THE ROLE OF PSYCHOLOGICAL DISCREPANCY IN BELIEF CHANGE: TESTING FOUR MODELS WITH A SINGLE MESSAGE

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

---

by
Luling Huang
May 2020

Examining Committee Members:

Deborah A. Cai, Advisory Co-Chair, Communication and Social Influence
Edward L. Fink, Advisory Co-Chair, Communication and Social Influence
Bruce W. Hardy, Communication and Social Influence
Stan A. Kaplowitz, Michigan State University

ABSTRACT

In belief change, message discrepancy is the difference between the belief position advocated in a message and a message receiver's initial belief position. Psychological discrepancy is the message discrepancy experienced by the receiver. Existing literature had assumed that a high level of psychological discrepancy discounted the weight of a message, which could make the message less effective. However, there were three alternative assumptions about the role of psychological discrepancy. The problem that this dissertation examined was: Does psychological discrepancy affect the weight of a message only, affect the scale value of the message only, affect neither, or affect both? To find out which of these four assumptions was more plausible, this dissertation derived competing hypotheses based on four mathematical models either through an analytic proof alone or a combination of an analytic proof and a computational approximation. This dissertation tested these hypotheses in an experiment with a 3 (high vs. moderate vs. low message scale value) × 3 (high vs. moderate vs. low upper bound) between-subjects design (*N* = 448 Mechanical Turk workers). The results showed that the weight-discounting model had the most supported hypotheses and fit the data the best, which indicated that psychological discrepancy affected the weight of a message only. This dissertation improves the understanding of the mechanism that leads to the outcomes posited by discrepancy models in persuasion research and provides additional empirical evidence for the scale value constancy assumption in information integration theory.

ACKNOWLEDGMENTS

I sincerely appreciate the support and help from several groups of persons throughout my dissertating process. First, I would like to thank Dr. Deborah Cai and Dr. Edward Fink for being such fantastic mentors. Dr. Cai guided my doctoral studies since my first day at Temple University and set the foundation for me to commit to quantitative communication research. Dr. Fink instilled the courage and perseverance in me to digest difficult readings and explore new ideas. Both Dr. Cai and Dr. Fink made sure that this dissertation would be a success by providing timely and thorough feedback. I would like to thank Dr. Stan Kaplowitz for his insightful comments and challenging questions. Some of his questions made me to refine this dissertation's research questions and rethink this dissertation's significance. Dr. Kaplowitz and Dr. Fink's empirical work on belief change and their ways of doing mathematical modeling in communication science heavily influenced this dissertation. I would also like to thank Dr. Bruce Hardy for his constructive feedback on this dissertation's broader implications. In addition to the excellent committee members, I also appreciate the input from Dr. Sungeun Chung at the early stages of this dissertation.

Next, I would like to thank the Office of Research and Graduate Studies at the Klein College for keeping my dissertating schedule on track, processing the dissertation-related paperwork, and providing funding support. I owe many thanks to Dr. Patrick Murphy, Nicole McKenna, and Kaitlin Fitzsimmons. When I worked on this dissertation, I also worked part-time at the Temple Library's Scholars Studio. Not only the Scholars Studio partially funded my dissertating period, I had the opportunities to develop the computational skills needed in my dissertation. I would like to thank my supervisors at the Scholars Studio: Dr. Peter Logan, Dr. Marcus Bingenheimer, Dr. Alex Wermer-Colan, Dr. Jennifer Grayburn, and Matt Shoemaker.

I also appreciate the support from my friends and colleagues in the Media and Communication Doctoral Program. I would like to give special thanks to Dr. Kun Xu, who gave helpful comments on this dissertation's research questions, study design, and data analysis.

Finally, I would like to thank my parents and grandparents for their invaluable support. They gave their all to help me succeed. My relative, Amy Lou, kindly tolerated me living at her house for seven years and provided spiritual support all along.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

In the field of persuasion, the essential problem is how the formation, change, or reinforcement of attitude, belief, and behavior occurs in the process of communication (Miller, 2012; O'Keefe, 1990). Many theories and models that examine the persuasion process have been proposed and tested (Dillard & Shen, 2012; Eagly & Chaiken, 1993). Some of them have developed mathematical models to study persuasion (Hunter, Danes, & Cohen, 1984; Hamilton, Hunter, & Boster, 1993). Unlike some theories in persuasion in which propositions about the relationship between concepts are stated verbally, the significance of using mathematical models is to state explicit assumptions, to derive hypotheses symbolically, and to make precise predictions (Kaplowitz, Fink, & Bauer, 1983, p. 233). The goal of mathematical modeling is not to make the study of persuasion seem to be more difficult, because the complexity of the process of social influence can be explained by simple rules expressed in mathematical forms (Fink, 1996; Latané, 1996). The discrepancy model (Fink & Cai, 2012; Kaplowitz & Fink, 1997) and information integration theory (Anderson, 1981, 2008) are parsimonious models used to explain and predict the process of persuasion.

This dissertation, built on Fink, Kaplowitz, and Bauer's (1983) psychological-discrepancy-discounting model, found a problem in one assumption with respect to the role of psychological discrepancy in that model. To illustrate a general case of the problem, two hypothetical stories can be compared. In the first story, Person A attempts to persuade Person B to donate $100 to Charity X; Person B's initial belief position about an approximate donation is $20 with a range (according to Person B) from $10 to $50; upon Person A's request, how would Person B evaluate this request (e.g., Is it reasonable? Does it seem too extreme?) and evaluate Person A (e.g., Is this person trustworthy)? Finally, what would be Person B's new belief position? In the second story, all elements in the first study stay the same except that before making the same request (i.e., asking Person B to donate $100), Person A mentions that the donation that Charity X has received so far ranges from $10 to $500. The question is: In which story is the request more effective?

According to the psychological-discrepancy-discounting model (Fink et al., 1983), the request in the second story would be more effective. Person B would perceive the request in the second story as less discrepant from the initial position (i.e., $20), because Person B's perspective (i.e., a person's range of belief positions that he or she takes into account; Ostrom & Upshaw, 1968) is wider in the second story than the perspective in the first story (the equation that explains this conclusion is provided in Chapter 2). Fink et al. (1983) assumed that, because the psychological discrepancy would be greater in the first story, the only effect of psychological discrepancy was that the weight (i.e., the importance) of the request would be discounted more than the weight in the second story. Therefore, the request in the second story would be more effective. However, does a change in psychological discrepancy (i.e., the same request being perceived as more or less discrepant from one's initial belief position) also imply a change of the perceived scale value of the request? That is, in Person B's perception, is it possible that the $100 request would be perceived as $300 (i.e., a contrast effect; Sherif & Hovland, 1961) in the first story? Is it also possible that the $100 request would be perceived as $80 (i.e., an assimilation effect; Sherif & Hovland, 1961) in the second story? Following this line of reasoning, the reader may ask: Why would the increased perceived position in the first story be less effective than the decreased perceived position in the second story? This is because both the scale value and the weight of the request may vary as a function of psychological discrepancy.

In short, there are several alternative assumptions to the one made in Fink et al. (1983) about the role of psychological discrepancy in belief change. Does psychological discrepancy affect the weight of a message only (Fink et al., 1983), affect the scale value of the message only, affect neither, or affect both? The purpose of this dissertation is to examine which of these four assumptions would be more plausible.

Except the original model proposed in Fink et al. (1983) that assumes that psychological discrepancy affects the weight of a message only, for each of the three alternative assumptions, a mathematical model was proposed. The three alternative models were specified in a way that they would generate similar observational consequences as those found in Fink et al.'s (1983) study. To compare and test the four mathematical models, this dissertation derived competing hypotheses. These hypotheses were tested in an experiment. The results showed that the weight-discounting model fit the data the best, which indicated that psychological discrepancy affected message weight only.

Table 1 provides a notation reference for the symbols used in the four models.

Table 1

Notation

Symbol	Definition	Example
s	Message scale value	A message advocates a 15% of tuition increase; s is 15%.
s_0	Subject's initial belief position	Before message receipt, I believe a reasonable percent of tuition increase is 5%; s_0 is 5%.
R	Response of belief position after message receipt	After message receipt, I believe a reasonable percent of increase is 10%; R is 10%.
$R - s_0$	Amount of belief change	My belief change after message receipt is 10% − 5% = 5%.
D	$D = \|s - s_0\|$. Positional discrepancy	Positional discrepancy is 15% − 5% = 10%.
w_0	Weight of subject's initial belief position	Before message receipt, tuition increase is not important for me; w_0 is 10 based on a magnitude scale.*
w	Weight of message	The message makes a strong argument for a tuition increase of 15%; w is 500 based on a magnitude scale.*
U	Upper bound of subject's belief position	I believe the maximally high tuition increase is 50%; U is 50%.
L	Lower bound of subject's belief position	I believe the maximally low tuition increase is 1%; L is 1%.
P	Perspective; $P = U - L$	P = 50% − 1% = 49%
k	A scaling constant	

Table 1

(continued)

Symbol	Definition	Example
ψ	Psychological discrepancy, where $\psi = kD/(U - L)$	I perceive that the psychological difference between the message's position and my own initial position is 250 based on a magnitude scale.*
$\Delta(\psi)$	A change function of psychological discrepancy. $\Delta(\psi) = e^{-\gamma\psi}$, or $\Delta(\psi) = 1 - e^{-\gamma\psi}$ depending on models.	
γ	A parameter that affects the rate of change in $\Delta(\psi)$ with respect to ψ.	
w_p	Perceived message weight; $w_p = w\Delta(\psi)$ if w_p is a function of ψ; otherwise $w_p = w$	If w_p is a decreasing function of ψ, I perceive that the message weight is 100 due to a large ψ.
s_p	Perceived message scale value; $s_p = s\Delta(\psi)$; otherwise $s_p = s$	If s_p is an increasing function of ψ, I perceive that the message scale value is 30% due to a large ψ.
k'	A scaling constant in the complex model	

Note. All of the variables are assumed to be positive.

*A magnitude scale, where 100 is a moderate degree of importance or difference, zero is not important or no difference at all, and there is no upper bound.

CHAPTER 2

WHAT IS THE PROBLEM?

The Basic Models

Anderson's (1981) Averaging Model

The information integration approach (Anderson, 1981) examines how people integrate their beliefs when they receive multiple messages. The belief after receiving *n* pieces of information is represented as a weighted average of one's prior belief and the beliefs of any incoming messages (1981, p. 145):

$$R_n = \sum_{i=0}^{n} w_i s_i, \text{ where } \sum_{i=0}^{n} w_i = 1, \tag{1}$$

where R_n is the belief after a subject processes *n* pieces of information, w_i is the *context-dependent* weight of the ith piece of information, and s_i is its scale value. The context-dependent weight means that the weight of a piece of information depends on other pieces of information. Here, the weights in Equation 1 sum to one (Anderson, 1981). And, R_n and s_i are assumed to be on the same unidimensional scale; s_0 is the subject's prior belief position on the same scale.

An alternative form of R_n is:

$$R_n = \frac{\sum_{i=0}^{n} w_i s_i}{\sum_{i=0}^{n} w_i}, \tag{2}$$

where w_i is the *independent* weight of the *i*th piece of information. The difference of the weights in Equation 2 from those in Equation 1 is that the independent weight of a piece of information does not depend on other pieces of information.

Conceptually, weight represents the importance (Himmelfarb, 1975) or "amount of information" (Anderson, 2008, p. 43; Saltiel & Woelfel, 1975) of an incoming message or of a person's prior beliefs. For an incoming message, the psychological meaning of weight can be its salience, relevance (Anderson, 1981), or informativeness (Fiske, 1980). Therefore, if one conceptualizes weight with regard to its effects, the weight of an incoming message can represent the effectiveness of this message. In other words, one can gauge the weight of a message by examining the amount of belief change induced by this message (Saltiel & Woelfel, 1975). For a person's prior beliefs before message receipt, weight can be the strength of a person's pre-existing beliefs (Anderson, 2008; Chung, Fink, Waks, Meffert, & Xie, 2012), which represents the resistance of this person's initial beliefs to change. If one conceptualizes weight with regard to its causes, one of such causes is the number of messages that accumulate to form a person's initial beliefs (Saltiel & Woelfel, 1975).

Relation to Linear Discrepancy Models

In the area of belief change within persuasion, discrepancy models posit that the amount of belief change is a function of discrepancy, where discrepancy is the difference between the advocated position in a message and a subject's initial position on some object, and the amount of belief change is assumed to be the difference between the subject's final position and the subject's initial position (Fink & Cai, 2012; Kaplowitz & Fink, 1997). The theoretical and practical significance of discrepancy models can be seen in this general research question: To optimize the effectiveness of a persuasive message, what level of the message's discrepancy should be used? For example, to persuade undergraduate students, who initially believe there should not be a tuition increase, to adopt a higher level of college tuition, is a message advocating a 50% increase more effective than a message advocating a 15% increase? The effectiveness here is defined as the amount of belief change generated by a message (i.e., the difference between a subject's final position after message receipt and his or her initial position before message receipt).[1]

Linear discrepancy models posit that discrepancy is linearly related to belief change (Anderson & Hovland, 1957). Thus, after message receipt and when one's belief position is in equilibrium, the amount of belief change on a tuition increase is a linear function of the level of the advocated tuition increase, where subject's initial position is held constant.[2] Information integration theory has the same logic as linear discrepancy models (Fink & Cai, 2012). The relation between the averaging model and linear discrepancy models can be formulated with two perspectives. First, Equation 2 indicates that *each successive belief change* ($R_i - R_{i-1}$ for $i \geq 1$) is a linear function of the discrepancy between s_i and R_{i-1}. Let j index the jth message among i messages:

$$R_i = \frac{\sum_{j=0}^{i-1} w_j}{\sum_{j=0}^{i} w_j} R_{i-1} + \frac{w_i s_i}{\sum_{j=0}^{i} w_j}. \tag{3}$$

Then, belief change between $R_i - R_{i-1}$ is:

$$R_i - R_{i-1} = \frac{w_i}{\sum_{j=0}^{i} w_j} (s_i - R_{i-1}), \tag{4}$$

where $s_i - R_{i-1}$ is the discrepancy between the advocated position of the ith message and one's belief position after receiving $i - 1$ messages.[3]

Second, the amount of belief change between R_i and R_0 is the weighted average of $s_1 - R_0$, $s_2 - R_0$, $s_3 - R_0$, . . ., $s_i - R_0$, where the independent weight for each discrepancy (i.e., $s_i - R_0$) is w_i. The proof for this conclusion can be done by subtracting s_0 from Equation 2. The simplest occasion for a message-induced belief change is when there is only one incoming message (see also Kaplowitz & Fink, 1991, p. 191), where

$R_1 - R_0 = \frac{w_1}{w_0+w_1}(s_1 - R_0)$.

Nonlinear Discrepancy Models

Later researchers found a nonlinear relationship between discrepancy and belief change (Aronson, Turner, & Carlsmith, 1963; Bochner & Insko, 1966; Chung, Fink, & Kaplowitz, 2008; Fink et al., 1983; Laroche, 1977; see Fink & Cai, 2012, and Kaplowitz & Fink, 1997, for a thorough discussion). The nonlinear relationship can either be monotonically increasing but decelerating or have an inverted-U shape (nonmonotonic). Either way, the interpretation makes intuitive sense: If a message advocates a too extreme position, such as a 50% tuition increase, the effectiveness of the message can be reduced compared with a message advocating only a moderately discrepant 15% increase. Several theoretical approaches can account for the nonlinear relationship between discrepancy and belief change, as summarized in Fink et al.'s (1983) work. For an extremely discrepant message, the social judgment approach (Sherif & Hovland, 1961) states that the message should be perceived by a subject as even more psychologically discrepant and thereafter rejected. The cognitive dissonance approach (Aronson et al., 1963) states that a subject should disparage the source of the message, disparage the content of the message, or seek social support, in order to reduce the dissonance generated by the extremely discrepant message (see also Bochner & Insko, 1966). The cognitive response approach (Brock, 1967) states that the subject should generate more counterarguments to a highly discrepant message than to a moderately discrepant one, which reduces the effects of the extremely discrepant message. Finally, the information integration approach (Anderson, 1981) states that a subject assigns less weight to a message due to less attention paid to it or due to its inconsistency with prior beliefs.

