of two subscales: familiarity with mediation and training experience in mediation. The present study only uses participants' responses of familiarity with mediation (See Appendix B on page 70).

2.2.4.1. *Familiarity with mediation*

Familiarity with mediation was measured in order to control for potential confounding effects of participants' familiarity with mediation in relation to the effect of different formats on interpretational accuracy and readers' experience. Familiarity with mediation is measured with two items. The items are, "Are you familiar with mediation?" and "Have you used any statistical method to test mediation?" These items are scored on a binary scale (*Yes / No*) where "Yes" is coded as 1 and "No" is coded as 0.

2.2.5. Demographics

The demographic questionnaire begins by measuring participants' academic status (Graduate student, Tenure-track faculty, or Contract faculty). Graduate students are asked to state their program of study (Masters or PhD), and what year they are in the program. Data are also collected on the following measures: highest education achieved and primary occupational

activity (e.g., teaching, conducting research, providing therapy, and other). For those who have graduated, they were asked to provide their graduating year in numbers (e.g. 2002), and whether they received a PhD or Psy.D. Additionally, questions on participants' primary research area, gender, and age are included (See Appendix C on page 76).

2.3. Design

This experimental study is a one-way between-subjects design, where the single factor has four levels, which are the different formats of presenting results from simple mediation.

2.3.1. Experimental stimuli.

The four different formats of presenting results from a simple mediation analysis have been piloted to ensure that participants do not require any specific knowledge about any particular area in psychology. Four different but comparable stimuli for each condition were developed based on the results from the empirical study by Garcia et al. (2012). Reflecting current practice, the four different combinations of formats used in this study are (i) no specific form other than presenting the results with text; (ii) text and a table with regression coefficients; (iii) text and a path diagram; or (iv) text, a table of coefficients, and a path diagram (see Appendix D on page 78).

All four conditions presented identical information of the results in a simple mediation analysis. However, within each condition, results were not duplicated across the different formats (Publication Manual of the American Psychological Association, 2010, p. 130, and p. 150). For instance, in the text only condition, the text contains four different types of information: (i) the concept of the simple mediation in the context of the research question; (ii) numerical estimates from MLR models (refer to Equations 1, 2, and 3 on page 4); (iii) numerical values for

the indirect effect, the z -statistic of the Sobel test, and the 95% percentile bootstrap CI; and (iv) interpretations of numerical estimates in the context of the research question (see Appendix D on page 78).

In the condition of the text and table of coefficients, the text contains only the concept of the simple mediation in the context of the research question, the numerical values for the indirect effect, the z -statistic of the Sobel test, the 95% percentile bootstrap CI, and the interpretations of numerical estimates in the context of the research question. The numerical estimates from MLR models are presented in the table of coefficients (see Appendix D on page 78).

In the condition of the text and path diagram, the text contains the concept of the simple mediation in the context of the research question, the numerical values for the indirect effect, the z -statistic of the Sobel test, the 95% percentile bootstrap CI, and the interpretations of numerical estimates in the context of the research question. The text in this condition is equivalent to the text in the condition of the text and table of coefficients. The numerical estimates from MLR models are presented in the path diagram.

Finally, in the condition of the text, table of coefficients, and path diagram, the text contains only the numerical values for indirect effect, the z -statistic of the Sobel test, the 95% percentile bootstrap CI, and the interpretations of numerical estimates in the context of the research question. The numerical estimates from MLR models are presented in the table of coefficients. The conceptual simple mediation model in the context of the research question without numerical estimates is presented in the path diagram (see Appendix D on page 78).

2.4. Procedure

The experiment was delivered through Qualtrics, which is a web-based software that can create surveys and collect data (Qualtrics.com). Participation in the experiment was performed through the Qualtrics website, which was accessed by the link provided in an email soliciting participation. Participants took part in the experiment individually, at their convenience, and without any restrictions regarding time and location when participating in the experiment. The order of the procedure is depicted in Figure 2.

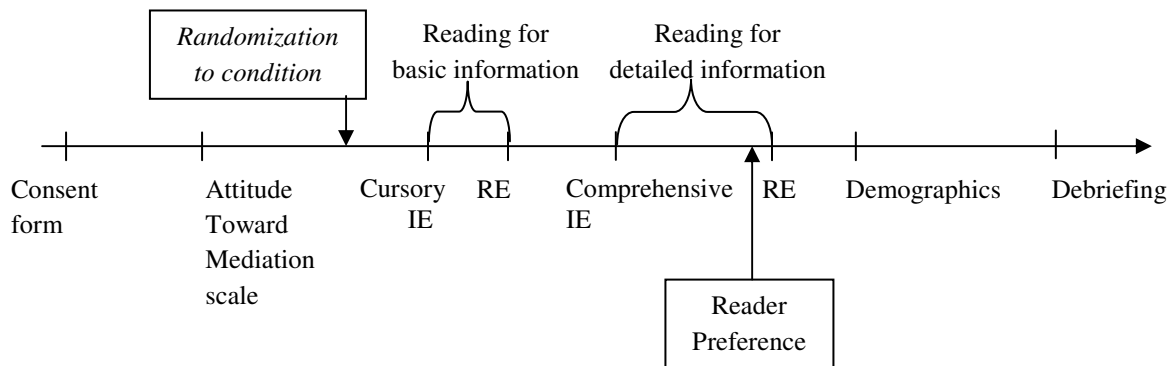


Figure 2. Procedure of the experiment.

IE = Interpretational Error scale; RE = Reader Experience items

Clicking on the link of the study will lead participants to the informed consent form (see Appendix E on page 85). After giving their consent, participants will fill out the Attitude Toward Mediation scale. Upon scale completion, participants are randomly assigned to one of the four conditions: (i) text only condition; (ii) text and table of coefficients condition; (iii) text and path diagram condition; and (iv) text, table of coefficients, and path diagram condition. As part of the instructions to this section of the study, the Interpretational Error scale includes a link to one of the randomly assigned formats of reported results.

When a participant copies and pastes the link to the experimental stimulus on the address box of an internet browser, which is different from the browser with the questionnaire, one of the four formats of reported results is shown as an Adobe Acrobat PDF file. As researchers often read journal articles in Adobe Acrobat PDF files, the present study provides the stimuli in Adobe Acrobat PDF file to reflect current practice. The items measuring cursory understanding are presented first, followed by the items measuring comprehensive understanding. The items measuring participants' experience are presented twice, after the items measuring cursory understanding and also after the items measuring comprehensive understanding (see Figure 2 on page 26).

The last section of the experimental study has participants' demographics questions, which are presented after participants complete the Interpretational Error scale and reader experience items. After participants complete the demographic questions, debriefing information is presented (see Appendix F on page 88). The debriefing information contains the aim of the study, the procedure of the experiment, expected results, the contact information of the primary investigator, and two references related to this study. On average, participants took 20 to 30 minutes to complete the study.

Chapter Three: Results

3.1. Preliminary Analysis

3.1.1. Missing data

The total number of participants who provided data was 98. Participants were removed when no response was recorded in the Interpretational Error scales but responses were only recorded in either the Attitude Toward Mediation scale ($n = 10$) or demographic questions ($n = 5$). If the responses of items measuring interpretational error in both cursory and comprehensive understanding were missing, these participants were removed even if responses of other items were recorded ($n = 2$). If the participants did not have any responses on familiarity and interpretational error in both phases of reading for basic and detailed information, these participants were also removed ($n = 10$).

If a participant responded to both items measuring cursory interpretational error in the phase of reading for basic information and the items measuring reader experience in the phase of reading for detailed information, this participant was included even if no responses were recorded in items measuring comprehensive interpretational error in reading for detailed information. These missing values for items measuring interpretational error in the phase of reading for detailed information were treated as participants' not selecting correct answers because this participant stopped responding when he or she faced the questionnaires asking for comprehensive understanding. Yet, this participant still responded to the items measuring reader experience in reading for detailed information that were placed after the items measuring comprehensive interpretational error in the reading for detailed information ($n = 1$). The missing values of this participant in items measuring comprehensive interpretational error in the reading

for detailed information were coded as 1 where 1 indicates an incorrect response. The final number of cases included in the data analyses is 71.

3.1.2. Correlations

Recall that there are two experimental phases where readers' understanding of basic information versus detailed information of the results is evaluated. In the first phase, cursory interpretational error reflecting basic information is measured. In the second phase, comprehensive interpretational error reflecting detailed information of the reported results is measured.

The association between the interpretational errors and the reader experience is examined only within each phase of reading for basic information versus for detailed information. The correlation coefficients among interpretational errors and reader experience are presented in Table 2 on page 30.

3.1.2.1. *Within the phase of reading*

[1] Reading for basic information

There were weak point biserial correlations between cursory interpretational error (IE_{bas}) and with perceived time ($TIME_{bas}$) and perceived ease ($EASE_{bas}$), $r_{IE_{bas}, TIME_{bas}} = -.03, p = ns$; and $r_{IE_{bas}, EASE_{bas}} = -.03, p = ns$. The point biserial correlation between cursory interpretational error and perceived difficulty ($DIFF_{bas}$) was small, $r_{IE_{bas}, DIFF_{bas}} = .16, p = ns$. Although difficulty and ease were correlated highly at $r_{DIFF_{bas}, EASE_{bas}} = -.85, p < .01$, their different correlations with cursory interpretational error warrant the separate treatments for difficulty and ease in the primary analysis.

Table 2. Correlation Between Interpretational Error and Reader Experience

| | (n = 71) | (n = 70) | (n = 70) | (n = 69) | (n = 70) | (n = 70) | (n = 71) | (n = 63) | (n = 63) | (n = 63) | (n = 62) | (n = 62) |
|---|----------|---------------|---------------|--------------|--------------|--------------|----------|---------------|---------------|----------|--------------|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Cursory Understanding | | | | | | | | | | | | |
| 1. <i>Interpretational Error</i> ^a | 1.00 | | | | | | | | | | | |
| 2. Perceived Time | -.03 | 1.00 | | | | | | | | | | |
| 3. Difficulty | .16 | .81** | 1.00 | | | | | | | | | |
| 4. Ease | -.03 | -.74** | -.85** | 1.00 | | | | | | | | |
| 5. Satisfaction | -.34** | -.19 | -.33** | .29* | 1.00 | | | | | | | |
| 6. Confidence | -.30** | -.16 | -.30* | .33** | .84** | 1.00 | | | | | | |
| Detail Understanding | | | | | | | | | | | | |
| 7. <i>Interpretational Error</i> | .28* | . | . | . | . | . | 1.00 | | | | | |
| 8. Perceived Time | . | .74** | .55** | -.53** | -.10 | -.03 | .01 | 1.00 | | | | |
| 9. Difficulty | . | .72** | .67** | -.67** | -.12 | -.06 | .07 | .85** | 1.00 | | | |
| 10. Ease | . | -.64** | -.70** | .71** | .23 | .12 | -.19 | -.72** | -.85** | 1.00 | | |
| 11. Satisfaction | . | -.06 | -.16 | .23 | .66** | .70** | -.43** | -.19 | -.22 | .27* | 1.00 | |
| 12. Confidence | . | -.10 | -.21 | .29* | .65** | .75** | -.35** | -.14 | -.20 | .24 | .89** | 1.00 |

Note. ^aBiserial correlation used between interpretational error in the cursory understanding and others. * $p < .05$, ** $p < .01$

Cursory interpretational error also has moderate point biserial correlations with satisfaction and confidence, $r_{IE_{bas}, SAT_{bas}} = -.34, p < .01$; and $r_{IE_{bas}, CON_{bas}} = -.30, p < .01$. Similar to the findings of Agus et al. (2015), these correlations suggest that when participants made more cursory interpretational errors, they would have less satisfaction and confidence.

The patterns of correlations among the items measuring reader experience show two distinct subsets. Specifically, the items of perceived time, difficulty, and ease form one cluster, whereas, the items of satisfaction and confidence form the second cluster. Within each subset, the correlations are strong. However, the correlations are weak between subsets.

[2] Reading for detailed information

Similar to the correlations within the phase of the reading for basic information, weak correlations between the comprehensive interpretational error (IE_{det}) and perceived time ($TIME_{det}$), difficulty ($DIFF_{det}$), and ease ($EASE_{det}$) were observed, $r_{IE_{det}, TIME_{det}} = .01, p = ns$; $r_{IE_{det}, DIFF_{det}} = .07, p = ns$; and $r_{IE_{det}, EASE_{det}} = -.19, p = ns$. Although difficulty and ease were correlated highly at $r_{DIFF_{det}, EASE_{det}} = -.85, p < .01$, their different correlations with comprehensive interpretational error warrant the separate treatments for difficulty and ease in the primary analysis. Comprehensive interpretational error also has moderate correlations with satisfaction and confidence, $r_{IE_{det}, SAT_{det}} = -.43, p < .01$; and $r_{IE_{det}, CON_{det}} = -.35, p < .01$. Similar to the findings of Agus et al. (2015), these correlations suggest that when participants made more comprehensive interpretational errors, they would have less satisfaction and confidence.

The patterns of correlations among the items measuring reader experience after being evaluated on detailed information also show two distinct subsets. The items of perceived time, difficulty, and ease form one cluster, whereas, the items of satisfaction and confidence form the

second cluster. The correlations within each subset are strong. However, the correlations are weak between subsets.

3.1.2.2. *Between reading for basic information and reading for detailed information*

[1] Interpretational error

The point biserial correlation between the interpretation error measuring cursory understanding in the reading for basic information (IE_{bas}) and the interpretational error measuring comprehensive understanding (IE_{det}) in the reading for detailed information is positive, and moderate, $r_{IE_{bas}, IE_{det}} = .28, p < .05$. This moderate relationship between the interpretation error measuring cursory understanding and the interpretational errors measuring comprehensive understanding suggests that results for these outcomes would be distinct.

[2] Interpretational error versus Reader experience

The pattern of correlations among variables representing constructs of reader experience between reading for basic information and reading for detailed information show similar relationships within the phases of reading. First, each of five variables in reader experience (i.e., perceived time, difficulty, ease, satisfaction, and confidence) in the reading for basic information is strongly correlated to the same constructs in the reading for detailed information. For instance, perceived time in reading for basic information is strongly related to the perceived time in reading for detailed information (see Table 2 on page 30).

Second, the reader experience variables were relatively stable across the two different phases which evaluate cursory and comprehensive understanding of the reported results from simple mediation. The correlations of the same variable between the phases (e.g., $TIME_{bas}$,

TIME_{det}) ranged from .64 to .75. These correlations are situated in the diagonal of the lower left quadrant of Table 2 on page 30. The pattern of correlations for the two clusters of reader experience items between the different phases which evaluated basic and detailed information suggested stability of these clusters.

Despite the stability of constructs in reader experience, the magnitudes of linear relationships between difficulty and interpretational error within the experimental phases of reading for basic information versus reading for detailed information are different from the magnitudes of linear relationships between ease and interpretational error within each experimental phases of reading. Within the phase of reading for basic information, the point biserial correlation between interpretational error (IE_{bas}) and difficulty (DIFF_{bas}) is stronger compared to the point biserial correlation between interpretational error and ease (EASE_{bas}), $r_{DIFF_{bas}, IE_{bas}} = .16, p = ns$; and $r_{EASE_{bas}, IE_{bas}} = -.03, p = ns$, respectively. However, within the phase of reading for detailed information, the correlation between interpretational error (IE_{det}) and difficulty (DIFF_{det}) is weaker compared to the correlation between interpretational error and ease (EASE_{det}), $r_{DIFF_{det}, IE_{det}} = .07, p = ns$; and $r_{EASE_{det}, IE_{det}} = -.19, p = ns$. Thus, items measuring difficulty and ease were tested separately in the data analyses because the item using the word “difficult” and the item using the word “easy” may not measure the same construct.

3.2. Primary Analysis

3.2.1. Data analysis plan

The present study examines the effect of four different presentation formats on the participants’ interpretational errors and reader experience. The following covariates of familiarity with the concept of mediation (FAM_{ex}), familiarity with the statistical methods to test

mediation (FAMst), and being conferred with a doctorate degree or not (PhD; no doctoral degree being conferred is coded as 0, and a doctoral degree being conferred as 1) are included in the primary analysis.

Planned contrasts in the ANCOVA model are used to examine the hypotheses regarding the dependent variables of interpretational error measured in the phase of evaluating detailed understanding and reader experience across both experimental phases of understanding.

An ANCOVA model, defined below, is fit to the data.

$$DV = B_0 + B_1FAMex + B_2FAMst + B_3PhD + B_4COND + e, \quad (6)$$

where DV denotes the dependent variables of the comprehensive interpretational error and reader experience measured in the two experimental phases; B_0 is the intercept, B_1 is the difference in the DV between high and low in familiarity with mediation, B_2 is the difference in the DV between high and low in familiarity with using statistical methods to test mediation, B_3 is the difference in the DV between having a doctorate degree and having no doctorate degree, COND is the condition of the four different combinations of presentation formats, and e is the error term.

Interpretational error in cursory understanding is not normally distributed because it only has three values (i.e., 0, 1, 2), which represent the number of incorrect answers. Only 3 participants had incorrect responses on all items measuring interpretational error in cursory understanding which has two items measuring interpretational error. Nineteen participants had incorrect responses in one of two items, and 49 participants had no incorrect response in both two items. Thus, the present study dichotomized the interpretational error in cursory understanding into “High” and “Low” where “High” included participants with any incorrect responses, and performed the planned contrasts in the logistic regression model. The logistic

regression modeled the probability of having high compared to low interpretational error. The predictors of the logistic model mirror those of Equation 6 on page 34.

The assumptions of normality and equal variance were examined for the ANCOVA models using Q-Q plots and residual plots. No extreme departure from normality and equal variance was observed.

3.2.2. Reading for basic information

Predicted means and their standard errors are presented in Table 3 below.

Table 3. Predicted Means from the Statistical Models and their Significance in Reading for Basic Information

| | Text only ^a | Text & Table ^a | Text & Diagram ^b | All formats ^b |
|-------------------------------------|------------------------|---------------------------|-----------------------------|--------------------------|
| | $M_{Predicted} (SE)$ | $M_{Predicted} (SE)$ | $M_{Predicted} (SE)$ | $M_{Predicted} (SE)$ |
| Interpretational error ^c | .49 (1.59) | .49 (1.59) | .42 (1.65) | .42 (1.65) |
| Reader experience ^d | | | | |
| Perceived time | 3.67 (.23) | 3.67 (.23) | 3.21 (.25) | 3.21 (.25) |
| Difficulty | 3.62 (.24) | 3.62 (.24) | 3.17 (.25) | 3.17 (.25) |
| Ease | 2.16 (.21) | 2.16 (.21) | 2.60 (.22) | 2.60 (.22) |
| Satisfaction | 3.07 (.22) | 3.07 (.22) | 2.92 (.23) | 2.92 (.23) |
| Confidence | 2.79 (.22) | 2.79 (.22) | 2.71 (.23) | 2.71 (.23) |

Note. ^a Conditions with no path diagram. ^b Conditions with a path diagram. ^c Predicted values are odds ratios. ^d 1 = “Strongly Disagree”; 2 = “Disagree”; 3 = “Neither Disagree nor Agree”; 4 = “Agree”; 5 = “Strongly Agree”

3.2.2.1. Interpretational error

[1]Covariates

The covariates of FAMex, FAMst, and PhD are not statistically significant in the logistic regression model¹, $B_1 = .57$, $t(59) = .70$, $p = .4865$; $B_2 = -1.22$, $t(59) = -1.61$, $p = .1124$; and $B_3 = 1.32$, $t(59) = 1.74$, $p = .0866$, respectively. There is no significant effect of familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree on interpretational error.

[2] Contrast: Conditions with no path diagram versus Conditions with a path diagram

The contrast of the conditions with a path diagram minus the conditions with no path diagram is also not statistically significant. The log odds estimate of difference in the contrast between the conditions with a path diagram and the conditions with no path diagram is $-.08$, $t(59) = -.24$, $p = .8100$, 95% CI $[-.66, .50]$. The odds of having higher interpretational error are .92 times lower in the conditions with a path diagram compared to the odds of having higher interpretational error in the conditions with no path diagram. No significant difference between the conditions of no path diagram and the conditions of a path diagram is observed.

3.2.2.2. Reader experience

(1) Perceived time, difficulty, and ease

[1]Covariates

Perceived time. The covariates of FAMex and FAMst are not significant in predicting perceived time, $B_1 = -.35$, $t(59) = -.92$, $p = .3627$; and $B_2 = -.09$, $t(59) = -.26$, $p = .7930$,

¹ Coefficient estimates B_1 , B_2 , and B_3 are log odds from the logistic regression.

respectively. The effects of familiarity with simple mediation and familiarity with using statistical methods to test mediation are not significant on perceived time. On the other hand, the effect of having a doctorate degree on participants' perceived time is significant, $B_3 = .86$, $t(59) = 2.41$, $p = .0193$. Participants with a doctoral degree have .86 units of longer perceived time, 95% CI [.14, 1.57], after they responded to items measuring interpretational error for a cursory understanding compared to the participants without a doctoral degree.

Difficulty. The covariates of FAMex and FAMst are not statistically significant in predicting difficulty, $B_1 = -.05$, $t(59) = -.12$, $p = .9088$; and $B_2 = -.37$, $t(59) = -1.11$, $p = .2718$, respectively. The effects of familiarity with simple mediation and familiarity with using statistical methods to test mediation are not significant on difficulty. On the other hand, the effect of having a doctorate degree on participants' difficulty is significant, $B_3 = .89$, $t(59) = 2.47$, $p = .0164$. Participants with a doctoral degree have .89 units of higher difficulty, 95% CI [.17, 1.61], after they responded to items measuring interpretational error for a cursory understanding compared to the participants without a doctoral degree.

Ease. The covariates of FAMex and FAMst are not statistically significant in predicting ease, $B_1 = .28$, $t(59) = .80$, $p = .4282$; and $B_2 = .01$, $t(59) = .05$, $p = .9639$, respectively. The effects of familiarity with simple mediation and familiarity with using statistical methods to test mediation are not significant on ease. On the other hand, the effect of having a doctorate degree on participants' ease is significant, $B_3 = -.75$, $t(59) = -2.34$, $p = .0229$. Participants with a doctoral degree have .75 units of less ease, 95% CI [-1.39, -.11], after they responded to items measuring interpretational error for a cursory understanding compared to the participants without a doctoral degree.