The Psychological-Discrepancy-Discounting Model

Fink et al. (1983) developed a model based on Anderson's (1981) averaging model and Laroche's (1977) model. Fink et al.'s (1983) model, a variant of Equation 2 above, made it possible to examine the nonlinear relationship between discrepancy and belief change, as well as the interdependence among messages in multiple message conditions.

First, the key equation in Laroche's (1977, p. 249) model is:

$$\Delta A = De^{-\gamma D}, \tag{5}$$

where ΔA is the amount of attitude change, D is the discrepancy between the advocated position in a message and the initial position held by a subject, and γ is a function of source credibility and the subject's degree of noninvolvement with the message's content; γ is assumed to be equal to or greater than 0. Both ΔA and D are assumed to vary between 0 and 1. Equation 5 indicates that when $\gamma = 0$, ΔA linearly increases with D. When $0 < \gamma \leq 1$, ΔA increases with a decelerating rate as D increases. When $\gamma > 1$, ΔA increases first and then decreases (i.e., it is nonmonotonic).

Fink et al. (1983) made a conceptual distinction between positional discrepancy (i.e., $s_i - R_{i-1}$ in Equation 4) and psychological discrepancy (i.e., the discrepancy *as experienced by the subject*). The authors assumed that the same amount of positional discrepancy could be experienced differently by a subject due to individual and contextual differences. The authors proposed that, keeping the positional discrepancy of a message constant, as the level of its psychological discrepancy increased, the subject would discount the message more, which lessens the effectiveness of that message. Fink et al.'s (1983) model indicates that belief change is a nonmonotonic function of *D* (1983, p. 418n22).

The following presents the functional forms of Fink et al.'s (1983) model. Let ψ be the psychological discrepancy of a message and $\Delta(\psi)$ be a function of ψ. Suppose a subject receives one message. Let R be the subject's belief position in equilibrium after receiving the message. Then,

$$R = \frac{w_0 s_0 + w\Delta(\psi)s}{w_0 + w\Delta(\psi)}, \tag{6}$$

and $\Delta(\psi) = e^{-\gamma\psi}$ (Fink et al., 1983, p. 418). (7)

Equation 7 is an exponential decay function when γ is assumed to be greater than 0. This indicates that as ψ increases, $\Delta(\psi)$ decreases, while keeping γ constant.

In Equation 6, if $\Delta(\psi)$ is constant, then $R = \frac{w_0 s_0}{w_0 + w\Delta(\psi)} + \frac{w\Delta(\psi)}{w_0 + w\Delta(\psi)} s$, which in effect is the linear model based on Equation 2. Ceteris paribus, R is a linearly increasing function of s. However, if $\Delta(\psi)$ is allowed to vary as a function of ψ, the relationship between R and s becomes nonmonotonic. More specifically, as s increases, R first increases and later decreases with the assumption of $\psi = \alpha D^{\beta}$ (Fink et al., 1983, p. 418n22).

What determines psychological discrepancy? In addition to the positional discrepancy, *D*, Fink et al. (1983, p. 418) suggested three more factors. First, messages without arguments or with weak arguments may have a greater psychological discrepancy (i.e., a greater level of discrepancy as experienced by a subject) than messages with strong arguments. Second, messages provided by low credible sources may have greater psychological discrepancy. Third, given multiple messages, psychological discrepancy of one message may change because of other messages surrounding it. For this latter contextual factor, Fink et al. (1983) drew upon Ostrom and Upshaw's (1968) perspective theory to elaborate on how to use different combinations and orders of an extremely discrepant message and a moderately discrepant one to manipulate psychological discrepancy.

Fink et al. (1983) hypothesized that the psychological discrepancy of a moderately discrepant message would be reduced by presenting an extremely discrepant message first. In contrast, the psychological discrepancy of an extremely discrepant message would not change regardless of context. The rationale was that the extremely discrepant message extended the boundary of one's perception about how discrepant a message could be. Therefore, a moderately discrepant message would seem to be less discrepant if presented after an extremely discrepant message than if it was presented alone. The reverse was not true because a moderately discrepant message would be less likely to extend the perceptual boundary for an extremely discrepant message. Based on this rationale, Chung and Fink (2017) provided an updated functional form of the psychological-discrepancy-discounting model based on perspective theory:

$$R = \frac{w_0 s_0 + w\Delta(\psi)s}{w_0 + w\Delta(\psi)}, \text{ where } \Delta(\psi) = e^{-\gamma\psi} \text{ and } \psi = k\frac{D}{U-L}. \tag{8}$$

In Equation 8, ψ is an increasing function of D and a decreasing function of the difference between the upper bound of a range of a person's belief positions that he or she takes into account, U, and the lower bound of this range, L. Similar to the assumption $\psi = \alpha D^{\beta}$, the relationship between R and D is also nonmonotonic when $\psi = kD/(U - L)$, where k is a positive constant (see Appendix A).

The results of Fink et al.'s (1983) experiments supported the proposed hypotheses. The stimuli included messages advocating two levels of a college tuition increase: 15% (moderately discrepant) and 50% (extremely discrepant). Subjects were randomly assigned to one of seven message conditions: no message as the control condition, single 15% message, single 50% message, 15% message then 50% message, 50% then 15%, 15% then 15%, and 50% then 50%. After message receipt, the subjects were also asked how much difference there was between the view in the message and their own view, which was a measure of psychological discrepancy. Also, the subjects responded to how much they believed that the tuition should be increased, which was a measure of belief position.

All but one hypothesis was supported based on the psychological-discrepancy-discounting model. The critical statistical test was a comparison of the linear discrepancy model with the model that takes psychological discrepancy into account. The results from a nonlinear regression indicated that for both the single message and the double message conditions, the psychological discrepancy model significantly increased the amount of explained variance from the linear model (Fink et al., 1983).

Relation of Perspective to Latitude of Acceptance

The term $U - L$ in Equation 8 is defined as the *perspective* (P) of a subject's belief. This construct was developed in perspective theory (Ostrom & Upshaw, 1968) and was similar to the concept of *latitude of acceptance* in social judgment theory (Sherif & Hovland, 1961; see also Eagly & Telaak, 1972; Granberg & Steele, 1974; Johnson, Lin, Symons, Campbell, & Ekstein, 1995). In Equation 8, as P increases, ceteris paribus, ψ decreases, and R decreases, which is consistent with Eagly and Telaak's (1972) finding that for all levels of positional discrepancy, the amount of attitude change for the participants with wide latitudes of acceptance was greater than the amount of attitude change for the participants with medium or narrow latitudes of acceptance. Equation 8 explains the psychological mechanism underlying Eagly and Telaak's (1972) finding, in which psychological discrepancy is a crucial mediating variable between latitude of acceptance and attitude change.

This dissertation does not use the concept of latitude of acceptance for two reasons. First, social judgment theory has two other categories of latitude: *latitude of rejection* and *latitude of noncommitment* (Sherif, Sherif, & Nebergall, 1965). The widths of the three categories of latitude indicate different levels of ego-involvement in a topic and relate to different individual levels of resistance to persuasion attempts, which implies interdependence between width of latitude and one's involvement in some topic (Sherif, Kelly, Rodges, Sarup, & Tittler, 1973), and between width of latitude and one's resistance to persuasion (Johnson, Maio, & Smith-McLallen, 2005, pp. 620-621). In the nonlinear discrepancy models proposed in this dissertation, the perspective of a subject's belief positions was assumed to be independent of involvement and resistance. Instead, a subject's level of involvement and resistance, or the strength of an individual's initial belief more generally, was represented by the value of his or her weight of initial belief position (w_0; Chung et al., 2012; Saltiel & Woelfel, 1975). That is, P and w_0 are independent in the models presented in this dissertation.

Second, although conceptually similar, the perspective (P) of a subject's belief positions and his or her latitude of acceptance are operationalized differently. In this dissertation, the manipulation of P is more similar to perspective theory's manipulation of perspective (i.e., $U - L$; Ostrom, 1970; see also Ostrom & Upshaw, 1968, p. 229). In an empirical testing of perspective theory (Ostrom, 1970), the subjects read a case history of a man guilty of a bomb threat and were asked to indicate what they believed a fair sentence should be for that man. Then, perspective was manipulated by telling the subjects that "the range of prison sentences provided by law" (1970, p. 283) for a bomb threat was either from one to five years (narrow perspective) or from one to 30 years (wide perspective). In the social judgment studies (Eagly & Telaak, 1972; Granberg & Steele, 1974), however, the width of the latitude of acceptance for a subject was measured by counting the number of statements that the subject indicated as acceptable among a list of ordered attitudinal statements on some issue. Note that a *count* of acceptable statements does not tell us precisely what the *amount* of the difference between a subject's U and L is (e.g., with year of sentence or percent of tuition increase).

Given the above two reasons, the concept of latitude of acceptance was not used in this dissertation.

A Philosophical Problem in Fink Et Al.'s (1983) Model

As stated above, the model with the psychological discrepancy discounting factor explained significantly more variance than the linear model. However, the term $w\Delta(\psi)s$ in Equation 6 suggests a philosophical problem: Does the discounting factor $\Delta(\psi)$ discount the *weight* assigned to the message (w) or the *scale value* of the message (s)? Whereas the assumption in Fink et al. (1983) was that only the weight was discounted (see also Kaplowitz, Fink, Armstrong, & Bauer, 1986, p. 510; Kaplowitz & Fink, 1997, p. 86), the multiplication form of $w\Delta(\psi)s$ alone suggests that either interpretation may be valid. In fact, there are two more alternative interpretations.

On the one hand, the weight discounting interpretation may be more plausible than the scale value discounting interpretation: With the inclusion of the psychological-discrepancy-discounting factor, the subjective scale value of a message remains constant and the weight gets discounted, so the net effect of the message on belief change is reduced. Although seemingly less plausible, it is also possible that, with the inclusion of the psychological-discrepancy-discounting factor, the weight of a message remains constant but the subjective scale value gets discounted, with the same reduced effect of the message on belief change. The difficulty in the scale value discounting interpretation is to answer what is meant by discounting a message's scale value. Does it mean the subject perceives the advocated position as closer to his or her own initial position? If this were true, then a message advocating a 50% tuition increase may be perceived as only advocating a 40% increase given that a subject's own initial position is zero. However, a stimulus based on numerical values (e.g., "a 50% increase" rather than "a huge increase") have relatively fixed meanings, which suggests that it is implausible for a subject to change a scale value represented by a number.

More importantly, there is a logical contradiction for the scale value discounting assumption. If this assumption is true, the greater the level of psychological discrepancy, the more the perceived message scale value is "dragged" towards one's initial position to reduce the effectiveness of the message. But this results in a lesser amount of psychological discrepancy. How can this situation, in which more psychological discrepancy results in less psychological discrepancy, be possible? The logical contradiction leads to the possible assumption that psychological discrepancy affects the message weight and message scale value separately (Anderson, 2008), so that when the psychological discrepancy is greater, the perceived message scale value is greater, and the message weight is less.

The Four Models

Four models are displayed in a 2 × 2 table (Table 2) by different roles for psychological discrepancy. The first model (Equation 8) is just the original psychological-discrepancy-discounting model (Fink et al., 1983). This model assumes that only the message weight is discounted due to the effect of psychological discrepancy, as can be seen in the product, $w\Delta(\psi)s$, in the denominator. In this dissertation, this model is referred to as the weight-discounting model. The following presents the functional forms of the other three models.

Table 2

Conceptual Table of the Four Models: The Role of Psychological Discrepancy

<table>
<tr><th></th><th></th><th colspan="4">Does psychological discrepancy affect message weight?</th></tr>
<tr><th></th><th></th><th colspan="2">No</th><th colspan="2">Yes</th></tr>
<tr><td rowspan="4">Does psychological discrepancy affect message scale value?</td><td rowspan="2">Yes</td><td colspan="2">Psychological-discrepancy-scale-value-pullback model</td><td colspan="2">Complex model</td></tr>
<tr><td>Method to determine sign of derivative</td><td>$\frac{\partial R}{\partial P}$: AP
$\frac{\partial R}{\partial s}$: AP</td><td>Method to determine sign of derivative</td><td>$\frac{\partial R}{\partial P}$: AP and CA
$\frac{\partial R}{\partial s}$: AP and CA</td></tr>
<tr><td rowspan="2">No</td><td colspan="2">Independent-psychological discrepancy model</td><td colspan="2">Psychological-discrepancy-weight-discounting model</td></tr>
<tr><td>Method to determine sign of derivative</td><td>Same as the scale-value-pullback model</td><td>Method to determine sign of derivative</td><td>$\frac{\partial R}{\partial P}$: AP
$\frac{\partial R}{\partial s}$: AP</td></tr>
</table>

Note. In this dissertation, the psychological-discrepancy-weight-discounting model psychological-discrepancy-scale-value-pullback model are sometimes referred to as the weight-discounting model and the scale-value-pullback model, respectively. For the method to determine the sign of a first partial derivative: AP and CA stand for analytic proof and computational approximation, respectively.

The Psychological-Discrepancy-Scale-Value-Pullback Model

Whereas Equation 8 assumes that the weight is discounted, an alternative model assumes that only the message scale value varies as a function of ψ:

$$R = \frac{w_0 s_0 + w\Delta(\psi)s}{w_0 + w}, \text{ where } \Delta(\psi) = e^{-\gamma\psi} \text{ and } \psi = k\frac{D}{U-L}. \tag{9}$$

Equation 9 has the same set of parameters as Equation 8 except that Δ(ψ) is removed from the denominator.

Note that in Equation 9, when $s_0 > s$, the effectiveness of a message increases as $\Delta(\psi)s$ is further away from s_0 than s is. Is this statement plausible? From the social judgment approach, the above statement is similar to a contrast effect (Sherif & Hovland, 1961). However, the contrast effect should lead to the rejection of a highly discrepant message, and it should lead to less belief change. However, Equation 9 indicates that a message with high psychological discrepancy leads to more belief change when $s_0 > s$.

For the scale-value-discounting model (Equation 9), the term "discounting" is problematic when $s_0 > s$. Discounting implies a reduction of scale value towards zero. This meaning works well for the weight-discounting-model and the scale-value-discounting model when $s_0 < s$, because it aligns with the assumption that the greater the psychological discrepancy, the less the effectiveness of a message. If this assumption also holds for the scale-value-discounting model when $s_0 > s$, Equation 9 needs to be adjusted as follows:

$$R = \frac{w_0 s_0 + w\Delta(\psi)s}{w_0 + w}, \text{ where } \psi = k\frac{D}{U-L}, \text{ and } \Delta(\psi) = \begin{cases} e^{-\gamma\psi} & \text{if } s_0 < s, \\ 1 - e^{-\gamma\psi} & \text{if } s_0 > s. \end{cases} \tag{10}$$

Equation 10 is the psychological-discrepancy-scale-value-pullback model. It is consistent with the idea that, as a message's psychological discrepancy increases, the force of resisting change in the advocated direction increases, as if there is a force pulling one's belief position back to the initial position, s_0. The essential idea is still consistent with the nonlinear relationship between positional discrepancy and belief change (Chung et al., 2008; Fink & Cai, 2012; Fink et al., 1983; Laroche, 1977). However, the scale-value-pullback model is a more complete model that takes the situation of $s_0 > s$ into account.

A Logical Contradiction in the Scale-Value-Pullback Model

According to the scale-value-pullback model, keeping all other parameters constant, the greater the psychological discrepancy, the closer the subjective scale value is to the initial position. However, how can this occur in a logical way? Consider two messages that have the same positional discrepancy, where the first message has a greater level of psychological discrepancy than the second message. Why does the first message, which is perceived as psychologically more discrepant than the second message, become closer to the initial position in terms of its subjective scale value? Note that the effectiveness of the first message is indeed less than that of the second message because of its subjective scale value being closer to the initial position, which is consistent with the original hypothesis of the relationship between psychological discrepancy and message effectiveness (Fink et al., 1983).

Therefore, the crucial question is whether the scale-value-pullback model should be eliminated only based on logic. Even if the empirical evidence provided by this dissertation supports the scale-value-pullback model, how can it logically be true? There are two arguments that defend the scale-value-pullback model.

First, if time is considered, it is possible that a highly psychologically discrepant message can *later* be perceived as closer to one's initial position. Without empirical evidence, this possibility cannot be *a priori* eliminated. Second, the paramorphic representation of mathematical models (Dawes & Corrigan, 1974) can explain this logical contradiction. The scale-value-pullback model does not mean that a subject actually changes the semantic meaning of a message in a logical way. Instead, the model functions *as if* the subject pulls the message scale value closer to his or her initial position even if the message's psychological discrepancy is high.

The Independent-Psychological-Discrepancy Model

The third model has the same functional form as the scale-value-pullback model (Equation 10). Therefore, the net effect of belief change as a function of psychological discrepancy is the same as the scale-value-pullback model. The difference between the independent-psychological-discrepancy model and the scale-value-pullback model is that the factor $\Delta(\psi)$ in Equation 10 is independent of message weight and independent of message scale value. Thus, one way to distinguish the two models is to measure the perceived message scale value to see if the perceived scale values of the messages with different levels of psychological discrepancy are significantly different from each other. If they are, the scale-value-pullback model is favored between the two models.

The Complex Model

The fourth model assumes that psychological discrepancy affects message weight and message scale value separately. The functional form when $s_0 < s$ is as follows:

$$R = \frac{w_0 s_0 + w\Delta(\psi)_{\mathrm{w}} \cdot [s\Delta(\psi)_{\mathrm{s}} + s_0]}{w_0 + w\Delta(\psi)_{\mathrm{w}}},$$

$$\text{where } \psi = k\frac{D}{U-L},\ \Delta(\psi)_{\mathrm{w}} = e^{-\gamma\psi},\ \text{and } \Delta(\psi)_{\mathrm{s}} = k'(1 - e^{-\gamma\psi}). \tag{11}$$

In Equation 11, $\Delta(\psi)_{\mathrm{w}}$ and $\Delta(\psi)_{\mathrm{s}}$ represent psychological discrepancy's effect on the message weight and the message scale value, respectively; k' is a positive constant that allows the possibility that the perceived scale value as a function of psychological discrepancy is greater than the explicit message scale value when $k' > 1$. And $\Delta(\psi)_{\mathrm{s}}$ describes a relationship in which the perceived message scale value increases as psychological discrepancy increases.

This model satisfies the following assumptions. First, the perceived message scale value that has high psychological discrepancy is greater than that of a message that has low psychological discrepancy. Together with the assumption of $\psi = kD/(U - L)$, this assumption is consistent with perspective theory (Judd & DePaulo, 1979; Ostrom & Upshaw, 1968). Also, it resolves the logical contradiction in the scale-value-pullback model, as explained above.

Second, the weight of a message that has high psychological discrepancy is discounted more than that of a message that has low psychological discrepancy. Thus, after discounting, the high psychological discrepancy message is weighted less than the low psychological discrepancy message. This is consistent with Fink et al.'s (1983) assumption.

Third, the perceived message scale value is assumed to be equal to or greater than the initial position, so the belief position after message receipt is always greater than the initial position when $s_0 < s$. In other words, the boomerang effect is not allowed in this model. This is consistent with Kaplowitz and Fink's (1997, p. 85) conclusion that the empirical evidence for the boomerang effect is scarce (but see Zhao & Fink, 2018).

Fourth, the relationship between positional discrepancy and belief change is nonmonotonic. A partial analytic proof and a computational demonstration of the nonmonotonic relationship are discussed in Chapter 5. The graph of the nonmonotonic relationship is presented in Figure 1.

The fifth assumption is the critical difference from the weight-discounting model and the scale-value-pullback model. In the complex model, the relationship between psychological discrepancy and belief change is no longer monotonic (i.e., the greater the psychological discrepancy, the less the belief change). Instead, the relationship between psychological discrepancy and belief change is nonmonotonic, where R first increases and then decreases, as ψ increases. Because $\psi = kD/(U - L)$, ceteris paribus, one can examine how R changes as a function of P. A partial analytic proof and a computational demonstration of the nonmonotonic relationship between R and P are discussed in Chapter 5. Ceteris paribus, initially the amount of belief change increases as P increases. After reaching a certain level of P, the amount of belief change decreases, as P increases. The graph of the nonmonotonic relationship is presented in Figure 2.

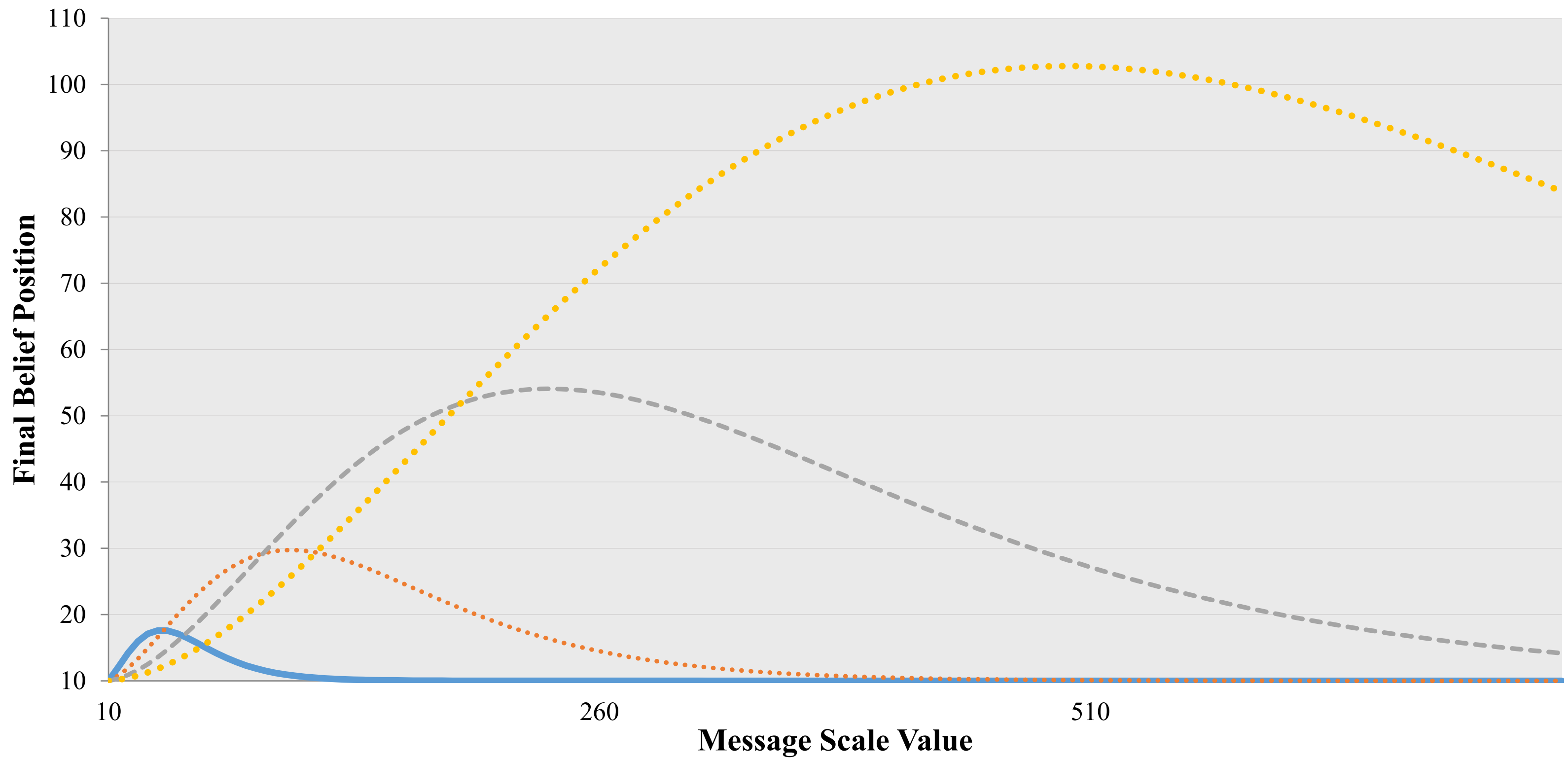


Figure 1. The complex model: Final belief position, R, as a function of message scale value, s, with hypothetical data. Message scale value varies from 10 to 730. The following parameters are kept constant: $w_0 = 0.3$, $s_0 = 10$, $w = 0.5$, $\gamma = -0.4$, $L = 8$, $k = k' = 1$. There are four values of U: $U = 120$ for the top curve, $U = 60$ for second curve, $U = 30$ for the third curve, and $U = 15$ for the bottom curve.

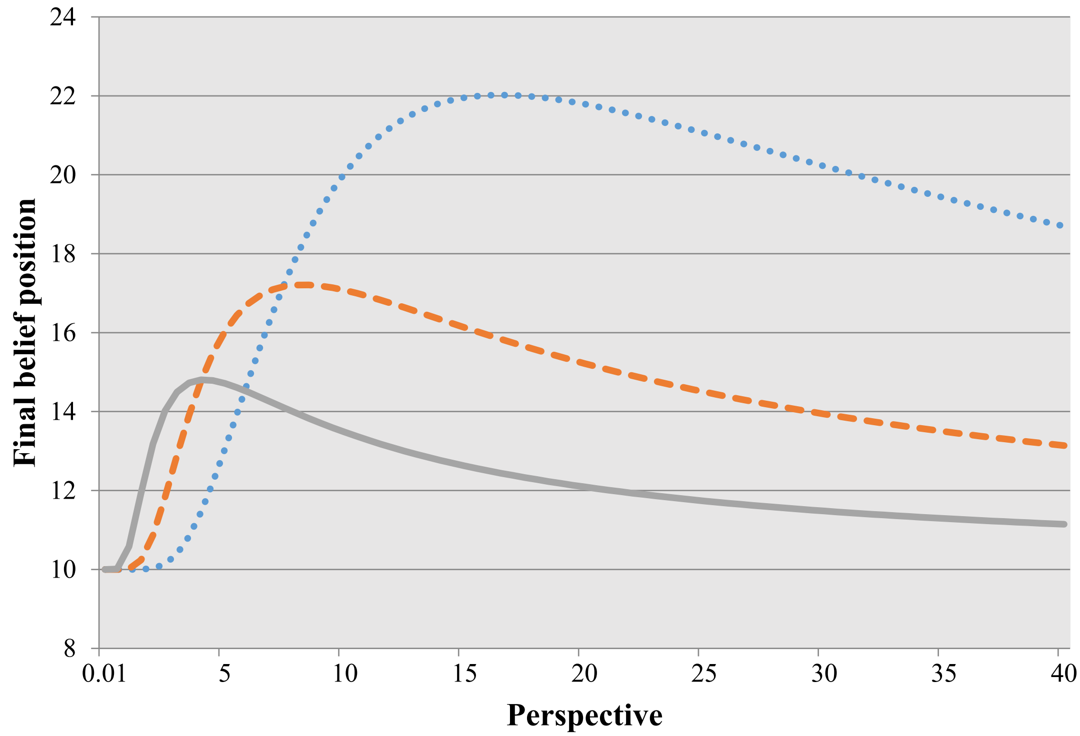


Figure 2. The complex model: Final belief position, *R*, as a function of perspective, *P*, with hypothetical data. The following parameters are kept constant: $w_0 = 0.3$, $s_0 = 10$, $w = 0.5$, $\gamma = -0.4$, and $k = k' = 1$. There are three values of *s*: $s = 50$ for the top curve, $s = 30$ for the middle curve, and $s = 20$ for the bottom curve.

CHAPTER 3

WHAT DOES THE LITERATURE SAY ABOUT THE PROBLEM?

Information Integration and Functional Measurement

In information integration theory (Anderson, 1981, 2008), the process of information integration has three components: valuation, integration, and response. Valuation turns a stimulus's scale value into a subjective value. Imagine there are two messages. Message A advocates a 50% tuition increase and Message B advocates a 15% increase. Suppose that Message A is evaluated by a subject as advocating a 40% increase and Message B is evaluated as advocating a 30% increase. Integration combines the subjective values through *cognitive algebra* (e.g., adding, multiplying, or averaging). Suppose that the subject follows the equal-weight averaging rule (i.e., the values of w_i are equal across messages in Equation 1) of cognitive algebra to combine the subjective message values. By ignoring a subject's initial belief, the resulting combined value for him or her is a $0.5 \times 40\% + 0.5 \times 30\% = 35\%$ increase. Response turns the implicit result from integration into ratings on some scale. The implicit response value (r) is defined as the result of integration (e.g., 35%), and the explicit response value (R) is defined as what the subject reports when asked about his or her belief position after message receipt (e.g., what percent tuition increase he or she believes is reasonable). Suppose that the relationship between r and R is expressed as $R = r - 5\%$. Then, when asked about what should be a reasonable tuition increase, the subject responds that he or she believes a 30% increase is reasonable.

In the above processes of valuation, integration, and response, there are three observable variables (i.e., 50% and 15% increase in the two messages plus the observed 30% increase of R). There are also three unobservable variables (i.e., the assumed 40% and 30% increase after the valuation process plus the assumed 35% increase of r). When we do not a priori know the algebraic rule that the subject follows and what the measurement scale of the response variable is, we cannot solve for the three unobserved variables (Anderson, 1981).

However, with a proper factorial design and an experimental procedure to collect data on the explicit response values, as well as by making assumptions on the integration and the response components (see the next section in this chapter), it is possible to accomplish three things. First, one can find whether the assumed integration rule is plausible. Second, one can find whether the assumed measurement scale is plausible. Third, the scale of the implicit stimulus values can be obtained by calculating the marginal means of the factorial table containing the measured explicit response values. The first of the above three tackles the theoretical problem of how people integrate multiple pieces of information. The other two address the scales that people use when they process stimuli and give responses. Thus, the contribution of information integration theory is to solve both the theoretical problem (in the integration component) and the measurement problem (in the valuation and the response components) at the same time (i.e., the logic of Anderson's, 1981, functional measurement).

The Two-Parameter Formulation and the Parallelism Theorem

In Equation 2, one piece of information can be formulated as the product of its scale value (*s*) and its weight (*w*). The inclusion of weight makes information integration theory flexible to interpret the violation of the *parallelism theorem* (Anderson, 1971a, 1981, 2008). To briefly introduce the parallelism theorem, suppose that a subject is asked to provide a response indicating his or her belief positions after receiving two messages. If each message is treated as a factor, and each factor has three levels on a unidimensional scale for message scale value, then a 3 × 3 design can be created, with three levels of Message I (as the row factor) by three levels of Message II (as the column factor). If it is a within-subject design, the subject will provide a response to each of the 9 message combinations.[4] Once the subject's responses are provided, there are nine data points for the subject. These data points can then be graphed by placing the column factor on the horizontal axis and letting the vertical axis be the ratings that measure the subject's explicit responses. The parallelism theorem states that, given the assumption that the subject's explicit response has a linear relationship with the implicit response, based on a strict additive model or an averaging model with equal weighting within factors (see the explanation in the following two paragraphs), the difference between any two levels of the row factor is constant across all levels of the column factor (Anderson, 1981).

The derivation below from Equation 12 to Equation 16 follows Anderson (1981). The purpose here is to lay out the logic of the parallelism theorem for later discussion. Let r_{ij} be the implicit response from a subject. Based on the strict additive model with

w_0s_0 omitted (because the exclusion of the initial belief does not affect the following derivation),

$$r_{ij} = w_{Ri}s_{Ri} + w_{Cj}s_{Cj}\,, \tag{12}$$

where w_{Ri} and w_{Cj} are the weight for the ith level of message scale value in the row factor and the weight for the jth level of message scale value in the column factor, and s_{Ri} and s_{Cj} are the ith level of message scale value in the row factor, and the jth level of message scale value in the column factor, respectively. The assumption that the explicit response has a linear relationship with the implicit response can be expressed as:

$$R_{ij} = C_0 + C_1 r_{ij}\,, \tag{13}$$

where C_0 and C_1 are parameters for the linear relationship between R_{ij} and r_{ij}. The difference in the explicit responses between the first and the second level of the row factor is

$$R_{1j} - R_{2j} = C_1(w_{R1}s_{R1} - w_{R2}s_{R2}). \tag{14}$$

Because $(w_{R1}s_{R1} - w_{R2}s_{R2})$ is constant across levels in the column factor, the lines connecting the data points across columns (one line for each row level) are parallel. The same applies to the other two differences, $R_{1j} - R_{3j}$ and $R_{2j} - R_{3j}$.