[2] Contrast: Conditions with no path diagram versus Conditions with a path diagram

Perceived time. The mean difference in perceived time between the conditions with a path diagram minus the conditions with no path diagram is -.23 units 95% CI[-.51, .05], $t(59) = -1.64$, $p = .1063$. There was no significant effect of a path diagram on perceived time after being evaluated on cursory understanding of reported results.

Difficulty. The mean difference in difficulty between the conditions with a path diagram minus the conditions with no path diagram is -.22 units 95% CI[-.51, .06], $t(59) = -1.56$, $p = .1234$. There was no significant effect of a path diagram on difficulty after being evaluated on cursory understanding of reported results.

Ease. The mean difference in ease between the conditions with a path diagram minus the conditions with no path diagram is .22 units 95% CI[-.03, .48], $t(59) = 1.78$, $p = .0807$. There was no significant effect of a path diagram on ease after being evaluated on cursory understanding of reported results.

In summary, no significant difference between the conditions with a path diagram and the conditions with no path diagram is observed on perceived time, difficulty, and ease when participants were evaluated with cursory understanding of the reported results. There was insufficient evidence to conclude an effect of a path diagram on participants' perceived time, difficulty, and ease in reading results for basic information.

(2) Satisfaction and confidence

[1]Covariates

Satisfaction. The covariates of FAMex, FAMst, and PhD are not statistically significant in predicting satisfaction, $B_1 = .41$, $t(59) = 1.13$, $p = .2640$; $B_2 = .33$, $t(59) = 1.07$, $p = .2890$; and $B_3 = -.36$, $t(59) = -1.08$, $p = .2860$, respectively. There is no significant effect of familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree on satisfaction after being evaluated for a cursory understanding of the reported results.

Confidence. The covariates of FAMex, FAMst, and PhD are not statistically significant in predicting confidence, $B_1 = .29$, $t(59) = .79$, $p = .4340$; $B_2 = .39$, $t(59) = 1.25$, $p = .2170$; and $B_3 = -.46$, $t(59) = -1.37$, $p = .1750$, respectively. There is no significant effect of familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree on confidence after being evaluated for a cursory understanding of the reported results.

[2] Contrast: Conditions with no path diagram versus Conditions with a path diagram

Satisfaction. The mean difference in satisfaction between the conditions with a path diagram minus the conditions with no path diagram is $-.08$ units 95% CI[$-.34$, $.19$], $t(59) = -.59$, $p = .5560$. There is nearly no effect of a path diagram on satisfaction after being evaluated for a cursory understanding of the reported results.

Confidence. The mean difference in confidence between the conditions with a path diagram minus the conditions with no path diagram is $-.04$ units 95% CI[$-.30$, $.23$], $t(59) = -.30$,

$p = .7670$. There is nearly no effect of a path diagram on confidence after being evaluated for a cursory understanding of the reported results.

In summary, no significant difference between the conditions with a path diagram and the conditions with no path diagram is observed for both satisfaction and confidence when participants were evaluated for their cursory understanding of the reported results. There was insufficient evidence to conclude an effect of a path diagram on participants' satisfaction and confidence in reading results for basic information.

3.2.3. Reading for detailed information

Three planned contrasts were examined under the phase of evaluating detailed information are: (i) the condition with text and table of coefficients versus the condition with text only; (ii) the condition with a path diagram (i.e., text and path diagram, and text, table of coefficients, and path diagram) versus the conditions with no path diagram (i.e., text only, and text and table of coefficients); and (iii) the condition with all three formats are included (i.e., text, table of coefficients, and path diagram) versus the other conditions (i.e. text only, text and table of coefficients, and text and path diagram). Predicted means and their standard errors are presented in Table 4 on page 41.

3.2.3.1. Interpretational error

[1] Covariates

The covariates of FAMex, FAMst, and PhD are not statistically significant in predicting comprehensive interpretational error, $B_1 = -2.17$, $t(57) = -1.20$, $p = .2365$; $B_2 = .81$, $t(57) = .51$, $p = .6110$; and $B_3 = 1.49$, $t(57) = .88$, $p = .3805$, respectively.

Table 4. Predicted Means from the Statistical Models and their Significance in Reading for Detailed

| Information | Text only | Text & Table | Text & Diagram | All formats |
|--------------------------------|----------------------|----------------------|----------------------|----------------------|
| | $M_{Predicted} (SE)$ | $M_{Predicted} (SE)$ | $M_{Predicted} (SE)$ | $M_{Predicted} (SE)$ |
| Interpretational error | 14.76 (1.39) | 18.48 (1.50) | 15.84 (1.47) | 17.00 (1.49) |
| Reader experience ^a | | | | |
| Perceived time | 4.04 (.28) | 3.65 (.32) | 3.23 (.31) | 3.53 (.31) |
| Difficulty | 4.14 (.27) | 3.65 (.30) | 3.22 (.29) | 3.26 (.29) |
| Ease | 1.99 (.26) | 2.09 (.29) | 2.42 (.28) | 2.50 (.28) |
| Satisfaction | 2.24 (.28) | 2.13 (.32) | 2.46 (.30) | 2.11 (.30) |
| Confidence | 2.08 (.27) | 1.99 (.31) | 2.43 (.30) | 2.07 (.30) |

Note. ^a 1 = “Strongly Disagree”; 2 = “Disagree”; 3 = “Neither Disagree nor Agree”; 4 = “Agree”; 5 = “Strongly Agree”

There is no significant effect of familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree on interpretational error measuring comprehensive understanding in reported results.

[2] Contrast 1: Text only versus Text and table of coefficients

The mean difference in comprehensive interpretational error between the text and a table of coefficients condition minus the text only condition is 1.86 units 95% CI[-3.72, .01], $t(57) = 2.00$, $p = .0507$. This difference is significant, where the text only condition is associated with less comprehensive interpretational error compared to the text and table of coefficients condition.

[3] Contrast 2: Conditions with no path diagram versus Conditions with a path diagram

The mean difference in interpretational error between the conditions with a path diagram minus the conditions with no path diagram is $-.78$ units 95% CI $[-2.37, 3.93]$, $t(57) = -.50$, $p = .6228$. The effect of a path diagram on comprehensive interpretational error is nonsignificant.

[4] Contrast 3: Text, table of coefficients, and path diagram versus Other conditions

The mean difference in interpretational error between the text, table of coefficients, and path diagram condition minus the other conditions (i.e., text only, text and table of coefficients, and text and path diagram) is $.86$ units 95% CI $[-3.65, 1.92]$, $t(57) = .62$, $p = .5372$. The effect of having all types of formats on comprehensive interpretational error is nonsignificant.

In summary, only a significant effect of the text only condition versus the text and table of coefficients condition on comprehensive interpretational error is observed.

3.2.3.2. Reader experience

(1) Perceived time, difficulty, and ease

[1]Covariates

Perceived time. The covariates of FAMex, FAMst, and PhD are not statistically significant in predicting perceived time, $B_1 = -.48$, $t(55) = -1.28$, $p = .2072$; $B_2 = .14$, $t(55) = .45$, $p = .6583$; and $B_3 = .42$, $t(55) = 1.22$, $p = .2283$, respectively. There is no significant effect of familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree on perceived time when participants were tasked to read for detailed information from the reported results.

Difficulty. The covariates of FAMex, FAMst, and PhD are not statistically significant in predicting difficulty, $B_1 = -.38$, $t(55) = -1.07$, $p = .2901$; $B_2 = .01$, $t(55) = .02$, $p = .9836$; and $B_3 = .52$, $t(55) = 1.60$, $p = .1154$, respectively. There is no significant effect of familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree on difficulty after reading the results for detailed information.

Ease. The covariates of FAMex, FAMst, and PhD are not statistically significant in predicting ease, $B_1 = .33$, $t(55) = .98$, $p = .3320$; $B_2 = -.06$, $t(55) = -.20$, $p = .8400$; and $B_3 = -.49$, $t(55) = -1.56$, $p = .1270$, respectively. There is no significant effect of familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree on ease after reading the results for detailed information.

[2] *Contrast 1: Text only versus Text and table of coefficients*

Perceived time. The mean difference in perceived time between the text and a table of coefficients condition minus the text only condition is $-.20$ units 95% CI $[-.58, .19]$, $t(55) = -1.01$, $p = .3176$. Although the difference is not significant, the observed results support the hypothesis of the text and table of coefficients condition having shorter perceived time compared to the text only condition.

Difficulty. The mean difference in difficulty between the text and a table of coefficients condition minus the text only condition is $-.24$ units 95% CI $[-.61, .12]$, $t(55) = -1.33$, $p = .1904$. Although the difference is not significant, the observed results support the hypothesis of the text and table of coefficients condition having less difficulty compared to the text only condition.

Ease. The mean difference in ease between the text and a table of coefficients condition minus the text only condition is $.05$ units 95% CI $[-.30, .40]$, $t(55) = .29$, $p = .7730$. The

difference between the text and a table of coefficients condition and the text only condition is not significant.

[3] *Contrast 2: Conditions with no path diagram versus Conditions with a path diagram*

Perceived time. The mean difference in perceived time between the conditions with a path diagram minus the conditions with no path diagram is -.61 units 95% CI[-.63, .03], $t(55) = -1.84$, $p = .0711$. Participants in the conditions with a path diagram (i.e., text and path diagram and text, table of coefficients, and path diagram) have .61 units of shorter perceived time compared to the participants in the conditions with no path diagram (i.e., text only and text and table of coefficients).

Difficulty. The mean difference in difficulty between the conditions with a path diagram minus the conditions with no path diagram is -.67 units 95% CI[-.65, -.03], $t(55) = -2.18$, $p = .0334$. The effect of a path diagram on difficulty is significant. Participants in the conditions with a path diagram (i.e., text and path diagram and text, table of coefficients, and path diagram) have .67 units of less difficulty compared to the participants in the conditions with no path diagram (i.e., text only and text and table of coefficients). The observed results support our expectation of the condition with a path diagram having less difficulty compared to the conditions with no path diagram.

Ease. The mean difference in ease between the conditions with a path diagram minus the conditions with no path diagram is .39 units 95% CI[-.10, .49], $t(55) = 1.30$, $p = .1980$. Although the effect of a path diagram is not significant, the observed result supports the hypothesis of the condition with a path diagram having higher ease compared to the conditions with no path diagram.

[4] Contrast 3: Text, table of coefficients, and path diagram versus Other conditions

Perceived time. The mean difference in perceived time between the text, table of coefficients, and path diagram condition minus the other conditions (i.e., text only, text and table of coefficients, and text and path diagram) is .22 units 95% CI[-.12, .27], $t(55) = .75$, $p = .4547$. The difference between the text, table of coefficients, and path diagram condition and the other conditions on participants' perceived time is not significant.

Difficulty. The mean difference in difficulty between the text, table of coefficients, and path diagram condition minus the other conditions (i.e., text only, text and table of coefficients, and text and path diagram) is .03 units 95% CI[-.17, .19], $t(55) = .11$ and $p = .9119$. The difference between the text, table of coefficients, and path diagram condition and the other conditions on participants' difficulty is not significant.

Ease. The mean difference in ease between the text, table of coefficients, and path diagram condition minus the other conditions (i.e., text only, text and table of coefficients, and text and path diagram) is .06 units 95% CI[-.16, .19], $t(55) = .23$ and $p = .8200$. The difference between the text, table of coefficients, and path diagram condition and the other conditions on participants' ease is not significant.