For the averaging model, Equation 12 becomes

$$r_{ij} = \frac{w_{Ri}s_{Ri} + w_{Cj}s_{Cj}}{w_{Ri} + w_{Cj}}\,. \tag{15}$$

Equation 14 becomes

$$R_{1j} - R_{2j} = C_1\left(\frac{w_{R1}s_{R1} + w_{Cj}s_{Cj}}{w_{R1} + w_{Cj}} - \frac{w_{R2}s_{R2} + w_{Cj}s_{Cj}}{w_{R2} + w_{Cj}}\right). \tag{16}$$

Consider Equation 16: If the assumption is that w_{Ri} is constant within the row factor and w_{Cj} is constant within the column factor, one can rewrite both $w_{\mathrm{R1}} + w_{\mathrm{Cj}}$ and $w_{\mathrm{R2}} + w_{\mathrm{Cj}}$ as $w_{\mathrm{R}} + w_{\mathrm{C}}$. Then the component in the parentheses of Equation 16 is simply $\frac{w_{\mathrm{R}}s_{\mathrm{R1}} - w_{\mathrm{R}}s_{\mathrm{R2}}}{w_{\mathrm{R}} + w_{\mathrm{C}}}$. This means that $R_{1\mathrm{j}} - R_{2\mathrm{j}}$ is constant across levels in the column factor. Again, the lines connecting the data points across columns (one line for each row level) are parallel.

If the assumption is that the values of w_{Ri} are different across levels in the row factor and the values of w_{Cj} are different across levels in the column factor, then $R_{1\mathrm{j}} - R_{2\mathrm{j}}$ is no longer constant across levels in the column factor in Equation 16. The averaging model with different weights within factors (differential weighting) can account for any nonparallelism observed in the data.

Change in Weight or Change in Scale Value?

One example of using the averaging model with differential weighting to explain nonparallelism can be negativity effects. A series of experiments (Anderson, 1965, 1968; Anderson & Alexander, 1971) showed that the response data points of the $\mathrm{LM^-}$ and $\mathrm{M^-L}$ message combinations (L is the unfavorable message and $\mathrm{M^-}$ is the mildly unfavorable message) were closer to that of the LL combination than to that of the $\mathrm{M^-M^-}$ combination (Anderson, 1981, p. 274, Figure 4.14). This result showed a deviation from the parallelism theorem. To interpret this pattern, information integration theory proposed that the message with a lower scale value had a heavier weight (Anderson, 1981, pp. 273–276). That is, the averaging model with the additional assumption that a lower scale value has a greater weight can account for nonparallelism.

It is also possible that the weights of messages are the same across the levels in each factor, but the L message is assigned a lower value (i.e., a more unfavorable scale value) by the subjects in the LM^- and the M^-L sets than in the LL set. If this reasoning is correct, it shows the same pattern as above. Note that the discussion so far has not addressed the "important, implicit assumption" (Anderson, 1981, p. 18), which states that the weight and the scale value of a single message are constant regardless of which other messages that single message is combined with. That is, taking Equation 16 as an example, for the jth level message in the column factor, w_{Cj} and s_{Cj} are assumed to be constant regardless of whether they are combined with $w_{R1}s_{R1}$ or $w_{R2}s_{R2}$ in the row factor. If this assumption is lifted, the parallelism theorem as shown above is difficult to achieve (1981, p. 19). However, as shown above in the example of the negativity effect, if either the weight or the scale value is constant, the model analysis is rather straightforward. The problem is which assumption—that an unfavorable message is assigned a greater weight or that an unfavorable message is assigned a more unfavorable scale value—is more plausible.

Evidence That Supported the Scale Value Constancy Assumption

Anderson (1981, 2008) claimed that the observed parallelism in experimental studies provided strong support for the assumption of "meaning constancy" (1981, p. 161) or "meaning invariance" (2008, p. 34). This assumption states that the weight and the scale value of a single message are constant regardless of which other messages that single message is combined with. The information integration approach has had two lines of inquiry that validated this meaning constancy assumption.

The first line of inquiry investigates the problem of how to interpret the "positive context effect" (Anderson, 1981, p. 165). In one study (Anderson, 1971b), the subjects were asked to give an overall impression of a person based on a set of three personal trait messages. Then subjects were asked to rate each trait separately. Empirical data showed that the subject's judgment of the single person trait (either M^+ for a moderately positive message or M^- for a moderately negative one) increased as a function of the scale values of the other two traits (either LL, M^-M^-, M^+M^+, or HH; 1971b, p. 79). One interpretation was that the scale value of one trait in a set of three changed depending on which two other traits were in the set. The other interpretation was that the single scale value did not change; rather, the judgment of the single trait was the result of information integration between the overall impression and the single trait. In other words, because the overall impression was measured before the single trait judgment, as the overall impression changed with different trait combinations, the single trait judgment also changed as a result of it being integrated with the overall impression.

The critical test for deciding between the above two interpretations was to compare the positive context effect between two conditions. In one condition, subjects were asked to write a paragraph about the person about whom the traits described before making an overall judgment. In the other condition, the subjects made judgments without writing a paragraph. It was assumed that in writing a paragraph, the subjects should think about the three traits together in greater detail, which should bring about more flexibility in the traits' meanings. Therefore, if the change of meaning hypothesis is true, the rating of a single trait in the paragraph-writing condition should be more displaced toward the other two traits than the rating of that trait in the no-paragraph condition, and the slope of the increasing ratings should also be greater. However, because it was found that the rating of the single trait and the slope of the increasing ratings were not greater in the paragraph-writing condition than in the no-paragraph condition, Anderson's (1971b) study did not support the change of meaning hypothesis.

The second line of inquiry examined the problem of how to interpret a primacy effect (Anderson, 1973; 1981, pp. 179-191). A primacy effect is defined as the observation that when multiple messages are presented to a subject in a sequence, the earlier messages induce more attitude change than the later messages. The change of meaning interpretation stated that earlier messages either changed the scale value of later messages (Asch, 1946) or discounted the weight of later messages if there was any inconsistency (Anderson & Jacobson, 1965). The competing interpretation was that attention decreased over time (Anderson & Hubert, 1963) without any change of meaning. The critical test of comparing the above two interpretations was to compare two conditions in a study where a series of adjectives about a hypothetical person's traits were presented. In one condition, the subjects were instructed to recall the adjectives after giving an overall impression. In the other, no such instruction was given. The rationale was that if the attention decrement interpretation was plausible, the primacy effect should be eliminated in the recall condition, because it was assumed that the subjects should pay an equal amount of attention to all adjectives in the recall condition. The experimental results showed that the primacy effect was significantly reduced in the recall conditions and even changed to a recency effect (1963, p. 381), whereas the primacy effect persisted in the no-recall conditions. Thus, the results supported the attention decrement interpretation, not the change of meaning interpretation, because if there were any meaning change, there should be no difference between the recall and no-recall conditions.

Tesser (1968) compared the change of weight and change of scale value interpretations of the primacy effect in impression formation. Subjects were asked to rate the scale value of each word in a set of six after giving an overall impression of a person described by those six words. The sequences of words were balanced so that each word appeared once in each one of the six positions. If the change of scale value hypothesis were plausible, the rated scale values of a single word should be different as a function of position. The differential weighting hypothesis states that the weights of the six words differ, which was tested by correlating each word's scale value with the overall impression. If the differential weighting hypothesis were plausible, the correlations of the earlier words to the overall impression should be higher than the correlations of the later words to the overall impression. Tesser's (1968) results did not support the change of scale value hypothesis and provided only partial support for the differential weighting hypothesis.

In sum, the information integration literature has supported the general assumption that the scale value of a message is not susceptible to change with different message combinations (Anderson, 2008; see also Fiske, 1980, p. 893; Ostrom & Davis, 1979, p. 2040). If there were any deviation from the parallelism theorem, the unequal weights within a factor in the averaging model should account for the deviation. Two studies on inconsistency discounting (Anderson & Jacobson, 1965; Himmelfarb & Anderson, 1975) were further reviewed because they were more relevant to the problem of discounting in Fink et al.'s (1983) study. These two studies claimed they supported the meaning constancy assumption. However, Anderson and Jacobson's (1965) paper was based on a problematic model analysis (see below). Also, part of Himmelfarb and Anderson's (1975) results were open to the alternative interpretation of meaning change (see below).

Evidence That Challenged the Scale Value Constancy Assumption

Anderson and Jacobson's (1965) study investigated the hypothesis that the inconsistency among a set of three personal trait adjectives contributes to the decreased weight for the minority adjective. For example, if the combination of three adjectives was HLL (H for favorable and L for unfavorable) and if the subjects were not explicitly told to weight the adjectives equally, the weight of the H message (i.e., the minority adjective) would be discounted due to the inconsistency between the scale value of the H message and the scale values of the two L messages. The results supported the authors' hypotheses. However, there was a problem in the model's analysis (1965, pp. 532-533). The authors assumed that the reduced overall impression under a discounting averaging process (e.g., H is discounted in HLL) from the overall impression under an equal weighting averaging process (e.g., H, L, and L were equally weighted) would mean that the *weight* of the inconsistent term would be discounted. Following the authors' illustration (p. 532), suppose that the H message has a scale value of 3a, the L message has a scale value of zero, and all messages have an equal independent weight of one. In an HLL set, the equal averaging process leads to an overall impression: $[(1 \times 3a) + (1 \times 0) + (1 \times 0)] / (1 + 1 + 1) = a$. With the discounting averaging process, the authors stated that the overall impression should be less than a by *assuming* that the *weight* of the H message was discounted: $[(0.5 \times 3a) + (1 \times 0) + (1 \times 0)] / (0.5 + 1 + 1) = 0.6a < a$. However, a reduced overall impression could also be found by shifting the H message's scale value toward the scale values of the two L messages and by keeping the weight of the H message constant (i.e., shifting $3a$ toward 0 and keeping all the independent weights as 1). For example, if the H message's scale value becomes $1.8a$, then the overall impression is: $[(1 \times 1.8a) + (1 \times 0) + (1 \times 0)] / (1 + 1 + 1) = 0.6a < a$.

Himmelfarb (1975) defended the assumption that the scale value should not change from a model-building perspective. He interpreted the meaning constancy assumption by analogy with the "mass constancy" in Newton's laws (1975, p. 582) and argued that the two assumptions shared "the same logical status" (p. 582). Then he suggested that to explain a change in weight, a third parameter must be included, because scale value is held constant. The inclusion of psychological discrepancy in Fink et al.'s (1983) model did just this, where psychological discrepancy was assumed to change message weight, whereas message scale value was held constant. Still, as mentioned above, psychological discrepancy could discount the scale value rather than the weight of a message. Although Himmelfarb (1975) presented a strong argument, the "bootless philosophical arguments" (p. 581) should be based on empirical evidence.

Himmelfarb and Anderson's (1975) study did present some evidence that supported the meaning constancy assumption. The weight of a message was manipulated by experimental instructions that varied the amount of information needed for opinion attribution. The subjects were presented with three essays (two militaristic ones plus one pacifistic essay) written by a hypothetical student, and they were asked to rate the degree of that student being militaristic versus pacifistic on a bipolar scale. There were three conditions: In the first condition, the subjects were told that the student wrote the three essays under the instruction of writing two essays that expressed his or her own feelings and one pacifistic essay. The weight of the pacifistic essay was assumed to be the lowest compared with the pacifistic essays in the other two conditions (see below) because it was not informative. In the second condition, the subjects were told that the three essays were written under the instruction of writing about the student's own feelings. The weight of the pacifistic essay was assumed to equal its independent weight and to be greater than that in the first condition. In the third condition, the subjects were told that the essays were written under the instruction of writing only militaristic essays. The weight of the pacifistic essay in this condition was assumed to be the highest, because it was supposedly informative about the student's true opinion. In the experimental design, the row factor was the militaristic essays and the column factor was the pacifistic essay. Each factor had three levels (mildly, moderately, and extremely militaristic or pacifistic) of scale value. Based on the averaging model with equal weighting within factors, it was hypothesized that the overall slope of curves on the graph of the results should increase as a function of the pacifistic essay's weight. Also, the vertical spread between curves should decrease as a function of the pacifistic essay's weight (p. 1066). These two hypotheses were both supported.

The significance of Himmelfarb and Anderson's (1975) study was that the hypothesis about the vertical spread did not involve any consideration of the pacifistic essay's scale value per se. In other words, the experimental design separated the weight of the pacifistic essay and its scale value to test one of the hypotheses. However, the hypothesis about the overall slope of curves still rested on the assumption that the difference between the scale value of an extremely pacifistic essay and that of a mildly pacifistic one did not change across conditions. The plausibility of this assumption may still be challenged. For example, in the condition in which the instruction is that the student is told to write three militaristic essays, the difference between a mildly pacifistic essay and an extremely pacifistic one may seem greater for the subjects (probably due to a more flexible evaluation on the pacifistic essay due to inconsistency) than the other two conditions, which suggests a change in the scale value.

In summary, the studies from the information integration approach have relied on the assumption that the scale value of a message should not change. If psychological discrepancy affects belief change, it is the weight that gets discounted, whereas the scale value is constant (i.e., the perceived message scale value equals the explicit message scale value). However, although the message scale value's constancy is more about whether a single message's scale value varies when it is combined with other messages, seldom have studies convincingly excluded the possibility that the message scale value in the integration process does not equal the explicit scale value as a result of message discrepancy.

At the theoretical intersection of the discrepancy model and information integration theory, the crucial argument of this dissertation is that, if the introduction of the concept of psychological discrepancy is necessary for explicating the nonlinear relationship between belief change and discrepancy, then it also necessarily implies that the perceived message scale value does not equal the explicit scale value. The rationale for this argument is that the conceptual distinction between positional discrepancy and psychological discrepancy has already been made. This dissertation proposes that message discrepancy is central to the valuation process in information integration. The perceived message scale value is constructed in the valuation process as a function of message discrepancy (both positional and psychological). Then, the perceived scale value, rather than the explicit scale value, is used in the integration process.

Why Is This Dissertation Significant?

The theoretical significance of this dissertation is twofold. First, from the perspective of the discrepancy model in persuasion, there have been few attempts to examine the psychological process underlying the original psychological-discrepancy-discounting model. Fink et al. (1983) remarked on their study's limitation: "A mechanism explaining the discounting function needs to be explicated and tested" (p. 429). Different model assumptions imply different psychological mechanisms. The original psychological-discrepancy-discounting model was based on the weight-discounting assumption. However, this assumption was not empirically validated.

Second, because the psychological-discrepancy-discounting model (Fink et al., 1983) is based on Anderson's (1971a, 1981) averaging model, this dissertation also has implications for information integration theory, by directly addressing one of the basic assumptions of the information integration approach: the scale value constancy assumption. This assumption (Anderson, 1981, p. 161) states that the scale value of a message is constant across contexts (e.g., a message advocating a 50% tuition increase has a fixed scale value of 50 regardless of whether it is presented alone or whether it is preceded with or followed by a 15% increase message). This assumption has been validated through a series of experiments (Anderson, 1971b, 1973, 1981, 2008; Anderson & Hubert, 1963; Anderson & Jacobson, 1965; Tesser, 1968; see also Fiske, 1980, p. 893; Ostrom & Davis, 1979, p. 2040). Note that most of these experiments examine whether a message scale value changed depending on different message combinations. This line of research does not address how the scale value of a message varies as a function of a participant's initial position, which is a function of discrepancy (both positional and psychological). The former focuses more on how the external environment, as a function of message combination, has an effect on message scale value, whereas the latter focuses more on how a subject's belief before message receipt has an effect on message scale value.

The focus on how a subject's belief before message receipt has an effect on message scale value points to one specific area that has not been investigated enough in information integration theory: the valuation process. Information integration theory treats the valuation process as only a measurement issue and is mute on any theorizing about how the valuation occurs (Anderson, 2008, p. 340). The concept of psychological discrepancy is appropriate to address how valuation occurs. However, as mentioned above, the crucial problem of whether psychological discrepancy has an effect on the weight or on the scale value of a message has not been investigated. Indeed, Anderson recognized that both of these assumptions are possible. In discussing the role that the prior attitude of a subject plays in the valuation process, he stated, "Note that valuation constructs not just [scale value], but also [. . .] weight, which is qualitatively different but essential" (2008, p. 89).