In summary, there is a significant effect where less perceived difficulty is observed with the presence of a path diagram when participants were reading results for detailed information. However, there are no significant effects of the text only condition versus text and table of coefficients condition as well as no significant effects of the text, table of coefficients, and a path diagram condition versus the other conditions on participants' perceived time, difficulty, and ease.

(2) *Satisfaction and confidence*

[1] *Covariates*

Satisfaction. The covariates of FAMex, FAMst, and PhD are not statistically significant in predicting satisfaction, $B_1 = .66$, $t(54) = 1.76$, $p = .0836$; $B_2 = -.08$, $t(54) = -.24$, $p = .8106$; and $B_3 = -.24$, $t(54) = -.70$, $p = .4896$, respectively. There is no significant effect of familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree on satisfaction in the reading for detailed information.

Confidence. The covariates of FAMex, FAMst, and PhD are not statistically significant in predicting confidence, $B_1 = .35$, $t(54) = .94$, $p = .3520$; $B_2 = .28$, $t(54) = .89$, $p = .3840$; and $B_3 = -.18$, $t(54) = -.53$, $p = .5980$, respectively. There is no significant effect of familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree on confidence when participants were reading results for detailed information.

[2] *Contrast 1: Text only versus Text and table of coefficients*

Satisfaction. The mean difference in satisfaction between the text and a table of coefficients condition minus the text only condition is $-.05$ units 95% CI $[-.45, .34]$, $t(54) = -.28$, $p = .7834$. The difference between the text and a table of coefficients condition and the text only condition on participants' satisfaction is not significant.

Confidence. The mean difference in confidence between the text and a table of coefficients condition minus the text only condition is $-.05$ units 95% CI $[-.44, .34]$, $t(54) = -.23$, $p = .8170$. The difference between the text and a table of coefficients condition and the text only condition on participants' confidence is not significant.

[3] *Contrast 2: Conditions with no path diagram versus Conditions with a path diagram*

Satisfaction. The mean difference in satisfaction between the conditions with a path diagram minus the conditions with no path diagram is .28 units 95% CI[-.19, .47], $t(54) = .85$, $p = .4001$. The difference between the conditions with a path diagram and the conditions with no path diagram on participants' satisfaction is not significant.

Confidence. The mean difference in confidence between the conditions with a path diagram minus the conditions with no path diagram is .40 units 95% CI[-.12, .52], $t(54) = 1.25$, $p = .2180$. The difference between the conditions with a path diagram and the conditions with no path diagram on participants' confidence is not significant.

[4] *Contrast 3: Text, table of coefficients, and path diagram versus Other conditions*

Satisfaction. The mean difference in satisfaction between the text, table of coefficients, and path diagram condition minus the other conditions (i.e., text only, text and table of coefficients, and text and path diagram) is -.26 units 95% CI[-.28, .10], $t(54) = -.92$, $p = .3625$. The difference between the text, table of coefficients, and path diagram condition and the other conditions on participants' satisfaction is not significant.

Confidence. The mean difference in satisfaction between the text, table of coefficients, and path diagram condition minus the other conditions (i.e., text only, text and table of coefficients, and text and path diagram) is -.27 units 95% CI[-.28, .10], $t(54) = -.97$, $p = .3340$. The difference between the text, table of coefficients, and path diagram condition and the other conditions on participants' confidence is not significant.

In summary, no significant results with respect to satisfaction and confidence is observed when participants were reading the results on simple mediation for detailed information. Although not significant, the results show opposite directions to what was expected in satisfaction and confidence, but close to zero. This study expected that participants in the text, table of coefficients, and path diagram condition would have higher satisfaction and confidence in understanding reported results compared to other conditions.

3.2.4. Reader Preference

The frequency of readers' preference in the type of presentation format is in Table 5 below. As hypothesized, among text, a table of coefficients, and a path diagram, the path diagram has the highest frequency of being selected as most useful presentation format to understand reported results in mediation analysis in general. The results in Table 5 only include the responses of participants who responded "Yes" to the item of "Are you familiar with mediation?"

Table 5. Frequency of the Preference in Presentation Format

| Preference | Count | Total ($N = 49$) % |
|-----------------------|-------|-------------------------|
| Text only | 6 | 10.90 |
| Table of coefficients | 4 | 7.30 |
| Path diagram | 39 | 89.10 |

The chi-square test of goodness of fit was $\chi^2(2) = 47.31, p < .001$, implying that some formats were endorsed more often than others. The path diagram is 6.5 times more likely to be selected as the preferred format compared to the text only, odds ratio estimate = 6.5, $p = .05$,

95% CI[.96, 5.33]. The path diagram is also 9.75 times more likely to be selected as the preferred format compared to the table of coefficients, odds ratio estimate = 9.75, $p = .05$, 95% CI[.96, 7.52]. However, the chi-square test of independence is nonsignificant when conditions of different formats are crossed with preference, implying that there is no significant association between participants' preferred format and the conditions of different formats, $\chi^2(6) = 7.18$, $p = .3040$.

Chapter Four: Discussion

Simple mediation is a highly popular model in psychological research as it is primarily used to investigate underlying causal mechanisms. After developing a simple mediation model in the context of a research question, researchers use statistical methods to test mediation. At least 14 approaches for testing simple mediation have been proposed (MacKinnon et al., 2002), and new developments of statistical methods to test mediation (i.e., mediation analysis) are continuously proposed (Preacher, 2015). The large number of approaches to test simple mediation adds to the variability in what and how results from simple mediation are reported (Gelfand et al., 2009). The present study focused on the variability in how results from simple mediation are reported (i.e., using combinations of text, table of coefficients, and path diagram).

Despite the high variability in how results from simple mediation are reported in psychological research, no standards or recommendations on how best to report results from simple mediation have been developed. It is important to develop recommendations for reporting results of simple mediation because high variability in reporting results of the simple mediation model can potentially increase interpretational errors associated with the simple mediation model. Additionally, how results are presented is also likely to influence readers' experience in understanding the study's conclusions. Thus, the present study examined whether different formats of presenting simple mediation results would affect the interpretational error and reader experience in interpreting the reported results of mediation analyses. The results from the present study are used to develop recommendations for how best to report results of a simple mediation to reduce interpretational error and facilitate positive reader experience.

Recall that four different combinations of formats where participants were randomly assigned to are (i) no specific form other than presenting the results with text; (ii) text and a table with regression coefficients; (iii) text and a path diagram; or (iv) text, a table of coefficients, and a path diagram. To control the effect of individual differences on interpretational error and reader experience, participants' familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree or not were also examined. The present study also incorporated two different contexts of reading research papers into two experimental phases: (i) reading for basic information; and (ii) reading for detailed information.

4.1. Reading for Basic Information

In reading for basic information, participants were asked to respond to the items measuring cursory understanding of simple mediation in the context of research presented in the stimuli. Then, participants were asked to respond to items measuring reader experience (see Figure 2 on page 26).

4.1.1. Covariates

Two different patterns were observed in the covariates when participants were gleaning for basic information. Participants' familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree or not did not have any effect on cursory interpretational error, satisfaction, and confidence while reading for basic information. On the other hand, the effect of having a doctorate degree was significant on perceived time, difficulty, and ease. The effects of familiarity with simple mediation and familiarity with using statistical methods to test mediation were not significant on perceived time, difficulty, and ease.

4.1.2. Contrast: Conditions with no path diagram versus Conditions with a path diagram

There is insufficient evidence to suggest that the presence of a path diagram influenced on cursory interpretational error and reader experience of perceived time, difficulty, ease, satisfaction, and confidence in reading for basic information. Thus, recommendations for presentation formats that will reduce cursory interpretational error and increase positive reader experience cannot be developed when the target audience is reading for the basic information.

4.2. Reading for Detailed Information

In reading for detailed information, participants were asked to interpret numerical values in the stimuli in the context of research. Participants were also asked to identify which statistical methods testing the simple mediation were used in the stimuli. Then, the items measuring reader experience were presented to participants (see Figure 2 on page 26).

4.2.1. Covariates

Familiarity with simple mediation, familiarity with using statistical methods to test mediation, and having a doctorate degree or not did not have any effect on comprehensive interpretational error and reader experience (i.e., perceived time, difficulty, ease, satisfaction, and confidence).

4.2.2. Contrast 1: Text only versus Text and table of coefficients

The text only condition had less comprehensive interpretational error compared to the text and table of coefficient condition, which is the opposite direction to the hypothesis of the present study. Although the table of coefficients had more organized numerical values compared

to the text, how the table of coefficients was organized in this study plausibly influenced participants' interpretational error in comprehensive understanding.

The difference between the text only condition and the text and table of coefficient condition, on the other hand, was not significant for reader experience (i.e., perceived time, difficulty, ease, satisfaction, and confidence). However, the direction of the effects on reader experience followed the pattern of correlations. Recall that the reader experience items formed two subsets where the first subset was composed of perceived time, difficulty, and ease, and the second subset was composed of satisfaction and confidence. The reader experience of the first subset followed what was expected with the presence of a table being associated with more positive reader experience (i.e., perceived time, difficulty, ease) compared to the text only.

Conversely, the second subset showed the opposite patterns to what was expected although the effect was very weak; participants in the text and table of coefficients condition had less satisfaction and confidence after being asked about a comprehensive understanding of the reported results compared to the participants in the text only condition. How the present study organized the table of coefficients in the stimuli plausibly decreased participants' satisfaction in understanding reported results and confidence in understanding reported results.

4.2.3. Contrast 2: Conditions with no path diagram versus Conditions with a path diagram

The presence of a path diagram did not significantly influence comprehensive interpretational error. However, the observed results were in the expected direction where participants in the conditions with a path diagram had less comprehensive interpretational error compared to the conditions with no path diagram.

While there was no effect of a path diagram on perceived time and ease, participants in the conditions with a path diagram had shorter perceived time and higher ease compared to the conditions with no path diagram as expected. On the other hand, there was a significant effect on difficulty. Participants in the conditions with a path diagram had significantly less difficulty in understanding the reported results compared to the participants in the conditions with no path diagram. A more realistic conceptual representation of the path diagram likely assisted participants to perceive less difficulty in the conditions with a path diagram compared to the conditions with no path diagram.

Despite of nonsignificant findings, satisfaction and confidence were higher in the conditions with a path diagram compared to the conditions with no path diagram as expected.

4.2.4. Contrast 3: Text, table of coefficients, and path diagram versus Other conditions

The condition with all three formats did not significantly decreased comprehensive interpretational error relative to the other conditions (i.e., text only, text and table of coefficients, text and path diagram). However, the direction of the effects warrants discussions. Contrary to the expectations, participants in text, table of coefficients, and path diagram condition had higher comprehensive interpretational error compared to the participants in the other conditions. More than two formats presenting information possibly contributed to the increase of interpretational error because there were three different formats to read compared to other conditions with one or two formats.

Similar to the comprehensive interpretational error, there were no significant differences in reader experience between the text, table of coefficients, and path diagram condition and the other conditions. Contrary to the expectations, participants in text, table of coefficients, and path

diagram condition had longer perceived time compared to participants in the other conditions. Participants possibly perceived a longer time taken to obtain the comprehensive understanding because there were three different formats to read compared to other conditions with one or two formats.