The takeaway from the above statement is that weight and scale value have the same conceptual status, although it might be argued that scale value is conceptually superior to weight. In this regard, Himmelfarb (1975) made a strong argument that scale value is analogous to the concept of mass in Newton's laws. In Newton's laws, mass is held constant. Himmelfarb's argument indicates that, in information integration, the only parameter that is allowed to vary is weight, not scale value. However, Himmelfarb's assertion, although convincing, because it is reasonable to have scale value as a foundation for model building (Anderson, 2008, pp. 88-89), is purely hypothetical without any empirical support. The empirical work (Himmelfarb & Anderson, 1975) that was used as the support for the scale value constancy assumption does not provide conclusive evidence (see above). Second, Anderson's statement opens up the possibility that psychological discrepancy can have an effect on weight and scale value separately at the same time. This assumption is consistent with the complex model explicated above. That is, the scale value of a message used in integration is the result of the valuation process when psychological discrepancy constructs the subjective scale value. At the same time, the weight of the message is also constructed as a function of psychological discrepancy.

In sum, based on the framework of the discrepancy approach, especially the psychological-discrepancy-discounting model, this dissertation makes a theoretical contribution to explaining the valuation process in information integration. It also refines the psychological-discrepancy-discounting model by providing empirical evidence of how psychological discrepancy affects belief change.

CHAPTER 4

AN INITIAL COMPARISON OF MODELS USING NONLINEAR REGRESSION

Are there any available data that can be used to test the differences among the four models? This chapter presents the results of nonlinear regressions conducted on the weight-discounting model, the scale-value-pullback model (together with the independent-psychological-discrepancy model, as these two share the same functional form), and the complex model by using Fink et al.'s (1983) data. The method of Fink et al.'s (1983, pp. 420-422) study is described first.

Method of Fink Et Al.'s (1983) Study

Procedure and Sample

Questionnaires were distributed to undergraduate students in sociology classes at Michigan State University in a fall 1979 study (n = 193) and a spring 1980 study (n = 114, six months after the fall study). The students completed questionnaires in class. The participants read either one or two messages advocating a certain level of tuition increase. Then, the participants indicated how psychologically discrepant each message was and what percentage increase of tuition they believed would be reasonable.

Topic of Stimuli

In a pilot study, 63 undergraduate students in two communication classes at the same university were asked how important each topic was from a list of 21 topics.

Tuition increase was found to be the most important. Over half of the 63 students believed there should be zero percent of tuition increase (M = 2.07%, SD = 3.59%).

In the pilot study, the researchers also asked 43 undergraduate students how much percentage tuition increase was considered to be moderately, substantially, and extremely discrepant from their own views. Among the 43 students, 93% of them indicated that a 50% tuition increase was extremely discrepant from their own views. Thus, the level of 50% increase was chosen to be the advocated position in the extremely discrepant message. Also, a level of 15% increase was considered to be the most extremely moderate and the most moderately extreme position. Thus, the level of 15% increase was chosen to be the advocated position in the moderately discrepant message.[5]

Manipulation of Advocated Position

See the description in the section that introduces Fink et al.'s (1983) model in Chapter 2.

Measures

There were two dependent variables. For the variable of psychological discrepancy, after reading each message, the participants indicated how different the advocated position in a message was from their own views. The participants were told that a moderate level of difference was 100, the lowest level of difference was 0, and there was no upper bound. Then, for the variable of belief position after message receipt, the participants were asked "What is your opinion? That is, what percent do you think tuition at public institution of higher education should be increased next year?" (Fink et al., 1983, p. 422).

Nonlinear Regression Analysis

To be consistent with Fink et al.'s (1983) analysis, the data from the two studies were combined in the nonlinear regression analyses reported below. The effect of history was controlled for by including T (0 for the fall study and 1 for the spring study) in the regression model equations (see Table 3). The linear model and the psychological-discrepancy-weight-discounting model were tested in Fink et al.'s (1983) study. These two models were included in the current analysis for comparing them with the two models I propose in this dissertation: the psychological-discrepancy-scale-value-pullback model and the complex model. Untransformed values were used in the analyses.

The regression equations are shown in Table 3. For the linear models, *b*s were the coefficients in the linear regression. For the weight-discounting model (Equation 6), the scale-value-pullback model (together with the independent-psychological-discrepancy model; Equation 10), and the complex model (Equation 11), the independent weight of a message was assumed to be w for both the single and double message models. Then, for the nonlinear models (Equations 6, 10, and 11), w was divided from both the numerators and the denominators. The resulting forms were the regression equations for these models in Table 3. B_1 represented the estimated value of w_0/w. Therefore, the ratio of the two parameters, w_0 and w, could be estimated with a single parameter B_1. Note that the separate estimated values for w_0 and w were not the primary concerns here. The term, $B_2 + B_4T$, represented the scale value of participant's initial belief position, s_0, as a function of T. And, B_3 represented $-\gamma$. It was assumed that $\Delta(\psi)=e^{B_3(\psi + u)}$, where u was

Table 3

Regression Model Equations Using Fink Et Al.'s (1983) Data

	Single message
Linear model	$\widehat{R} = b_1 + b_2T + b_3s$
Weight-discounting model	$\widehat{R} = \dfrac{B_1(B_2 + B_4T) + se^{B_3(\psi+11.5)}}{B_1 + e^{B_3(\psi+11.5)}}$
Scale-value-pullback model	$\widehat{R} = \dfrac{B_1(B_2 + B_4T) + se^{B_3(\psi+11.5)}}{B_1 + 1}$
Complex model (restricted; $k' = 1$)	$\widehat{R} = \dfrac{B_1(B_2 + B_4T) + e^{B_3(\psi+11.5)}[s(1\text{-}e^{B_3(\psi+11.5)}) + (B_2 + B_4T)]}{B_1 + e^{B_3(\psi+11.5)}}$
Complex model (full)	$\widehat{R} = \dfrac{B_1(B_2 + B_4T) + e^{B_3(\psi+11.5)}[sB_5(1\text{-}e^{B_3(\psi+11.5)}) + (B_2 + B_4T)]}{B_1 + e^{B_3(\psi+11.5)}}$
	Double message
Linear model	$\widehat{R} = b_1 + b_2T + b_3s_1 + b_3s_2$
Weight-discounting model	$\widehat{R} = \dfrac{B_1(B_2 + B_4T) + s_1e^{B_3(\psi_1+11.5)} + s_2e^{B_3(\psi_2+11.5)}}{B_1 + e^{B_3(\psi_1+11.5)} + e^{B_3(\psi_2+11.5)}}$
Scale-value-pullback model	$\widehat{R} = \dfrac{B_1(B_2 + B_4T) + s_1\mathrm{e}^{B_3(\psi_1+11.5)} + s_2\mathrm{e}^{B_3(\psi_2+11.5)}}{B_1 + 1 + 1}$

Table 3

(continued)

	Double message
Complex model (restricted; $k' = 1$ for both messages)	$\widehat{R} = \frac{B_1(B_2+B_4T)+e^{B_3(\psi_1+11.5)}\left[s_1\left(1-e^{B_3(\psi_1+11.5)}\right)+(B_2+B_4T)\right]+e^{B_3(\psi_2+11.5)}\left[s_2\left(1-e^{B_3(\psi_2+11.5)}\right)+(B_2+B_4T)\right]}{B_1+e^{B_3(\psi_1+11.5)}+e^{B_3(\psi_2+11.5)}}$

Note. For the equations with a discounting factor $e^{B_3(\psi_1+11.5)}$, they were derived from dividing the independent weight of a message w (assumed to be equal across messages in the double-message condition) from the numerators and the denominators in the double-message versions of Equations 6, 10, and 11. The rationale was noted in the main text. This method was the same one used by Fink et al. (1983, p. 426). A symbol with a carat (^) is for a value without the error terms.

empirically determined to be 11.5 (Fink et al., 1983, p. 426n32). In the double-message condition, s_1 and s_2 represented the scale value of the first message and the second message, respectively; ψ_1 and ψ_2 represented the psychological discrepancy of the first message and the second message, respectively. For the complex model, at first, k' was assumed to be 1.00. Then, k' was allowed to vary to test the full complex model.

The results based on the data from the single-message condition are summarized in Table 4. For model evaluation and selection, the F test of explained variance and R^2_{adj} are not appropriate for a nonlinear regression (Spiess & Neumeyer, 2010), but the information is given here to give a sense of the model fits. AIC, a more appropriate index for model selection (Spiess & Neumeyer, 2010), is the Akaike information criterion. A lower AIC of an estimated model indicates less information loss of the estimated model from the true model, which means higher quality of the estimated model. A formula for calculating AIC based on least squares estimation can be found in Burnham and Anderson (2004, pp. 268-269). AIC_c is the AIC corrected for a small sample size (Burnham & Anderson, 2004, pp. 269-270). BIC is the Bayesian information criterion. Similar to AIC, the best model is the one that has the lowest BIC (Burnham & Anderson, 2004, p. 275). AIC, AIC_c, and BIC all penalize model complexity. Note that the analytic logic presented in Fink et al. (1983), where the amounts of the explained variance in different models were compared using the F test for comparing nested models, was no longer appropriate. The models in Table 3 were not nested. I used the Nonlinear Regression module of SPSS (2011) Version 20.0.0 to conduct the nonlinear regression analyses. For the scale-value-pullback model and the complex model,

Table 4

Analysis of Four Models for the Single-Message Condition Using Fink Et Al.'s (1983) Data

	Sum of squares	*df*	*F*	*p*	R^2_{adj}	SE_{est}[a]	*Sk*[b]	AIC	AIC_c	BIC
Total	2007.70	71								
Explained by linear model	144.01	2	2.67	.077	.045	5.20	.77	446.59	447.19	455.70
Explained by weight-discounting model	258.69	3	3.35	.024	.091	5.07	0.41	444.02	444.93	455.40
Explained by scale-value-pullback model*	260.31	3	3.38	.023	.092	5.07	0.43	443.95	444.86	455.33
Explained by complex model (k' = 1)*	465.76	3	6.85	< .001	.198	4.76	0.22	434.94	435.85	446.33
Explained by the full complex model*	477.08	4	5.22	.001	.193	4.78	0.20	436.41	437.71	450.07

Note. *n* = 72. AIC is the Akaike information criterion; AIC_c is the AIC corrected for a small sample size; BIC is the Bayesian information criterion.

*Data not published in Fink et al. (1983).

[a]SE_{est} is the standard error of the estimate, or the root mean squared error, which equals to $[\sum e^2/(n - v)]^{1/2}$, where *e* is the residual, and *v* is the number of estimated parameters.

[b]*Sk* is the skewness of residuals. The standard errors of the skewness all equaled to 0.28.

the start values for the parameters were chosen based on the estimated values from the original weight-discounting model (1983, Table 6, p. 428). For the test of the full complex model, two start values for k' (B_5) were tried. One is greater than 1.00 (2.00), and the other is less than 1.00 (0.50). Both start values reached the same optimal solution. For all nonlinear models, the constraint of $B_1 > 0$ was imposed.

The results showed that the restricted complex model (with $k' = 1$) fit the data the best according to AIC, AIC_c, and BIC. The full complex model came as a second close. With the same set of parameters, the restricted complex model (with $k' = 1$) fit the data better than the scale-value-pullback model and the weight-discounting model. Therefore, for the single-message condition, the restricted complex model with $k' = 1$, was the best performing model statistically.

The same analysis was conducted for the double-message condition. The results are summarized in Table 5. In contrast to the single-message condition, the restricted complex model, with $k' = 1$ for both messages, fit the data slightly worse than the scale-value-pullback model and the weight-discounting model according to AIC, AIC_c, and BIC. The less parsimonious complex models (by freeing k' for either of the messages or both of the messages) for the double-message condition fit worse than the restricted complex model reported above. Therefore, the results for those models are not reported in Table 5. Therefore, for the double-message condition, the conclusion was that the fully restricted complex model, with $k' = 1$ for both messages, the weight-discounting model, and the scale-value-pullback model performed almost equally well regarding model fit.

Table 5

Analysis of Four Models for the Double-Message Condition Using Fink Et Al.'s (1983) Data

	Sum of squares	*df*	*F*	*p*	R^2_{adj}	SE_{est}[a]	*Sk*[b]	AIC	AIC_c	BIC
Total	3994.93	105								
Explained by linear model	231.73	2	3.17	.046	.040	6.04	0.50	687.19	687.59	697.84
Explained by weight-discounting model	1156.67	3	13.86	< .001	.269	5.28	0.41	659.29	659.89	672.61
Explained by scale-value-pullback model*	1184.41	3	14.33	< .001	.275	5.25	0.44	658.25	658.85	671.57
Explained by complex model (k' = 1 for both messages)*	1145.89	3	13.67	< .001	.266	5.29	0.34	659.69	660.29	673.01

Note. n = 106. AIC is the Akaike information criterion; AIC_c is the AIC corrected for a small sample size; BIC is the Bayesian information criterion.

*Data not published in Fink et al. (1983).

[a]SE_{est} is the standard error of the estimate. See the note under Table 4.

[b]*Sk* is the skewness of residuals. The standard errors of the skewness all equaled to 0.24.

Discussion

For the single-message condition, the above results indicate that the restricted complex model, where psychological discrepancy has an effect on message weight and message scale value separately, is more plausible statistically than the weight-discounting model, the scale-value-pullback model, and the independent-psychological-discrepancy model. The restriction of $k' = 1$ in the complex model means that the perceived message scale value is always less than the explicit message scale value. In the full complex model tested for the single message condition, the estimated k' is less than 1, which also indicates that the perceived message scale value is less than the explicit message scale value. This empirical evidence is inconsistent with the social judgment approach's contrast effect in which an extreme message is perceived as even more extreme than the explicit message scale value *so that* it is rejected (Sherif & Hovland, 1961). The complex model proposed by this dissertation instead posits that the perceived message scale value moves closer to one's initial position as a function of psychological discrepancy. More importantly, the perceived scale value of a message with high psychological discrepancy is still further away from that of a message with low psychological discrepancy. The complex model has a major advantage over the scale-value-pullback model, because the complex model resolves a logical problem within the scale-value-pullback model.

However, the data from the double-message condition do not indicate that the complex model is statistically more plausible than the weight-discounting model or the scale-value-pullback model, even though the complex model indeed fit the data better than the linear model. It is plausible that the valuation process is more complex in the double message condition than in the single message condition, which may require considerable modification to the functional form in Equation 11, if the terms for a second message are included. One change in the model formulation is to consider what the reference point for positional and psychological discrepancy should be for the second message. For the first message, the reference point is simply subject's initial belief position. However, after the first message is received, the reference point for the positional and psychological discrepancy of the second message may be the result of information integration of the initial belief position and the advocated belief position in the first message, rather than the initial belief position alone. Future model formulation for Equation 11 for two or more messages should consider the distinction between two processes (Hogarth & Einhorn, 1992): the step-by-step process (i.e., where a reference point changes at each step of information integration) and the end-of-sequence process (i.e., where a reference point remains constant).

Overall, the nonlinear regression analyses suggest that more work needs to be done. This dissertation derives and tests competing hypotheses regarding the relationship between s (message scale value) and R (subject's final belief position) and the relationship between P (perspective) and R among the four models. The competing hypotheses can be tested with an experimental design in which P and s were manipulated independently. Also, in the experiment, the perceived message weight and the perceived message scale value can be directly measured (Eagly & Telaak, 1972; Zalinski & Anderson, 1989). By keeping the independent message weight, w, and the explicit scale value, s, constant, different levels of psychological discrepancy should lead to values of perceived weight and scale value in a way that is predicted by the most plausible model.

CHAPTER 5

HYPOTHESES

This chapter presents three approaches that aim to derive competing hypotheses among the four models. For the first approach, the algebraic expressions of the first partial derivatives of R with respect to both s and P were calculated using the *SymPy* package (Meurer et al, 2017) in the Python programming language, Version 2.7.14 (Python Software Foundation, 2017). The goal of calculating the first partial derivatives of R with respect to both s and P is to examine whether R changes differently across the four models as a function of s and as a function of P. For example, if the first partial derivative of R with respect to s based on a model is positive, this model predicts a positive relationship between R and s, ceteris paribus; if the first partial derivative of R with respect to P based on a model is initially positive and then becomes negative as P increases, this model predicts that as P increases, R first increases and then decreases, ceteris paribus. The sign of a first partial derivative was determined via either an analytic proof alone or a combination of an analytic proof and a computational approximation. Table 2 shows which relationship predicted by which model was determined with either an analytic proof alone or with a combination of an analytic proof and a computational approximation.