The results of the second subset of the reader experience items measuring satisfaction and confidence also had opposite directions to what was expected. Participants in the text, table of coefficients, and path diagram condition had less satisfaction and less confidence compared to the participants in the other conditions when they were reading results for detailed information. Similar to the first subset, participants were possibly less satisfied and less confident in obtaining the comprehensive understanding of the results because there were three different formats to read compared to other conditions with one or two formats.

4.3. Recommendations

Due to the insufficient evidence in the results to support the hypotheses in the present study, it is premature to make definite recommendations for how best to present reported results of a simple mediation analysis which will reduce interpretational error and facilitate positive reader experience. Nonetheless, the direction of the observed nonsignificant effects in the results showed that including a path diagram in reporting the results in a simple mediation model may provide benefits in reducing comprehensive interpretational error and increasing positive reader experience when participants are reading for both basic and detailed information. The path diagram was also 6.5 times and 9.75 times more likely to be selected as most useful presentation format compared to the text only and to the table of coefficients, respectively. Additionally, participants in the conditions with a path diagram significantly perceived less difficulty in

understanding the reported results compared to the participants in the conditions with no path diagram after reading for detailed information. However, future studies are needed to make more definite recommendations for how best to report results of a simple mediation analysis.

4.4. Limitations

Three limitations might account for the nonsignificant results in the present study. First, the small sample size ($N = 71$) in this study likely decreased the statistical power in detecting the difference in the contrasts. Second, a small number of items used to measure both interpretational errors in participants' cursory understanding and reader experience suggests the potential for relatively large measurement error. The Interpretational Error scale only has two items measuring interpretational error in participants' cursory understanding and one item measuring each type of the reader experience (i.e., perceived time, difficulty, ease, satisfaction, and confidence) for each phase of reading for basic versus detailed information. Future studies can increase the sample size and add more items measuring interpretational error in participants' cursory understanding and more items measuring each construct in reader experience so as to reduce measurement error.

Finally, the online format of the present study may have decreased experimental control. Participants took part in the experiment individually, at their convenience, and without any restrictions because the experiment was conducted through the Qualtrics website. This lack of experimental control may reduce the size of the expected effects due to potential variability in how participants provided their information. Future studies can use a different study format in order to increase experimental control.

4.5. Future Directions

The present study examined potential differences among four different presentation formats (i.e., text only, text and table of coefficients, text and path diagram, and text, table of coefficients, and path diagram) on interpretational error and reader experience with two experimental phases of reading (i.e., reading for basic information and reading for detailed information).

In reading for basic information, no significant results were observed. Yet, in reading for detailed information, participants in the text and table of coefficients condition had significantly more interpretational error compared to the participants in the text only condition. This significant result was opposite to what was expected. The present study hypothesized that the text and table of coefficients condition would have less interpretational error compared to the text only condition because the table of coefficients used a more organized way in presenting numerical values compared to the way in presenting numerical values in the text. Additionally, participants' satisfaction and confidence in understanding reported results in the stimuli were lower in the text and table of coefficients condition compared to the text only condition in the observed results. These results might be due to the specific feature of the table of coefficients in the stimuli of the present study. Future studies should examine whether different ways in presenting numerical values in a table of coefficients may affect on interpretational error, satisfaction, and confidence in reading for detailed information.

Future studies should also be designed to use equivalence tests to examine whether the effects of different presentation formats are equivalent on interpretational error and reader experience. Nonsignificant differences in the contrasts do not imply that the effects of different

presentation formats on interpretational error and reader experience are equivalent. Thus, using equivalence test will examine whether differences in four different conditions of presentation formats on interpretational error and reader experience are practically nonsignificant and considered equivalence.

References

- Agus, M., Peró-Cebollero, M., Penna, M. P., & Guàrdia-Olmos, J. (2015). Towards the development of problems comparing verbal-numerical and graphical formats in statistical reasoning. *Quality & Quantity: International Journal of Methodology*, *49*, 691-709. doi: 10.1007/s11135-014-0018-7
- American Psychological Association. (2011). *Publication manual of the American Psychological Association*. (6th ed.) Washington, DC: American Psychological Association.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173. doi: 10.1037/0022-3514.51.6.1173
- Biesanz, J. C., Falk, C. F., & Savalei, V. (2010). Assessing mediational models: Testing and interval estimation for indirect effects. *Multivariate Behavioral Research*, *45*, 661-701. doi: 10.1080/00273171.2010.498292
- Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, *20*, 15-140.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*, 1-73. doi:10.1016/0010-0277(95)00664-8

El-Sheikh, M., & Elmore–Staton, L. O. R. I. (2004). The link between marital conflict and child adjustment: Parent–child conflict and perceived attachments as mediators, potentiators, and mitigators of risk. *Development and Psychopathology, 16*, 631-648.

doi:10.1017/S0954579404004705

Gaissmaier, W., Wegwarth, O., Skopec, D., Müller, A., Broschinski, S., & Politi, M. C. (2012).

Numbers can be worth a thousand pictures: Individual differences in understanding graphical and numerical representations of health-related information. *Health Psychology, 31*, 286-296.

doi:10.1037/a0024850

Garcia, D., Kerekes, N., & Archer, T. (2012). A will and a proper way leading to happiness:

Self-directedness mediates the effect of persistence on positive affectivity. *Personality and Individual Differences, 53*, 1034-1038. doi:10.1016/j.paid.2012.07.025

Gelfand, L. A., Mensinger, J. L., & Tenhave, T. (2009). Mediation analysis: A retrospective

snapshot of practice and more recent directions. *Journal of General Psychology, 136*, 153-

176. doi:10.3200/GENP.136.2.153-178

Gelsema, T. I., van der Doef, M., Maes, S., Akerboom, S. & Verhoeven, C. (2005). Job stress in

the nursing profession: The influence of organizational and environmental conditions and job characteristics. *International Journal of Stress Management, 12*, 222-240. doi:

10.1037/1072-5245.12.3.222

Kamphoff, C. S., Gill, D. L., & Huddleston, S. (2005). Jealousy in sport: Exploring jealousy's relationship to cohesion. *Journal of Applied Sport Psychology, 17*, 290-305. doi:

10.1080/10413200500313578

- Kashy, D. A., Donnellan, M. B., Ackerman, R. A., & Russell, D. W. (2009). Reporting and interpreting research in PSPB: Practices, principles, and pragmatics. *Personality and Social Psychology Bulletin*, *35*, 1131-1142. doi: 10.1177/0146167208331253
- MacKinnon, D. P., Lockwood, C. M., & Hoffman, J. (1998, June). A new method to test for mediation. In *annual meeting of the Society for Prevention Research, Park City, UT* (pp. 268-274).
- 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*, 83-104. doi:10.1037/1082-989X.7.1.83
- MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593-614. doi:10.1146/annurev.psych.58.110405.085542
- Piscotty, R. J., Jr., Kalisch, B., & Gracey-Thomas, A. (2015). Impact of healthcare information technology on nursing practice. *Journal of Nursing Scholarship*, *47*, 287-293.
doi:10.1111/jnu.12138
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments & Computers*, *36*, 717-731. doi: 10.3758/BF03206553
- Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology*, *66*, 825-852. doi: 10.1146/annurev-psych-010814-015258

- Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, 5, 359-371. doi:10.1111/j.1751-9004.2011.00355.x
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7, 422-445. doi: 10.1037/1082-989X.7.4.422
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13, 290-312.
- Sobel, M. E. (1986). Some new results on indirect effects and their standard errors in covariance structure models. *Sociological methodology*, 16, 159-186. Retrieved from <http://smx.sagepub.com/>
- Stoddard, S. A., Varela, J. J., & Zimmerman, M. A. (2015). Future expectations, attitude toward violence, and bullying perpetration during early adolescence: A mediation evaluation. *Nursing Research*, 64, 422-433. doi:10.1097/NNR.0000000000000127
- Tasir, Z., El, A. A., Abd Halim, N. D., & Harun, J. (2012). Relationship between teachers' ICT competency, confidence level, and satisfaction toward ICT training programmes: A case study among postgraduate students. *TOJET: The Turkish Online Journal of Educational Technology*, 11(1), 138-144. Retrieved from <http://eric.ed.gov/?id=EJ976576>

- Tormala, Z. L., Clarkson, J. J., & Henderson, M. D. (2011). Does fast or slow evaluation foster greater certainty? *Personality and Social Psychology Bulletin*, *37*, 422-434. doi: 10.1177/0146167210397378
- Tukey, J. W. (1990). Data-based graphics: Visual display in the decades to come. *Statistical Science*, *5*, 327-339. doi: 10.1214/ss/1177012101
- Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594-604. doi:10.1037/0003-066X.54.8.594
- Zhang, J., Jemmott, John B., I., II, & Jemmott, L. S. (2015). Mediation and moderation of an efficacious theory-based abstinence-only intervention for African American adolescents. *Health Psychology*, *34*, 1175-1184. doi:10.1037/hea0000244

Appendix A: Interpretational Error Scale

In this section, you will read an excerpt that highlights the results of a psychological journal article that makes use of mediation analysis. The questions that follow are about the excerpt below and concern your interpretation of the results reported. Please copy and paste <http://tinyurl.com/jkexce-d> in the address box in a new internet browser window.

**The "x" indicates the correct response.*

| Please select True or False. | True | False |
|---|------|-------|
| 1. The effect of Self-Directedness on positive affect is mediated by Persistence. | | x |
| 2. Self-Directedness significantly mediated the relationship between Persistence and Positive Affect. | x | |

| | SD | D | NDA | A | SA |
|--|-------------------|----------|----------------------------|-------|----------------|
| The following statements are about the impact of how the results were presented. | Strongly Disagree | Disagree | Neither Disagree nor Agree | Agree | Strongly Agree |
| 3. How the results were presented made it difficult to understand them. | | | | | |
| 4. How the results were presented took more time than usual to understand them. | | | | | |
| 5. How the results were presented made it easy to understand them. | | | | | |

| | SD | D | NDA | A | SA |
|---|-------------------|----------|----------------------------|----------|----------------|
| The following statements are about your understanding of the reported results | Strongly Disagree | Disagree | Neither Disagree nor Agree | Agree | Strongly Agree |
| 6. I am satisfied with my understanding of these results. | | | | | |
| 7. I am confident in my understanding of these results. | | | | | |

| | SD | D | NDA | A | SA |
|---|-------------------|----------|----------------------------|----------|----------------|
| The following statements are about your overall perceived familiarity with mediation. | Strongly Disagree | Disagree | Neither Disagree nor Agree | Agree | Strongly Agree |
| 8. I am satisfied with my knowledge of mediation. | | | | | |
| 9. I am confident in with my knowledge of mediation. | | | | | |

| | |
|---|------|
| 10. What is the estimated total effect of Persistence (X) on Positive Affect (Y)? | |
| | 0.23 |
| | 0.36 |
| | 0.26 |
| x | 0.34 |
| | 0.08 |

| | |
|---|------|
| 11. What is the estimated effect of Persistence (X) on Self-Directedness (M)? | |
| x | 0.23 |
| | 0.36 |
| | 0.26 |
| | 0.34 |
| | 0.08 |

| | |
|--|------|
| 12. What is the estimated direct effect of Persistence (X) on Positive Affect (Y)? | |
| | 0.23 |
| | 0.36 |
| x | 0.26 |
| | 0.34 |
| | 0.08 |

| | |
|--|------|
| 13. What is the estimated effect of Self-Directedness (M) on Positive Affect (Y), controlling for the effect of Persistence (X)? | |
| | 0.23 |
| x | 0.36 |
| | 0.26 |
| | 0.34 |
| | 0.08 |

| | |
|--|------------------------|
| 14. What is the estimated indirect effect of Persistence (X) on Positive Affect (Y) through Self-Directedness (M)? Please select all that apply. | |
| x | 0.23×0.36 |
| | 0.26×0.36 |
| x | 0.08 |
| | $0.34 - (0.23 + 0.36)$ |
| x | $0.34 - 0.26$ |

| | |
|--|---|
| 15. Please read five statements below. Please identify the statement with a different concept. | |
| | A product of the effect of Persistence (X) on Self-Directedness (M) by the effect of Self-directedness (M) on Positive Affect (Y), controlling for Persistence (X). |
| | The indirect effect of Persistence (X) on Positive Affect (Y) through Self-Directedness (M). |
| x | The effect of Persistence (X) on Positive Affect (Y) given the presence of Self-Directedness (M). |
| | The difference between the total and direct effects of Persistence (X) on Positive Affect (Y). |
| | The effect of Persistence (X) on Positive Affect(Y) through Self-Directedness (M). |

| | |
|--|---|
| 16. From the results presented previously, what does $(.23 \times .36 = .08)$ represent? Please select all that apply. | |
| x | The effect of Persistence (X) on Positive Affect (Y) through Self-Directedness (M). |
| x | The indirect effect of Persistence (X) on Positive Affect (Y). |
| | The expected change in Positive Affect (Y) due to a one unit change in Persistence (X), controlling for Self-Directedness (M). |
| x | The estimated difference between the total and direct effect of Persistence (X) on Positive Affect (Y) controlling for Self-Directedness (M). |
| x | The product of the effect of Persistence (X) on Self-Directedness (M), and the effect of Self-Directedness (M) on Positive Affect (Y), controlling for Persistence (X). |

| | |
|--|---|
| 17. What does Sobel test evaluate? Please select all that apply. | |
| x | The significance of the product of the effect of Persistence (X) on Self-Directedness (M) by the effect of Self-Directedness (M) on Positive Affect (Y), controlling for Persistence (X). |
| x | The significance of the difference between the total effect of Persistence (X) on Positive Affect (Y) and the direct effect of Persistence (X) on Positive Affect (Y). |
| x | The significance of the joint effect of Persistence (X) on Self-Directedness (M), and the effect of Self-Directedness (M) on Positive Affect (Y), controlling for Persistence (X). |
| | The significance of the effect of Persistence (X) on Positive Affect (Y), controlling for Self-Directedness (M). |
| x | The significance of the indirect effect of Persistence (X) on Positive Affect(Y). |

| | |
|---|--|
| 18. Which of the following statements give a correct interpretation of the percentile bootstrap confidence interval (CI) presented in the text? Please select all that apply. | |
| x | We are 95% confident that the effect of Persistence (X) on Positive Affect (Y) through Self-Directedness (M) in the population lies between (.02, .13). |
| x | The mediated effect of Persistence (X) on Positive Affect (Y), through Self-Directedness (M), is significantly different from zero. |
| x | We are 95% confident that the difference between the total and direct effects of Persistence (X) on Positive Affect (Y) in the population lies between (.02, .13). |
| | The direct effect of Persistence (X) on Positive Affect(Y) in the population is significantly different from zero. |
| x | We are 95% confident that the product of the effect of Persistence (X) on Self-Directedness (M) by the effect of Self-Directedness (M) on Positive Affect (Y) in the population lies between (.02, .13). |

| | |
|---|--|
| 19. Which of the following statements give a correct interpretation of the results? Please select all that apply. | |
| x | Self-Directedness (M) mediates the effect of Persistence (X) on Positive Affect (Y). |
| x | A unit increase in Persistence (X) is associated with an increase of .23 units in Self-Directedness (M), which in turn influences Positive Affect (Y) such that a unit increase in Self-Directedness (M) is associated with an increase of .36 in Positive Affect (Y). |
| x | Positive Affect (Y) is expected to increase by .08 indirectly through Self-Directedness (M) per unit change in Persistence (X). |
| x | The difference between the total and direct effect of Persistence (X) on Positive Affect (Y) is statistically significant. |
| x | The effect of Persistence (X) on Positive Affect (Y) decreases by .08 when Self-Directedness (M) is added into the model. |

| | | | |
|--|-------------------------------|--|--|
| 20. Which of the following statistical methods of testing mediation had estimates reported in the text? Please select all that apply. | | | |
| x | Causal steps | | Not use of "Causal steps" |
| x | Sobel test | | Not use of "Sobel test" |
| | Difference in coefficients | | Not use of "Difference in coefficients" |
| x | Bootstrap confidence interval | | Not use of "Bootstrap confidence interval" |
| x | Product of coefficients | | Not use of "Product of coefficients" |

| | |
|---|-----------------------|
| 21. Which device is the most useful for understanding reported results of mediation analysis? | |
| x | Text |
| x | Table of coefficients |
| x | Path diagram |

Appendix B: Attitudes Toward Mediation Scale

This section is about one's familiarity with the framework of mediation investigating underlying mechanism, which encompasses both theoretical concepts of and statistical approaches for testing mediation. Theoretical concepts of mediation relate to the psychological theories about underlying processes or causal mechanisms. Statistical approaches for testing mediation relate to the statistical models and procedures that are applied to data to evaluate mediation.

| Please select one choice for each statement. | Yes | No |
|--|------------|-----------|
| 1. Are you familiar with mediation? | | |
| 2. Are you aware that there is a distinction between the theoretical concept of mediation and statistical methods for testing mediation? | | |
| 3. Are you aware of the concepts behind testing mediation when using a statistical method to evaluate mediation? | | |
| 4. Have you used any statistical method to test mediation? | | |

| 5. In your research, who would mainly perform mediation analysis? | |
|---|---|
| | Myself |
| | My graduate student(s) / lab member(s) who is/are (a) graduate student(s) |
| | My collaborator(s) |
| | Someone I consult for statistics |

| | SD | D | NDA | A | SA |
|---|-------------------|----------|----------------------------|----------|----------------|
| Please select one choice for each statement. | Strongly Disagree | Disagree | Neither Disagree nor Agree | Agree | Strongly Agree |
| 6. It is difficult to understand the theoretical concept of mediation. | | | | | |
| 7. I can explain the theoretical concept of mediation to my colleagues. | | | | | |
| 8. It is difficult to understand the various statistical methods for testing mediation. | | | | | |
| 9. I can explain how to statistically test mediation to my colleagues. | | | | | |
| 10. Mediation analysis is <i>popular</i> in my area of research. | | | | | |
| 11. Mediation analysis is <i>useful</i> in my area of research. | | | | | |
| 12. Mediation analysis is <i>underused</i> in my area of research. | | | | | |
| 13. Mediation analysis is <i>overused</i> in my area of research. | | | | | |
| 14. Influential papers in my area of research use mediation. | | | | | |

This section is about your training experience in the framework of mediation, which encompasses both theoretical concepts of and statistical approaches for testing mediation. Theoretical concepts of mediation relate to the psychological theories about underlying processes or causal mechanisms. Statistical approaches to mediation relate to the statistical models and procedures that are applied to data to evaluate mediation.

| I learned about theoretical concepts of mediation by... | Yes | No | NA |
|---|-----|----|----|
| 15 (a). attending talks at conferences where the primary topic was about mediation. | | | |
| 15 (b). attending talks at conferences which touched on mediation as a side topic . | | | |
| 16 (a). attending workshops or short courses where the primary topic was about mediation. | | | |
| 16 (b). attending workshops or short courses which touched on mediation as a side topic . | | | |
| 17 (a). taking or auditing courses for credit at my institution where the primary topic was about mediation. | | | |
| 17 (b). taking or auditing courses for credit at my institution which touched on mediation as a side topic . | | | |
| 18 (a). without taking courses, reading textbooks where the primary topic was about mediation. | | | |
| 18 (b). without taking courses, reading textbooks which touched on mediation as a side topic . | | | |
| 19 (a). reading methodological journal articles where the primary topic was about mediation. | | | |
| 19 (b). reading methodological journal articles which touched on mediation as a side topic . | | | |

| I learned about theoretical concepts of mediation by... | Yes | No | NA |
|--|-----|----|----|
| 20. reading journal articles in my area of research. | | | |
| 21. visiting websites on mediation (e.g. personal blogs). | | | |
| 22. using social media (e.g. Facebook group in mediation). | | | |
| 23. interacting with colleagues or other lab members. | | | |

| I learned about statistical methods for testing mediation by... | Yes | No | NA |
|---|-----|----|----|
| 24 (a). attending talks at conferences where the primary topic was about mediation. | | | |
| 24 (b). attending talks at conferences which touched on mediation as a side topic . | | | |
| 25 (a). attending workshops or short courses where the primary topic was about mediation. | | | |
| 25 (b). attending workshops or short courses which touched on mediation as a side topic . | | | |
| 26 (a). taking or auditing courses for credit at my institution where the primary topic was about mediation. | | | |
| 26 (b). taking or auditing courses for credit at my institution which touched on mediation as a side topic . | | | |
| 27 (a). without taking courses, reading textbooks where the primary topic was about mediation. | | | |
| 27 (b). without taking courses, reading textbooks which touched on mediation as a side topic . | | | |
| 28 (a). reading methodological journal articles where the primary topic was about mediation. | | | |
| 28 (b). reading methodological journal articles which touched on mediation as a side topic . | | | |

| I learned about statistical methods for testing mediation by... | Yes | No | NA |
|--|-----|----|----|
| 29. reading journal articles in my area of research. | | | |
| 30. visiting websites on mediation (e.g. personal blogs). | | | |
| 31. using social media (e.g. Facebook group in mediation). | | | |
| 32. interacting with colleagues or other lab members. | | | |

| | SD | D | NDA | A | SA |
|--|-------------------|----------|----------------------------|-------|----------------|
| Theoretical mediation should be taught... | Strongly Disagree | Disagree | Neither Disagree nor Agree | Agree | Strongly Agree |
| 33. at the undergraduate level. | | | | | |
| 34. at the graduate level. | | | | | |
| 35. in a research methods course. | | | | | |
| 36. in a statistics course. | | | | | |
| 37. in a short course or workshop. | | | | | |

| | SD | D | NDA | A | SA |
|--|-------------------|----------|----------------------------|-------|----------------|
| Statistical methods for testing mediation should be taught... | Strongly Disagree | Disagree | Neither Disagree nor Agree | Agree | Strongly Agree |
| 38. at the undergraduate level. | | | | | |
| 39. at the graduate level. | | | | | |
| 40. in a research methods course. | | | | | |
| 41. in a statistics course. | | | | | |
| 42. in a short course or workshop. | | | | | |

| | SD | D | NDA | A | SA |
|---|-------------------|----------|----------------------------|----------|----------------|
| The following statements are about your overall perceived familiarity with mediation. | Strongly Disagree | Disagree | Neither Disagree nor Agree | Agree | Strongly Agree |
| 43. I am satisfied with my knowledge of mediation. | | | | | |
| 44. I am confident in my knowledge of mediation. | | | | | |