The second approach derived competing hypotheses regarding boomerang effects. The sign of the algebraic expression of the difference, $R - s_0$, was examined. For the third approach, if the perceived weight and the perceived scale value of a message by a subject can be measured, competing hypotheses can also be derived across the four models.

In this dissertation, the focus of testing the four models is on the single message condition and the condition of $s > s_0$. I will focus on the single-message condition to thoroughly understand the role of psychological discrepancy in belief change; therefore, I start with simpler models. As stated above, the model building for the condition of two or more messages needs to consider the processing manner (step-by-step vs. end-of-sequence; Hogarth & Einhorn, 1992). This requires a sufficient amount of revision in the functional forms of the four models. To confine the scope of this dissertation to a manageable one, the building and testing of models for the condition of two or more messages will be left for future studies. The reason to focus on the condition of $s > s_0$ is that by only considering this condition, competing hypotheses among the four models can be derived.

Differentiating the Weight-Discounting Model and the Scale-Value-Pullback Model

First, for the weight-discounting model and the scale-value-pullback model, derivation was performed in terms of the relationship between P and R. Derivation results were summarized in Table 6. The derivation shows that the two models predict the same relationship between P and R, where the symbol of P_H represents a high level of P, and P_L represents a low level of P:

H1a: The wider the perspective, the greater the final position.

Table 6

Summary of Derivation for H1a

Model	Model equation	Partial derivative of R with respect to P
Weight-discounting model	$R = \frac{w_0 s_0 + w\Delta(\psi)s}{w_0 + w\Delta(\psi)}$, where $\Delta(\psi) = e^{-\gamma\psi}$ and $\psi = kD/P$	$\frac{\partial R}{\partial P} = \frac{\gamma k D^2 e^{\gamma\psi} w w_0}{P^2(w_0^2 e^{2\gamma\psi} + 2ww_0 e^{\gamma\psi} + w^2)}$
Scale-value-pullback model ($s > s_0$)	$R = \frac{w_0 s_0 + w\Delta(\psi)s}{w_0 + w}$, where $\Delta(\psi) = e^{-\gamma\psi}$ and $\psi = kD/P$	$\frac{\partial R}{\partial P} = \frac{\gamma k s D w}{P^2 e^{\gamma\psi}(w_0 + w)}$
Weight-discounting model	Sign of $\frac{\partial R}{\partial P}$	$\frac{\partial R}{\partial P} > 0$
	Predicted relationship between R and P	$R_{P_H} > R_{P_L}$
Scale-value-pullback model ($s > s_0$)	Sign of $\frac{\partial R}{\partial P}$	$\frac{\partial R}{\partial P} > 0$
	Predicted relationship between R and P	$R_{P_H} > R_{P_L}$

Note. $s > s_0$; $P = U - L$; $D = s - s_0 > 0$; P_H represents a high level of P (i.e., low psychological discrepancy, ceteris paribus); P_L represents a low level of P (i.e., high psychological discrepancy, ceteris paribus). All parameters are assumed to be positive. The predicted relationship between R and P assumes that all the parameters except P are constants.

Second, what is the relationship between s and R? For the weight-discounting model, it has been shown that the relationship between s and R is nonmonotonic (with R first increasing and then decreasing, as s increases; see Appendix A). For the scale-value-pullback model, ceteris paribus, the relationship between s and R depends on the value of P (see Appendix B). Thus, the derivation shows that the two models predict the following relationships between s and R:

H2a (weight-discounting): As the message scale value increases, the final position initially increases, then decreases.

H2b (scale-value-pullback): In the wide perspective condition, as the message scale value increases, final position initially increases, then decreases; in the narrow perspective condition, as the message scale value increases, final position decreases.

Third, is a boomerang effect (i.e., $R < s_0$) possible in the two models? Derivation results (Appendix C) show that for the weight-discounting model, there is no boomerang effect (i.e., R is always greater than s_0); however, for the scale-value-pullback model, a boomerang effect is predicted when the value of s is large. Therefore, the hypotheses are:

H3a (weight-discounting): The mean of the final position of a subject's belief, R, is greater than the mean of the initial position of the same belief, s_0.

H3b (scale-value-pullback): The proportion of cases that have a boomerang effect in the most extreme message scale value condition is greater than the proportions in the less extreme message scale value conditions.

In addition to the above rationale for differentiating the two models, another approach is to measure the perceived importance and the perceived scale value of a message from the subjects (Eagly & Telaak, 1972; Zalinski & Anderson, 1989). The rationale is straightforward. Consider $w\Delta(\psi)s$: Holding w constant, if the weight of a message is discounted, and $w_p = w\Delta(\psi)$ is assumed to be the perceived weight of the message, then the value of w_p decreases as the level of psychological discrepancy increases. Meanwhile, holding the positional discrepancy constant, the perceived scale value of the message, s_p, should be constant as psychological discrepancy increases. However, if the scale value of a message is pulled back, and $s_p = s\Delta(\psi)$ is assumed to be the perceived scale value of the message, then the value of s_p decreases for $s > s_0$ and increases for $s < s_0$, as the level of psychological discrepancy increases. In addition, holding the positional discrepancy constant, w_p should be constant as psychological discrepancy increases.

A few words must be given here regarding the specific hypotheses regarding the relationship between P and w_p and the relationship between P and s_p. The way these hypotheses are stated takes the experimental design and the data analytic procedure of this dissertation into account. More details of the experiment will be given in later chapters, but for here it would be sufficient to let the reader know that the experiment was a between-subjects factorial 3 (levels of s) × 3 (levels of U) design, in which s_0 and L were kept constant across subjects. Therefore, by also measuring subject's s_0 and L, the empirical values of $D = |s - s_0|$ and $P = U - L$ for each subject could be obtained. A direct way to test, for example, $w_p = w\Delta(\psi) = we^{-\gamma\psi}$ (predicted by the weight-discounting model) is to linearize the relationship by taking the natural logarithm on both sides of the equation. After transforming the observed values of w_p, this linear regression model can be fitted: $\ln(w_p) = \ln(w) - \gamma\psi + \epsilon$, where ϵ is the error term. By taking the experimental manipulation and the multiplicative model of $\psi = kD/P$ into account, the final linear regression model fitted can be $\ln(w_p) = a + b_1 D + b_2(1/P) + b_3(D/\mathrm{P}) + \epsilon$ (see Blanton & Jaccard, 2006, for a recommended way to test this multiplicative model). If $w_p = w\Delta(\psi) = we^{-\gamma\psi}$ and $\psi = kD/P$ are plausible, the estimated b_1, b_2, and b_3 would all be negative, and b_3 would be significantly different from zero. A similar way can be used to test the relationship between P and s_p.

In what follows regarding H4 and H5, when I state that an independent variable predicts a dependent variable, I mean that there is a linear relationship between the independent variable and the dependent variable. Also, a no effect can be tested using equivalence testing (Weber & Popova, 2012; the details of equivalence testing will be presented in Chapter 7).

The psychological-discrepancy-weight-discounting model implies the following hypotheses regarding the relationship between P and w_p and the relationship between P and s_p:

H4a: The inverse of perspective negatively predicts the natural logarithm of perceived message weight when message discrepancy is at its sample mean. This hypothesis is about b_2 mentioned above. The conditional clause, when message discrepancy is at its sample mean, is included, because b_2 represents the simple effect of $1/P$ on $\ln(w_p)$ when the mean-centered D is zero.

H4b: The multiplicative term of message discrepancy and the inverse of perspective negatively and significantly predicts the natural logarithm of perceived message weight. This hypothesis is about b_3 mentioned above.

H5a: The inverse of perspective does not predict the natural logarithm of perceived message scale value.

H5b: The multiplicative term of message discrepancy and the inverse of perspective does not predict the perceived message scale value.

The psychological-discrepancy-scale-value-pullback model implies the following hypotheses regarding the relationship between P and w_p, and the relationship between P and s_p:

H4c: The inverse of perspective does not predict the natural logarithm of perceived message weight.

H4d: The multiplicative term of message discrepancy and the inverse of perspective does not predict the perceived message weight.

H5c: The inverse of perspective negatively predicts the natural logarithm of the perceived message scale value when message discrepancy is at its sample mean (see the explanation for H4a, above).

H5d: The multiplicative term of message discrepancy and the inverse of perspective negatively and significantly predicts the natural logarithm of perceived message scale value.

Differentiating the Independent-Psychological-Discrepancy Model and the Scale-Value-Pullback Model

Because the independent-psychological-discrepancy model and the scale-value-pullback model share the same functional form, the two models can only be differentiated based on the perceived message scale value. If the independent-psychological-discrepancy model is plausible, the empirical data should support the same hypotheses as the scale-value-pullback model except for H5c and H5d. The independent-psychological-discrepancy model predicts H5a and H5b instead.

Testing the Complex Model

If the complex model is plausible, H4a and H4b regarding the perceived message weight should be supported. And, regarding the perceived message scale value, the following hypothesis should be supported:

H5e: The inverse of perspective positively predicts the perceived message scale value when message discrepancy is at its sample mean (see the explanation for H4a, above).

H5f: The multiplicative term of message discrepancy and the inverse of perspective positively and significantly predicts the perceived message scale value.

Appendix D shows that the relationship between s and R is nonmonotonic when keeping P constant. Therefore, the complex model predicts H2a. Appendix E shows that the relationship between P and R is nonmonotonic by keeping positional discrepancy constant. Therefore, the complex model posits the following hypothesis:

H1b: As perspective increases, final position initially increases, then decreases.

The complex model also posits H2a and H3a.

Table 7 displays which model posits which hypothesis.

Table 7

Hypotheses Predicted by the Four Models

Competing hypotheses	Weight-discounting model	Scale-value-pullback model	Independent model	Complex model
H1a	×	×	×	
H1b				×
H2a	×			×
H2b		×	×	
H3a	×			×
H3b		×	×	
H4a	×			×
H4c		×	×	
H4b	×			×
H4d		×	×	
H5a	×		×	
H5c		×		
H5e				×
H5b	×		×	
H5d		×		
H5f				×

Note. A verbal statement of each hypothesis can be found in the main text (Chapter 5). The symbol × means that a specific hypothesis is predicted by a certain model. Each block separated by horizontal lines is composed of competing hypotheses.

CHAPTER 6

METHOD

This dissertation manipulated message scale value and perspective independently to examine the hypotheses stated in Chapter 5. An experiment that was a between-subjects 3 (high vs. moderate vs. low message scale value) × 3 (high vs. moderate vs. low *U*) factorial design was conducted. The manipulation of message scale value followed the method used in Kaplowitz and Fink (1991); the manipulation of upper bound was similar to the method used in Ostrom (1970). Participants were randomly assigned to one of the 3 × 3 = 9 conditions.

To have relatively consensual strength and position of an initial belief, the topic used by Kaplowitz and Fink (1991; see also Ostrom, 1970) was appropriate. In Kaplowitz and Fink's (1991) study, a pilot study was conducted in which college undergraduate students indicated how many years of imprisonment would be reasonable for several crimes. The crime that had the lowest standard deviation was selected. For this dissertation, armed robbery was used as the crime, because it had the lowest standard deviation in Kaplowitz and Fink's (1991) pilot study, and it was successfully used in two belief change studies testing discrepancy models (Chung et al., 2008; Kaplowitz & Fink, 1991).

Here is an overview of the experiment used here. The experiment recruited its participants through Amazon Mechanical Turk (MTurk hereafter). To save money, TurkPrime (Litman, Robinson, & Abberbock, 2017) was used to recruit MTurk workers in batches of nine, because Amazon charges an additional fee for a survey task that asks for 10 or more workers. Qualtrics (2018) was used to create and host the online questionnaires for the subjects to complete. On Qualtrics (2018), the participants were presented a description of a bogus armed robbery case in the U.S. in which the convict was called Convict X. Then the participants read a sentencing guideline that gave a recommended length of imprisonment (in years) for Convict X. This was an attempt to get everybody's s_0 fixed at a common value. Next, the participants were told a bogus maximum length of imprisonment for armed bank robbery in the past (i.e., upper bound manipulated). Then the participants read a sentencing decision made by a bogus judge that specified an actual length of imprisonment for Convict X (i.e., message scale value manipulated). Finally, the participants were asked to report an appropriate length of imprisonment (i.e., R measured) and their evaluations of the judge's decision (i.e., ψ, w_p, and s_p measured).

Before conducting the main study, six pilot studies were conducted to find out the subject's s_0, U, and L about sentencing Convict X (Pilot Study 1), the appropriate values of U to be used as the upper-bound manipulation (Pilot Studies 2 to 4), and the appropriate values of s to be used as the message-scale-value manipulation (Pilot Studies 5 to 6). The procedures and the results of the six pilot studies will be reported first. The details of the main study will then follow. The logarithm in the following text means the logarithm to the base of the constant, e (i.e., the natural logarithm). The significance level (alpha) is .05 for hypothesis testing (two-tailed).

In all pilot studies and the main study, all participants went through the following procedure. On Qualtrics (2018), after reviewing the IRB-approved consent form, the potential subjects were given the opportunity to either click the *I consent* button or the *I want to leave the study now* button. If a subject consented to proceed, he or she then read a brief overview of this study's purpose. The study was allegedly about the public beliefs towards the criminal justice system in the U.S. and the public opinions on a specific federal crime—armed bank robbery. The subjects then were asked six questions about their personal experience in the criminal justice system (see Appendix F), followed by the stimulus message and the relevant measures dependent upon the goal of a study (see the details in the section for each of the studies in this chapter). At the end of the questionnaire, the subjects were debriefed about the purpose of this dissertation. Then the subjects had the option to either allow or forbid me to use their recorded responses for the said purpose of this dissertation. At the very end of the questionnaire on Qualtrics (2018), a unique code was generated for a subject who had completed the survey. The subject then submitted this unique code to MTurk. Once I received this code, I checked on Qualtrics (2018) to make sure that this subject's responses were complete. If nothing seemed too abnormal (e.g., too many questions skipped), I approved that subject's work through TurkPrime, and the subject received the payment specified in the consent form.

The Qualtrics (2018) questionnaire for the main study is presented in Appendix F. This questionnaire for the main study includes all the materials used in the pilot studies. If the wording of a question was different in a pilot study from the main study, the question will be included in the section for this pilot study.

Data winsorization and transformation were used to deal with outliers and nonnormal distributions. Unless otherwise noted in the following sections (see also the section on Pilot Study 1 for a different way to handle outliers from the procedure outlined here), the winsorization and transformation were performed in the following procedure. If the skewness of a variable was significantly different from zero and is positive (commonly found in this dissertation's data), determined by comparing the ratio of the coefficient over its standard error against $z_{.025} = 1.96$, any values of that variable greater than the 95th percentile were recoded to the value at the 95th percentile (i.e., data winsorized). If the skewness was still significantly different from zero, a positively (common in this dissertation's data) skewed variable was transformed by taking its square, cube, sixth, twelfth root, and its natural logarithm when necessary (Fink, 2009). The logic of the procedure is that transforming a variable in the above sequence reduces the skewness gradually. Once a transformation results in a skewness that is not significantly different from zero, no further transformation is performed.

Pilot Studies

Pilot Study 1

The purpose of this pilot study was to collect two kinds of information that the main study needed for manipulating message scale value (see Kaplowitz & Fink, 1991) and upper bound (see Ostrom, 1970). The first kind of information was the subject's initial position on the topic (i.e., the most appropriate sentence in number of years for an armed bank robbery). The initial position was used as the recommended length of imprisonment in a fictitious sentencing guideline for all subjects in the main study. Also, via the manipulation of message scale value, the positional discrepancy levels could be manipulated according to the initial position. The second kind of information that the main study needed was the subjects' average initial U (the upper bound of the subject's belief position, i.e., the maximum harsh sentence in number of years for an armed bank robbery) and L (the lower bound of the subject's belief position, i.e., the maximum lenient sentence in number of years for an armed bank robbery). In the main study, to manipulate perspective, the harshest sentence from the previous trials (i.e., the U presented in the stimuli) was manipulated to have greater values than subject's initial U, where subject's initial L was manipulated and assumed to be constant.