Appendix C: Demographic Questions

| |
|--|
| The followings are about demographics questions. |
|--|

| | |
|---------------------|--|
| 1. Are you a _____? | |
|---------------------|--|

| | |
|--------------------------|----------------------|
| <input type="checkbox"/> | Graduate student |
| <input type="checkbox"/> | Postdoctoral fellow |
| <input type="checkbox"/> | Tenure-track faculty |
| <input type="checkbox"/> | Contract faculty |

| | |
|-----------------------------|--|
| 2. What program are you in? | |
|-----------------------------|--|

| | |
|--------------------------|---------|
| <input type="checkbox"/> | Masters |
| <input type="checkbox"/> | PhD |

| | | | | | | | |
|------------------------------------|--|--|--|--|--|--|--|
| 3. Which year of study are you in? | | | | | | | |
|------------------------------------|--|--|--|--|--|--|--|

| | | | | | | | |
|--------------------------|---|--------------------------|---|--------------------------|------------|--------------------------|---|
| <input type="checkbox"/> | 1 | <input type="checkbox"/> | 2 | <input type="checkbox"/> | 3 | <input type="checkbox"/> | 4 |
| <input type="checkbox"/> | 5 | <input type="checkbox"/> | 6 | <input type="checkbox"/> | 7 and more | | |

| | |
|--|--|
| 4. What is your highest achieved degree? | |
|--|--|

| | |
|--------------------------|-------------------------------|
| <input type="checkbox"/> | Bachelor of Science |
| <input type="checkbox"/> | Bachelor of Arts |
| <input type="checkbox"/> | Master of Science |
| <input type="checkbox"/> | Master of Arts |
| <input type="checkbox"/> | Doctorate |
| <input type="checkbox"/> | Other (Please specify) _____. |

| | |
|---|----------------------|
| 5 (a). Which doctoral degree did you receive? | |
| | Doctor of Philosophy |
| | Doctor of Psychology |

| | |
|--|--|
| 5 (b). Which year did you receive your doctoral degrees? | |
| Please write the year in numbers (e.g., 1990) _____. | |

| | |
|---|--------------------------------|
| 6. Which activity do you spend most of your time? | |
| | Teaching |
| | Conducting research |
| | Providing therapy |
| | Other (Please describe) _____. |

| | |
|--|--------------------------------|
| 7. What is your primary research area? | |
| | Clinical |
| | Clinical Developmental |
| | Cognitive / Neuroscience |
| | Developmental |
| | History & Theory |
| | Quantitative Methods |
| | Social / Personality |
| | Other (Please describe) _____. |

| | |
|-----------------------------|------------|
| 8. What is your gender? | |
| | Male |
| | Female |
| | Undeclared |
| 9. What is your age: _____. | |

Appendix D: Experimental Stimuli

Text only condition

Background

Persistence is an individual difference that, when applied to the pursuit of need-satisfying goals, leads to increased subjective well-being as operationalized by positive affect. Self-directedness, defined as the ability to develop good habits and behave in accordance with long-term values and goals, is examined in relation to persistence and positive affect. Specifically, self-directedness was examined as a mediator (M) in the relationship between persistence (X) and positive affect (Y). Cross-sectional data were collected from 304 high school students who filled out a questionnaire containing measures of Persistence, Self-Directedness and Positive Affect.

Preliminary Data Analysis

To test for mediation, a series of three regression equations were fit to the data. In the first model, Positive Affect (Y) was regressed onto Persistence (X), and Persistence accounted for a significant amount of variance in Positive Affect, $R^2 = .12$, $p < .001$. The total effect of Persistence (X) on Positive Affect (Y) was significant (standardized coefficient $b^* = .34$, $t(302) = 6.70$, $p < .001$). In the second model, Self-Directedness (M) was regressed onto Persistence (X), and Persistence contributed a significant amount of variance to Self-Directedness, $R^2 = .05$, $p < .001$. The effect of Persistence (X) on Self-Directedness (M) was significant ($b^* = .23$, $t(302) = 4.23$, $p < .001$). In the third and final model, Positive Affect (Y) was simultaneously regressed onto both Persistence (X) and Self-Directedness (M), and the model accounted for 23.9% of the variance in Positive Affect, $p < .001$. The direct effect of Persistence (X) on Positive Affect (Y), controlling for the effect of Self-Directedness (M) on Positive Affect (Y), was significant ($b^* = .26$, $t(301) = 5.14$, $p < .001$). Additionally, the effect of Self-Directedness (M) on Positive Affect (Y), controlling for the effect of Persistence (X) on Positive Affect (Y), was significant ($b^* = .36$, $t(301) = 7.45$, $p < .001$).

Mediation Analysis

Mediation is present when the effect of Persistence (X) on Positive Affect (Y) occurs through the mediator, Self-Directedness (M). When Self-Directedness (M) was included in the model, the total effect of Persistence (X) on Positive Affect (Y) ($b^* = .34$) was reduced to the direct effect of $b^* = .26$, which remained significant. The Sobel test for mediation was significant, $z = 3.55$, $p < .001$, and the indirect effect ($ab = .08$) had a 95% percentile bootstrap CI = [.02, .13].

Text and table of coefficients condition

Background

Persistence is an individual difference that, when applied to the pursuit of need-satisfying goals, leads to increased subjective well-being as operationalized by positive affect. Self-directedness, defined as the ability to develop good habits and behave in accordance with long-term values and goals, is examined in relation to persistence and positive affect. Specifically, self-directedness was examined as a mediator (M) in the relationship between persistence (X) and positive affect (Y). Cross-sectional data were collected from 304 high school students who filled out a questionnaire containing measures of Persistence, Self-Directedness and Positive Affect.

Preliminary Data Analysis

To test for mediation, a series of three regression equations were fit to the data. All estimates from these three regression equations are presented in Table 1. In the first model, Positive Affect (Y) was regressed onto Persistence (X), and Persistence accounted for a significant amount of variance in Positive Affect. The total effect of Persistence (X) on Positive Affect (Y) was significant. In the second model, Self-Directedness (M) was regressed onto Persistence (X), and Persistence contributed a significant amount of variance to Self-Directedness. The effect of Persistence (X) on Self-Directedness (M) was significant. In the third and final model, Positive Affect (Y) was simultaneously regressed onto both Persistence (X) and Self-Directedness (M), and the model accounted significantly for the variance in Positive Affect. The direct effect of Persistence (X) on Positive Affect (Y), controlling for the effect of Self-Directedness (M) on Positive Affect (Y), was significant. Additionally, the effect of Self-Directedness (M) on Positive Affect (Y), controlling for the effect of Persistence (X) on Positive Affect (Y), was significant.

Mediation Analysis

Mediation is present when the effect of Persistence (X) on Positive Affect (Y) occurs through the mediator, Self-Directedness (M). When Self-Directedness (M) was included in the model, the total effect of Persistence (X) on Positive Affect (Y) ($b^* = .34$) was reduced to the direct effect of $b^* = .26$, which remained significant. The Sobel test for mediation was significant, $z = 3.55$, $p < .001$, and the indirect effect ($ab = .08$) had a 95% percentile bootstrap CI = [.02, .13].

Table 1.

Estimates for the effect of Persistence and Self-Directedness on Positive Affect

| Model | b^* | t | df | R^2 |
|--|-------|------|------|-------|
| Positive Affect (Y) Persistence (X) | 0.34 | 6.70 | 302 | 0.12 |
| Self-Directedness (M) Persistence (X) | 0.23 | 4.23 | 302 | 0.05 |
| Positive Affect (Y) Persistence (X) | 0.26 | 5.14 | 301 | 0.24 |
| Self-Directedness (M) | 0.36 | 7.45 | 301 | |

Note. b^* = standardized regression coefficients and R^2 = squared multiple correlation for the model. All estimates were significant at $p < .001$.

Text and path diagram condition

Background

Persistence is an individual difference that, when applied to the pursuit of need-satisfying goals, leads to increased subjective well-being as operationalized by positive affect. Self-directedness, defined as the ability to develop good habits and behave in accordance with long-term values and goals, is examined in relation to persistence and positive affect. Specifically, self-directedness was examined as a mediator (M) in the relationship between persistence (X) and positive affect (Y). Cross-sectional data were collected from 304 high school students who filled out a questionnaire containing measures of Persistence, Self-Directedness and Positive Affect.

Preliminary Data Analysis

To test for mediation, a series of three regression equations were fit to the data. Path diagrams that represent these regression equations are presented in Figure 1. In the first model, Positive Affect (Y) was regressed onto Persistence (X), and Persistence accounted for a significant amount of variance in Positive Affect, $R^2 = .12, p < .001$. The total effect of Persistence (X) on Positive Affect (Y) was significant, $t(302) = 6.70, p < .001$. In the second model, Self-Directedness (M) was regressed onto Persistence (X), and Persistence contributed a significant amount of variance to Self-Directedness, $R^2 = .05, p < .001$. The effect of Persistence (X) on Self-Directedness (M) was significant, $t(302) = 4.23, p < .001$. In the third and final model, Positive Affect (Y) was simultaneously regressed onto both Persistence (X) and Self-Directedness (M), and the model accounted for 23.9% of the variance in Positive Affect, $p < .001$. The direct effect of Persistence (X) on Positive Affect (Y), controlling for the effect of Self-Directedness (M) on Positive Affect (Y), was significant, $t(301) = 5.14, p < .001$. Additionally, the effect of Self-Directedness (M) on Positive Affect (Y), controlling for the effect of Persistence (X) on Positive Affect (Y), was significant, $t(301) = 7.45, p < .001$.

Mediation Analysis

Mediation is present when the effect of Persistence (X) on Positive Affect (Y) occurs through the mediator, Self-Directedness (M). When Self-Directedness (M) was included in the model, the total effect of Persistence (X) on Positive Affect (Y) ($b^* = .34$) was reduced to the direct effect of $b^* = .26$, which remained significant. The Sobel test for mediation was significant, $z = 3.55, p < .001$, and the indirect effect ($ab = .08$) had a 95% percentile bootstrap CI = [.02, .13].

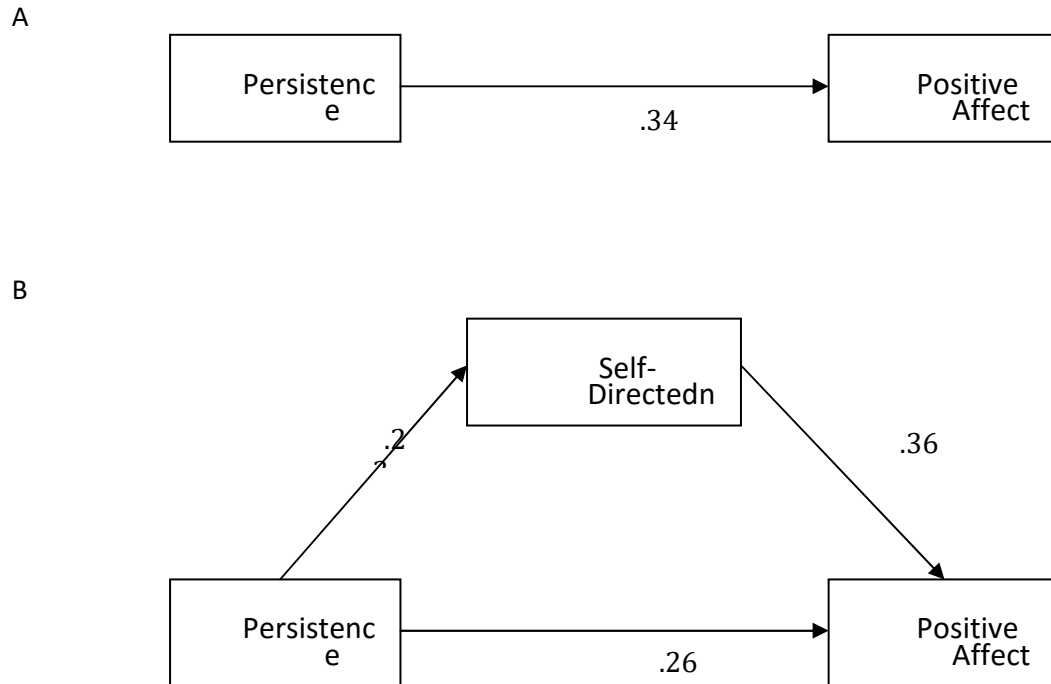


Figure 1. Path diagrams depicting models in simple mediation. Panel A depicts the direct effect of Persistence on Positive Affect. Panel B depicts the effect of Persistence on Positive Affect through Self-Directedness. Estimates are standardized regression coefficients which are significant at $p < .001$.