Procedure

The survey was launched via TurkPrime (Litman et al., 2017) at 11:25 a.m., Eastern Standard Time, March 5, 2019. To be eligible to participate in the study, an MTurk worker must have had 5,000 or more approved MTurk tasks and a 98% or above approval rate of the completed tasks. These criteria were set to improve data quality (Peer, Vosgerau, & Acquisti, 2013). The intended sample size for the first pilot study was 36. The data collection process was completed at 4:28 p.m., Eastern Standard Time on the same day. By then, 36 completed responses had been recorded on Qualtrics (2018). All 36 subjects' responses were approved for compensation. Each of the 36 subjects was paid $0.40. After debriefing, two subjects refused to allow their responses to be used, so these two subjects were excluded from analysis. For the remaining 34 subjects, the average time to complete the survey on Qualtrics was 243.53 seconds (SD = 145.66), which equals 4.06 minutes. The shortest and the longest completion times were 95 seconds (1.58 minutes) and 847 seconds (14.12 minutes), respectively. The median was 200.00 seconds, which equals 3.33 minutes.

The survey asked the most appropriate sentence first. The subjects were asked, “What do you believe is the most appropriate sentence (in number of years) for the crime of armed bank robbery?” Then, the survey randomized the order of asking the subject’s *U* and *L*. For *U*, the subjects were asked, “What do you believe is the reasonable harshest sentence (in number of years) for a convict of armed bank robbery?” For *L*, “What do you believe is the most reasonable lenient sentence (in number of years) for a convict of armed bank robbery?” The survey also forced the subjects to give a number greater than or equal to the most appropriate sentence when asked about *U* and to give a number less than or equal to the most appropriate sentence when asked about *L*. No data were missing.

Results

The most appropriate sentence. The most appropriate sentence ranged from 2 to 35 years (M = 12.97, SD = 9.35, N = 34). The median was 10 years. The distribution of this variable had a skewness of 0.87 (SE = 0.40) and a kurtosis of -0.23 (SE = 0.79). The skewness was significantly different from zero (0.87 / 0.40 = 2.18 > $z_{.025}$ = 1.96). Thus, this variable was transformed by taking the logarithm of the original value. The distribution of the transformed variable had a skewness of -0.49 (SE = 0.40) and a kurtosis of -0.52 (SE = 0.79). The mean of the transformed variable was 2.26 (SD = 0.85). I untransformed the mean (i.e., 2.26) by exponentiating it. The result was 9.58. With the same method, the untransformed median was 10.00. The transformation decreased the discrepancy between the mean and median. It could be concluded that 10 years was an appropriate value to use in the sentencing guidelines in the second pilot study and in the main study.

The subject's U. In the raw data (N = 34), there were two extreme values (999,999 and 999 years). The third greatest value was 120 years. Therefore, the two extreme values were both recoded as 150 years. This value was chosen, although quite arbitrarily, to reflect a reasonable extremely great value of U. A systematic way to winsorize data at the 95th percentile was not followed until Pilot Study 2. This might be a flaw of Pilot Study 1.

The SPSS boxplot showed four outliers (100 years, 120 years, and two with 150 years) that were outside the 1.5 × the interquartile range (IBM, 2011, p. 700). Two of the four outliers were presented with the harshest-sentence question first, and two were presented with the-most-lenient-sentence question first. After these four outliers were excluded, the subject's U ranged from 2 to 50 years (M = 18.57, SD = 14.04, n = 30). The median was 15 years. The distribution of this variable had a skewness of 1.25 (SE = 0.43) and a kurtosis of 0.64 (SE = 0.83). Again, I reemphasize that a systematic way to winsorize data at the 95th percentile was not followed until Pilot Study 2. Excluding those four outliers might be another flaw of Pilot Study 1.

With the remaining 30 subjects, I took the logarithm of the original value. The distribution of the transformed variable had a skewness of -0.53 (SE = 0.43) and a kurtosis of 0.61 (SE = 0.83). The mean of the transformed variable was 2.64 (SD = 0.82). I untransformed the mean (i.e., 2.64) by exponentiating it. The result was 14.01. With the same method, the untransformed median was 15.00. After transformation, the discrepancy between the mean and the median was reduced. Therefore, it could be concluded that 15 years was an appropriate value to represent the upper bound of subject's belief position.

The subject's L. With the same 30 subjects that was analyzed for the subject's *U*, the values of the subject's *L* ranged from 0.5 to 30 years (M = 7.78, SD = 7.09, n = 30). The median was 5.5 years. The distribution of this variable had a skewness of 1.38 (SE = 0.43) and a kurtosis of 1.93 (SE = 0.83). A logarithmic transformation was done on the variable. The distribution of the transformed variable had a skewness of -0.32 (SE = 0.43) and a kurtosis of -0.72 (SE = 0.83). The mean of the transformed variable was 1.60 (SD = 1.05). I untransformed the mean (i.e., 1.60) by exponentiating it. The result was 4.95. With the same method, the untransformed median was 5.48.

The effect of two questions' order. For *U* (transformed), an independent-samples *t* test indicated that subjects were harsher when they were presented the question for *U* first (M = 3.00, SD = 0.66, n = 13) than when they were presented the question for *L* first (M = 2.36, SD = 0.83, n = 17), $t(28) = -2.27$, $p < .05$, d = 0.60. The Levene's test for equality of variances was not significant, $F(1, 28)$= 0.24, p = .63.

For *L* (transformed), an independent-samples *t* test indicated that subjects were more lenient, although not significantly at the .05 alpha level, when they were presented the question for *L* first (M = 1.31, SD = 0.97) than when they were presented the question for *U* first (M = 1.97, SD = 1.07), $t(28) = -1.74$, $p < .10$, d = 0.45. The Levene's test for equality of variances was not significant, $F(1, 28)$ = 0.10, p = .75.

Because the second pilot study and the main study will only manipulate the upper bound, U, the manipulation check will only include the question about the subject's U. The order effect indicated above should not be a major concern. However, the order effect suggests that if ever both questions for U and L are asked in the same questionnaire, one should always randomize the order of the two questions to account for any possible systematic variation in dependent variables due to order.

Discussion

Based on the results from the first pilot study, the second pilot study and the main study used 10 years as the recommend length of imprisonment in the bogus sentencing guidelines. In the main study, to manipulate positional discrepancy (via manipulating the message scale value and assuming an equal initial position across subjects), the three levels of sentence in the judge's decision could be 15, 23, and 35 years. The reason that these values were chosen was to make the increase in the message scale value at each step was an approximately equal ratio: $15/10 \approx 23/15 \approx 35/23$ (see Kaplowitz & Fink, 1991). Note that the value of the first level, 15 years, was chosen to make sure that the values of the second level (23 years) and the third level (35 years) were not too harsh. This consideration was due to a decision I made before conducting the second pilot study that all the values used for the upper bound manipulation (see the next paragraph) should be equal to or greater than the values used for the message scale value manipulation. The reason for this decision was that I wanted to avoid showing a subject a judge's decision (e.g., when $s = 35$) that was unprecedented (e.g., when $U = 23$) for all conditions.

The results of the first pilot study also showed that 15 years could be assumed to be most subjects' upper bounds. Thus, in the second pilot study and in the main study, to manipulate upper bound, the three levels of the harshest sentence in previous trials could be 35, 49, and 68 years, with $49/35 \approx 68/49$. The value of the first level, 35 years, was chosen in consideration of the decision I mentioned in the previous paragraph.

Pilot Study 2

This pilot study examined the values of the upper bound (U) that should be used to manipulate perspective ($P = U - L$). Following Ostrom's (1970) manipulation of perspective, perspective (P) was manipulated by fixing the value of L and only changing the value of U across the three levels of perspective. The second pilot study was a one-way four-level between-subjects design. For all subjects, the message scale value was fixed as 23 years. Three groups of subjects were told that the maximum years of the sentence for armed bank robbery from allegedly previous trials in the U.S. had been 35, 49, and 68 years, respectively, whereas the control subjects received the same stimulus message (see below) as the others, except that they were not given the maximum years of the sentence for armed bank robbery.

A stimulus message was composed of three parts. The first part presented a (bogus) U.S. federal sentencing guideline for armed bank robbery that explicitly states that an appropriate sentence was 10 years (see also Chung et al., 2008; Kaplowitz & Fink, 1991). The length of this sentence was determined based on the results of the first pilot study. The second part of the stimulus message presented different (bogus) values of maximum years of the sentence for armed bank robbery from allegedly previous trials in the U.S. (i.e., manipulating *U*), by which perspective was manipulated. The third part presented a (bogus) decision on Convict X (including a speech and the number of the years of the sentence) made by Judge Walters (a fictitious character who was described as a judge in a U.S. district court), in which the message scale value was presented (*s*; fixed as 23 years in the second pilot study for all subjects). To keep the message weight constant, the messages across different conditions were the same except for the value of the previous maximum sentence. The subjects were also told that the judge usually made his or her own decision regardless of what the sentencing guidelines said or what the sentences were in previous trials.

Procedure

The survey was launched via TurkPrime (Litman et al., 2017) at 10:30 a.m., Eastern Standard Time, April 4, 2019. To be eligible to participate in the study, an MTurk worker must have had 5,000 or more approved MTurk tasks and a 98% or above approval rate of the completed tasks. The intended sample size for the second pilot study was 99. The data collection process was completed at 7:28 a.m., Eastern Standard Time, April 6, 2019. By then, 98 completed responses had been recorded on Qualtrics (2018). One MTurk worker submitted the task without responding to the Qualtrics survey. This worker was excluded for analysis. All of the remaining 98 subjects' responses were approved for compensation. Each of the 98 subjects was paid $1.35. After debriefing, no subjects refused to allow their responses to be used. For the 98 subjects, the average time to complete the survey on Qualtrics was 1,076.70 seconds (*SD* = 622.26), which equals 17.95 minutes. The shortest and the longest completion times were 280.00 seconds (4.67

minutes) and 4,821.00 seconds (80.35 minutes), respectively. The median was 924.00 seconds, which equals 15.40 minutes.

Results

The subject's initial position. The question for the subject's initial position (s_0) was, "What do you believe is the most appropriate sentence (in number of years) for the crime of armed bank robbery?" The subjects' initial positions ranged from 0 to 40 years (M = 11.65, Mdn = 10.00, SD = 6.54, N = 98). The distribution of this variable had a skewness of 2.42 (SE = 0.24) and a kurtosis of 7.99 (SE = 0.48). The 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles were 5.00, 5.00, 10.00, 10.00, 12.75, 20.00, and 20.60. The five highest values were 40, 40, 35, 32, and 20 (multiple cases); the lowest values were 0, 2, 3, and 5 (multiple cases). Similar to the results of Kaplowitz and Fink's (1991, p. 197) study, 85.71 percent of subjects (84 out of 98) gave their responses between 5 and 15 years.

This variable was winsorized by recoding any values greater than 21 as 21 (the integer that was the closest the 95th percentile). After winsorization, subject's initial position ranged from 0 to 21 years (M = 11.01, Mdn = 10.00, SD = 4.41, N = 98). The distribution of this variable had a skewness of 0.55 (SE = 0.24) and a kurtosis of 0.49 (SE = 0.48). The skewness was still significantly different from zero ($0.55/0.24 = 2.29 > z_{.025} = 1.96$). Therefore, the winsorized variable was transformed by taking the square root, and the skewness and the kurtosis of the square root of the winsorized variable were -0.81 (SE = 0.24) and 3.63 (SE = 0.48), respectively. Taking the square root made the distribution even more different from normal distribution. Thus, the winsorized-only variable was used in the later stages of data analysis.

The subject's U. This was a crucial measure for the manipulation check in the second pilot study. The subjects were asked, "What do you believe is a reasonable *harshest* sentence (in number of years) for a convict of armed bank robbery?" Against my expectation, the pattern of group means did not reflect the pattern of the manipulated values. Tables 8 to 12 showed the group means and the skewness statistics after a series of attempts at data winsorization and transformation. A linear increase was expected in the group means from the control group to Group 3. However, not considering the raw data (the first column in Tables 8 to 12), a common pattern across different winsorizations and transformations was that there was always a dip of group means at Group 2.

The next question was whether the dip was significant in an inferential test. Excluding the control group, a one-way ANCOVA on the subject's upper bound with the condition being the predictor and the initial position being the covariate was conducted. The dependent variable was based on the data in the last column in Table 12. The unadjusted group means were 3.33 (SD = 0.48, n = 24), 3.18 (SD = 0.53, n = 24), and 3.61 (SD = 0.66, n = 22) for the 35-, 49-, and 68-year sentencing group, respectively. The adjusted group means accounting for the subject's initial position were 3.28 (SE = 0.10), 3.25 (SE = 0.10), and 3.60 (SE = 0.10), respectively. Again, the pattern of these adjusted means reflected a similar one to that of the unadjusted group means. The Levene's test of equality of error variances was significant, $F(2, 67) = 3.17$, $p = .048$.

Table 8

Pilot Study 2: Group Means and Skewness Coefficients of Subject's Upper Bound With the Raw Data, the Winsorized Data, and the Transformed Winsorized Data 1

	Raw data	Winsorized data	Square root of winsorized data	Cube root of winsorized data	Sixth root of winsorized data	Natural logarithm of winsorized data
Control $n = 28$	53.320 / 5.281	21.038 / 3.904	4.521 / 2.559	2.714 / 2.035	1.642 / 1.510	2.956 / 1.002
Group 1 $n = 24$	30.210 / 0.353	30.208 / 0.353	5.442 / 0.168	3.075 / 0.094	1.748 / 0.013	3.334 / -0.076
Group 2 $n = 24$	26.500 / 2.603	26.292 / 2.428	4.988 / 0.820	2.887 / 0.125	1.689 / -0.635	3.106 / -1.426
Group 3 $n = 22$	44.180 / 0.711	43.773 / 0.644	6.362 / 0.236	3.392 / 0.041	1.830 / -0.199	3.586 / -0.491

Note. The values before / are the group means, and the values after / are the skewness statistics. Data winsorized at $U = 95$.

Table 9

Pilot Study 2: Group Means and Skewness Coefficients of Subject's Upper Bound With the Raw Data, the Winsorized Data, and the Transformed Winsorized Data 2

	Raw data	Winsorized data	Square root of winsorized data	Cube root of winsorized data	Sixth root of winsorized data	Natural logarithm of winsorized data
Control n = 28	53.320 / 5.281	20.857 / 3.753	4.512 / 2.408	2.711 / 1.900	1.641 / 1.395	2.954 / 0.910
Group 1 n = 24	30.210 / 0.353	30.208 / 0.353	5.442 / 0.168	3.075 / 0.094	1.748 / 0.013	3.334 / -0.076
Group 2 n = 24	26.500 / 2.603	26.083 / 2.240	4.977 / 0.682	2.884 / 0.012	1.688 / -0.722	3.104 / -1.488
Group 3 n = 22	44.180 / 0.711	43.318 / 0.579	6.338 / 0.188	3.384 / -0.003	1.828 / -0.238	3.582 / -0.526

Note. The values before / are the group means, and the values after / are the skewness statistics. Data winsorized at U = 90.

Table 10

Pilot Study 2: Group Means and Skewness Coefficients of Subject's Upper Bound With the Raw Data, the Winsorized Data, and the Transformed Winsorized Data 3

	Raw data	Winsorized data	Square root of winsorized data	Cube root of winsorized data	Sixth root of winsorized data	Natural logarithm of winsorized data
Control $n = 28$	53.320 / 5.281	20.964 / 3.824	4.530 / 2.598	2.720 / 2.146	1.644 / 1.703	2.966 / 1.286
Group 1 $n = 24$	30.210 / 0.353	30.208 / 0.353	5.442 / 0.168	3.075 / 0.094	1.748 / 0.013	3.334 / -0.076
Group 2 $n = 24$	26.500 / 2.603	26.458 / 2.469	5.057 / 1.442	2.924 / 1.084	1.704 / 0.733	3.175 / 0.397
Group 3 $n = 22$	44.180 / 0.711	43.500 / 0.615	6.371 / 0.321	3.400 / 0.199	1.834 / 0.062	3.605 / -0.093

Note. The values before / are the group means, and the values after / are the skewness statistics. Data winsorized at $U = 90$ and $U = 9$.