Text, table of coefficients, and path diagram condition

Background

Persistence is an individual difference that, when applied to the pursuit of need-satisfying goals, leads to increased subjective well-being as operationalized by positive affect (See Panel A of Figure 1). Self-directedness, defined as the ability to develop good habits and behave in accordance with long-term values and goals, is examined in relation to persistence and positive affect. Specifically, self-directedness was examined as a mediator (M) in the relationship between persistence (X) and positive affect (Y) (See Panel B of Figure 1). Cross-sectional data were collected from 304 high school students who filled out a questionnaire containing measures of Persistence, Self-Directedness and Positive Affect.

Preliminary Data Analysis

To test for mediation, a series of three regression equations were fit to the data. All estimates from these three regression equations are presented in Table 1. In the first model, Positive Affect (Y) was regressed onto Persistence (X), and Persistence accounted for a significant amount of variance in Positive Affect. The total effect of Persistence (X) on Positive Affect (Y) was significant. In the second model, Self-Directedness (M) was regressed onto Persistence (X), and Persistence contributed a significant amount of variance to Self-Directedness. The effect of Persistence (X) on Self-Directedness (M) was significant. In the third and final model, Positive Affect (Y) was simultaneously regressed onto both Persistence (X) and Self-Directedness (M), and the model accounted significantly for the variance in Positive Affect. The direct effect of Persistence (X) on Positive Affect (Y), controlling for the effect of Self-Directedness (M) on Positive Affect (Y), was significant. Additionally, the effect of Self-Directedness (M) on Positive Affect (Y), controlling for the effect of Persistence (X) on Positive Affect (Y), was significant.

Mediation Analysis

Mediation is present when the effect of Persistence (X) on Positive Affect (Y) occurs through the mediator, Self-Directedness (M). When Self-Directedness (M) was included in the model, the total effect of Persistence (X) on Positive Affect (Y) ($b^* = .34$) was reduced to the direct effect of $b^* = .26$, which remained significant. The Sobel test for mediation was significant, $z = 3.55$, $p < .001$, and the indirect effect ($ab = .08$) had a 95% percentile bootstrap CI = [.02, .13].

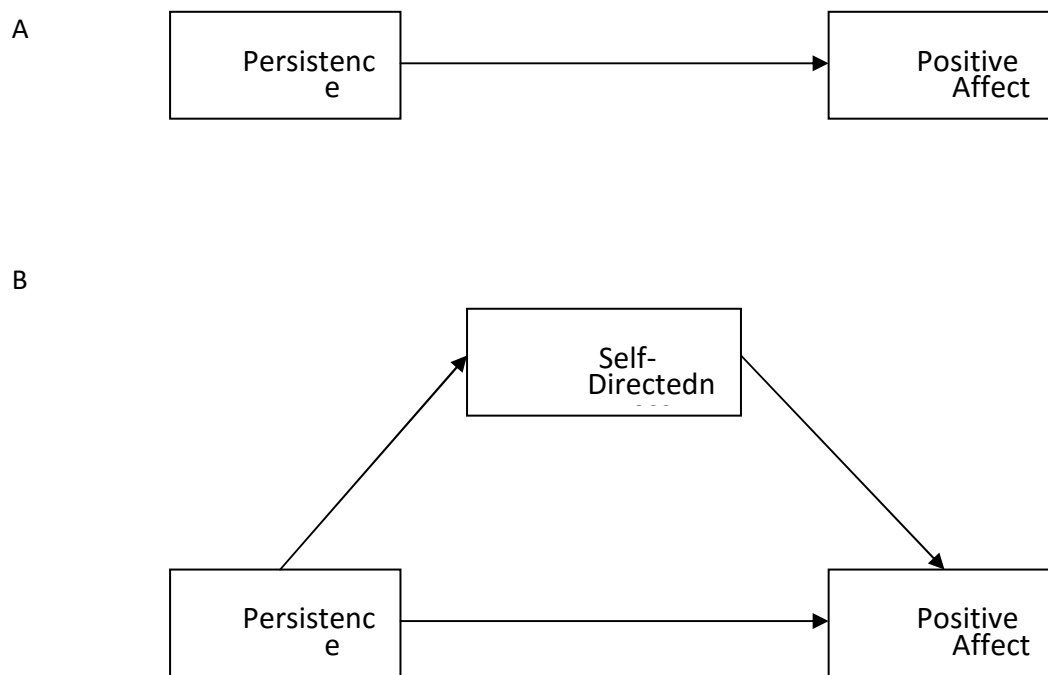


Figure 1. Diagrams depicting models in simple mediation. Panel A depicts the direct effect of Persistence on Positive Affect. Panel B depicts the effect of Persistence on Positive Affect through Self-Directedness.

Table 1.

Estimates for the effect of Persistence and Self-Directedness on Positive Affect

| Model | b^* | t | df | R^2 |
|---|--------------|--------------|------------|-------|
| Positive Affect (Y) Persistence (X) | 0.34 | 6.70 | 302 | 0.12 |
| Self-Directedness (M) Persistence (X) | 0.23 | 4.23 | 302 | 0.05 |
| Positive Affect (Y) Persistence (X) Self-Directedness (M) | 0.26 0.36 | 5.14 7.45 | 301 301 | 0.24 |

Note. b^* = standardized regression coefficients and R^2 = squared multiple correlation for the model. All estimates were significant at $p < .001$.

Appendix E: Informed Consent Form

Study Name

Evidence-based recommendations of reporting results from mediation analysis: A focus on ease of interpretation and maximum accuracy

Researchers:

Y. Jamie Kim, B.A. Honours; jkim793@yorku.ca

011 Behavioural Science Building, Department of Psychology, York University
4700 Keele Street, Toronto, Ontario, Canada M3J1P4

Pek, Jolynn, PhD; pek@yorku.ca

322 Behavioural Science Building, Department of Psychology, York University
4700 Keele Street, Toronto, Ontario, Canada M3J1P4

Sponsors: York University, Toronto, Ontario, Canada

Purpose of the Research

The experiment examines how different forms of presenting results in mediation analyses affects the extent of interpretational errors, readers' ease of interpretation, and readers' understanding in reported results.

What You Will Be Asked to Do in the Research

Please answer the questions which are mainly about statistical mediation. The experiment will take approximately 20 -30 minutes to complete.

Risks and Discomforts

You may feel mildly anxious when you answer questions about results of the mediation analysis.

Benefits of the Research and Benefits to You

There is no direct benefit to you as a participant. However, researchers using mediation would benefit in the future by making use of recommendations for result reporting derived from this study.

Voluntary Participation

Your participation in this study is completely voluntary and you may choose to stop participating at any time. Your decision not to volunteer will not influence the nature of the ongoing relationship you may have with the researchers or the nature of your relationship with York University.

Withdrawal from the Study

You can stop participating in this study at any time, for any reason, if you so decide. Your decision to stop participating, or to refuse to answer particular questions, will not affect your relationship with the researchers, York University, or any other group associated with this project. In the event you withdraw from the study, all associated data collected will be immediately destroyed wherever possible.

Confidentiality

All information you supply during the research will be held in confidence, and unless you specifically indicate your consent, your name will not appear in any report or publication of the research. If you agree to waive your anonymity and have your name included in any final reports or publications of the research, please put check mark below. Your data will be collected through Qualtrics, an online survey platform. Your data will be safely stored as an encrypted file in a password protected computer. Only research staff will have access to this information. The data will be kept for five years and will be deleted permanently after the retention period ends. Confidentiality will be provided to the fullest extent possible by law.

I as a participant agree to waive my anonymity and have my name included in any final reports / publications of the research.

Questions about the Research

If you have questions about the research in general, or about your role in the study, please feel free to contact the primary investigator, Y. Jamie Kim either by telephone at (416) 736-2100, extension 40264 or by email (jkim793@yorku.ca). This research has been reviewed by the Human Participants Review Subcommittee, York University's Ethics Review Board, and conforms to the standards of the Canadian Tri-Council Research Ethics guideline. If you have any questions about this process or about your rights as a participant in the study, please contact the Sr. Manager & Policy Advisor for the Office of Research Ethics, 5th Floor, York Research Tower, York University (telephone 416-736-5914 or email ore@yorku.ca).

I consent to participate in “Evidence-based recommendations of reporting results from mediation analysis: A focus on ease of interpretation and maximum accuracy” conducted by Y. Jamie Kim. I have understood the nature of this study and wish to participate. You are not waiving any of your legal rights by selecting “I agree”. Your choice of “I agree” indicates your consent.

I agree.

I disagree.

Appendix F: Debriefing

Evidence-based recommendations of reporting results from mediation analysis:

A focus on ease of interpretation and maximum accuracy

The aim of this study is to obtain evidence-based standards for reporting results of simple mediation so as to reduce errors in interpretation and maximize accuracy.

In the experiment, participants were randomly assigned into one of four conditions: text only, text and table of coefficients, text and path diagram, and text, table of coefficients, and path diagram. In each condition, participants were asked to answer questions pertinent to the presentation types of each condition.

We expect the difference in the number of interpretational errors and readers' understanding of results among four conditions. Study results would inform on best practices for reporting results of simple mediation.

Thank you for your participation! If you have any concerns about anxiety induced from interpreting and answering the questions about the results of mediation analysis, or have further questions about the study, please feel free to contact the primary investigator, Y. Jamie Kim at jkim793@yorku.ca. Additionally, if you have any concerns about any aspect of the study, you may contact the Sr. Manager & Policy Advisor for the Office of Research Ethics, 5th Floor, York Research Tower, York University (telephone 416-736-5914 or email ore@yorku.ca).

Additional Readings:

Gelfand, L. A., Mensinger, J. L., & Tenhave, T. (2009). Mediation analysis: A retrospective snapshot of practice and more recent directions. *Journal of General Psychology*, *136*, 153-176. doi:10.3200/GENP.136.2.153-178

Kashy, D. A., Donnellan, M. B., Ackerman, R. A., & Russell, D. W. (2009). Reporting and interpreting research in PSPB: Practices, principles, and pragmatics. *Personality and Social Psychology Bulletin*, *35*, 1131-1142. doi: 10.1177/0146167208331253