Table 11

Pilot Study 2: Group Means and Skewness Coefficients of Subject's Upper Bound With the Raw Data, the Winsorized Data, and the Transformed Winsorized Data 4

	Raw data	Winsorized data	Square root of winsorized data	Cube root of winsorized data	Sixth root of winsorized data	Natural logarithm of winsorized data
Control *n* = 28	53.320 / 5.281	21.143 / 3.969	4.539 / 2.744	2.723 / 2.277	1.645 / 1.816	2.968 / 1.378
Group 1 *n* = 24	30.210 / 0.353	30.208 / 0.353	5.442 / 0.168	3.075 / 0.094	1.748 / 0.013	3.334 / -0.076
Group 2 *n* = 24	26.500 / 2.603	26.667 / 2.647	5.068 / 1.573	2.927 / 1.195	1.705 / 0.823	3.178 / 0.469
Group 3 *n* = 22	44.180 / 0.711	43.955 / 0.679	6.394 / 0.366	3.407 / 0.239	1.835 / 0.096	3.610 / -0.063

Note. The values before / are the group means, and the values after / are the skewness statistics. Data winsorized at $U = 95$ and $U = 9$.

Table 12

Pilot Study 2: Group Means and Skewness Coefficients of Subject's Upper Bound With the Raw Data, the Winsorized Data, and the Transformed Winsorized Data 5

	Raw data	Winsorized data	Square root of winsorized data	Cube root of winsorized data	Sixth root of winsorized data	Natural logarithm of winsorized data
Control n = 28	53.320 / 5.281	21.321 / 4.095	4.548 / 2.877	2.726 / 2.399	1.646 / 1.921	2.970 / 1.464
Group 1 n = 24	30.210 / 0.353	30.208 / 0.353	5.442 / 0.168	3.075 / 0.094	1.748 / 0.013	3.334 / -0.076
Group 2 n = 24	26.500 / 2.603	26.875 / 2.813	5.078 / 1.699	2.931 / 1.302	1.705 / 0.911	3.180 / 0.537
Group 3 n = 22	44.180 / 0.711	44.364 / 0.745	6.415 / 0.410	3.413 / 0.277	1.837 / 0.129	3.614 / -0.034

Note. The values before / are the group means, and the values after / are the skewness statistics. Data winsorized at U = 100 and U = 9.

Adjusted for the subject's initial position, the group means did differ significantly, $F(2, 66) = 3.63$, $p = .032$, $\eta^2 = .099$. The linear trend was significant, $F(1, 66) = 5.05$, $p = .028$, $\eta^2 = .071$. The quadratic trend was not significant, $F(1, 66) = 2.33$, $p = .132$, $\eta^2 = .034$.

Although the results of the inferential test looked promising, an alarming fact was that Group 2's mean still dipped even after adjusting for the subject's initial position. The significant linear trend shown in the last paragraph was just an artifact of Group 3's mean being much greater than the means of Group 1 and Group 2, whereas the means of Group 1 and Group 2 were too close to each other. The manipulation in the second pilot study was problematic.

Psychological discrepancy. To further answer the question of whether the manipulation worked, I examined psychological discrepancy (see the question for measuring this variable in the section for the main study's method and in Appendix F). The expected result was that psychological discrepancy (ψ) was a decreasing function of upper bound (U) according to the relationship expressed in this equation: $\psi = kD/(U - L)$. Across all subjects, the subject's psychological discrepancy ranged from 0 to 500 ($M = 90.11$, $Mdn = 50.00$, $SD = 99.96$, $N = 98$). This variable had a skewness of 1.99 ($SE = 0.24$) and a kurtosis of 4.74 ($SE = 0.48$). After winsorizing this variable at the 95th percentile (300), the winsorized variable had a skewness of 1.23 ($SE = 0.24$) and a kurtosis of 0.59 ($SE = 0.48$). The skewness was still significantly different from zero ($1.23/0.24 = 5.13 < z_{.025} = 1.96$). Therefore, the winsorized variable was transformed by taking its square root, and the skewness and the kurtosis of the square root of the winsorized variable were 0.40 (*SE* = 0.24) and -0.62 (*SE* = 0.48), respectively. Both of the skewness and the kurtosis were now not significantly different from zero. Therefore, the square root of the winsorized psychological discrepancy was used below.

Excluding the control group, a one-way ANCOVA on the subject's psychological discrepancy with the condition being the predictor and the subject's initial position being the covariate was conducted. The unadjusted group means were 8.34 (*SD* = 4.77, *n* = 24), 8.88 (*SD* = 4.84, *n* = 24), and 6.63 (*SD* = 3.28, *n* = 22) for the 35-, 49-, and 68-year sentencing group, respectively. The adjusted group means accounting for the subject's initial position were 8.60 (*SE* = 0.86), 8.54 (*SE* = 0.86), and 6.72 (*SE* = 0.89), respectively. The Levene's test of equality of error variances was not significant, $F(2, 67) = 0.67$, $p = .515$. Adjusted for the subject's initial position, the group means did not differ significantly, $F(2, 66) = 1.48$, $p = .236$, $\eta^2 = .043$. The linear trend was not significant, $F(1, 66) = 2.32$, $p = .132$, $\eta^2 = .034$. The quadratic trend was not significant either, $F(1, 66) = 0.68$, $p = .414$, $\eta^2 = .010$. Again, these results indicated that the manipulation in Pilot Study 2 was problematic.

Discussion

The second pilot study used 35, and 49, and 68 years as the maximum sentences in previous trials of the same crime. Using the subject's upper bound for checking the manipulation, the results showed that there was a puzzling dip in the group means at Group 2 (i.e., the 49-year sentencing group), whereas a linear increase in the group means from Group 1 to Group 3 was expected. Controlling for the subject's initial position, the linear increase was indeed significant, and the quadratic trend was not significant.

However, I could hardly conclude that the manipulation was successful, because even the group means adjusted for initial position had a small dip at Group 2. The results became even less promising by using psychological discrepancy for checking the manipulation. Controlling for the subject's initial position, the linear decrease was not significant, nor was the quadratic trend. The only positive sign was that there was a hump at Group 2 in the unadjusted group means, which was consistent with the assumed negative relationship between upper bound and psychological discrepancy. All of the above results pointed to a priority at this point of this dissertation: Adjust the numbers used for the manipulation of upper bound. Once the subject's upper bound showed the expected linear increase in group means, it was highly possible that psychological discrepancy will show the expected linear decrease.

Pilot Study 3

This pilot study further examined the values of the upper bound (*U*) that would be used to manipulate perspective in the main study. For all subjects, the message scale value was fixed at 23 years. Three groups of subjects were told that the maximum number of years of the sentence for armed bank robbery from allegedly previous trials in the U.S. had been 35, 50, and 70 years, respectively, whereas the control subjects received the same stimulus message as the others, except that they were not given the maximum number of years of the sentence for armed bank robbery. The questionnaire for Pilot Study 3 was the same as the one used in Pilot Study 2 except for the different values of the upper-bound manipulation, an additional check question on whether a subject processed the information about *U*, and the omission of the measures of w_p and s_p to save time and money.

Procedure

The survey for Pilot Study 3 was launched via TurkPrime (Litman et al., 2017) at 10:14 a.m., Eastern Standard Time, May 15, 2019. To be eligible to participate in the study, an MTurk worker must have had 5,000 or more approved MTurk tasks and a 98% or above approval rate of the completed tasks. The intended sample size for the third pilot study was 99. The data collection process was completed at 10:04 p.m., Eastern Standard Time, May 16, 2019. By then, 99 completed responses had been recorded on Qualtrics (2018). Each of the 99 subjects was paid $0.80. After debriefing, no subjects refused to allow their responses to be used. For the 99 subjects, the average time to complete the survey on Qualtrics was 963.37 seconds (*SD* = 750.05), which equals 16.06 minutes. The shortest and the longest completion times were 356.00 seconds (5.93 minutes) and 5,094.00 seconds (84.90 minutes), respectively. The median was 706.00 seconds, which equals 11.77 minutes.

Results

Right before the measure of subject's *U*, the non-control participants were asked, "What was the documented harshest sentence (in years) for the crime of armed bank robbery in the United States?" This question checked whether a subject had processed the information of the maximum years of sentence (either 35, 50, or 70 years). Except for one subject in the 50-year group who answered 10 years, all of the subjects gave the correct answer. The subject who gave the incorrect answer was excluded in the following analysis.

For the subject's U, the subjects were asked, "What do you believe is a reasonable *harshest* sentence (in number of years) for a convict of armed bank robbery?" Across all subjects, the subject's U ranged from 5 to 999 years (M = 44.46, Mdn = 25.00, SD = 100.78, n = 98). The distribution of this variable had a skewness of 8.97 (SE = 0.24) and a kurtosis of 85.21 (SE = 0.48). The 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles were 11.90, 15.00, 20.00, 25.00, 50.00, 60.00, and 99.05. This variable was winsorized at the 5th percentile and the 95th percentile. After winsorization, the subject's U ranged from 12 to 99 years (M = 34.28, Mdn = 25.00, SD = 22.04, n = 98). The distribution of this variable had a skewness of 1.55 (SE = 0.24) and a kurtosis of 2.15 (SE = 0.48).

Next, the winsorized variable was transformed in different ways. The group means and skewness statistics after each transformation are presented in Table 13. The log transformation brought all groups' skewness statistics down to a value below one (see the first column from the right in Table 13). The group means indicated that the manipulation of subject's upper bound was only partially successful at best. The major problem was that the means of the 35-year group and the 50-year group were almost the same, which was a similar problem found in Pilot Study 2. Figure 3 shows the boxplots by conditions.

Based on the log-transformed winsorized data, an omnibus one-way ANOVA revealed that the control (M = 3.10, SD = 0.37, n = 27), 35-year (M = 3.28, SD = 0.57, n = 24), 50-year (M = 3.29, SD = 0.54, n = 22), and 70-year (M = 3.80, SD = 0.56, n = 25)

Table 13

Pilot Study 3: Group Means and Skewness Coefficients of Subject's Upper Bound With the Raw Data, the Winsorized Data, and the Transformed Winsorized Data

	Raw data	Winsorized data	Square root of winsorized data	Cube root of winsorized data	Sixth root of winsorized data	Natural logarithm of winsorized data
Control *n* = 27	23.74 / 1.51	23.74 / 1.51	4.89 / 1.09	2.87 / 0.95	1.68 / 0.78	3.10 / 0.63
Group 1 *n* = 24	68.96 / 4.84	31.54 / 2.18	5.46 / 1.52	3.07 / 1.28	1.73 / 1.01	3.28 / 0.77
Group 2 *n* = 22	29.91 / 0.18	30.50 / 0.32	5.45 / 0.16	3.07 / 0.10	1.74 / 0.02	3.29 / -0.06
Group 3 *n* = 25	56.12 / 1.58	51.60 / 0.60	7.03 / 0.16	3.64 / 0 .003	1.89 / -0.17	3.80 / -0.32

Note. In each cell, the value before the virgule is the group mean, and the value after the virgule is the skewness. The maximum years of sentence were 35, 50, and 70 in Group 1, Group 2, and Group 3, respectively. Data winsorized at $U = 99$ and $U = 12$.

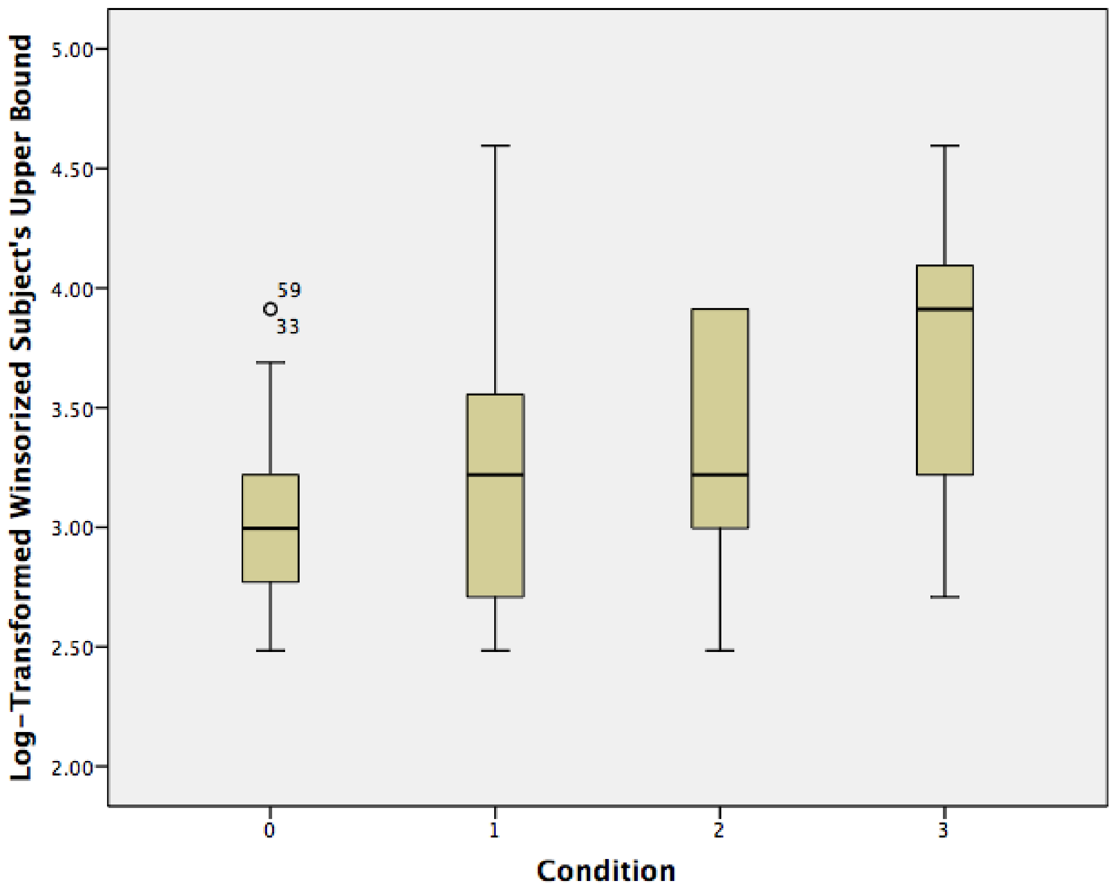


Figure 3. Pilot Study 3: The result of the manipulation check on subject's upper bound. The maximum sentence was 35, 50, and 70 years in Group 1, Group 2, and Group 3, respectively. Omnibus one-way ANOVA: $F(3, 94) = 9.07$, $p < .001$, $\eta^2 = .225$. Sample size by condition: $n = 27$ (control), $n = 24$ ($U = 35$), $n = 22$ ($U = 50$), and $n = 25$ ($U = 70$).

groups differed significantly, $F(3, 94) = 9.07$, $p < .001$, $\eta^2 = .225$. The Levene's test for equality of variances was not significant, $F(3, 94) = 2.06$, $p = .111$. Planned contrasts indicated, as expected, a significant linear increase in subject's upper bound, $F(1, 94) = 22.32$, $p < .001$, but no significant quadratic, $F(1, 94) = 2.70$, $p = .104$, or cubic, $F(1, 94) = 2.06$, $p = .154$, effect. However, a post hoc Tukey test comparing all pairs of group means revealed more severe problems. Among the six pairs of group mean differences, three involving the 70-year group were significant, whereas the remaining three differences were not significant. In other words, the control, 35-year, and 50-year groups did not significantly differ from each other, whereas the 70-year group had a significantly greater upper bound than each of the other three groups.

I tried removing the systematic variation due to the subject's initial position by regressing the log-transformed winsorized subject's *U* on the log-transformed winsorized subject's initial position. The regression model's R^2 was .406. The boxplots of the unstandardized residuals across conditions are shown in Figure 4. The pattern was similar to the one in Figure 3. I also tried examining the possible confounding effects of the six binary variables related to one's criminal justice experience (see Appendix F for the questions). I ran a series of two-way ANOVAs with one predictor being the condition and the other predictor being one of the six binary variables. Any significant interaction would be a sign worth further examination. None of the interactions was significant.

Discussion

The primary goal Pilot Study 3 was to examine if the manipulation of upper bound was effective. The key measure for the manipulation check was subject's upper

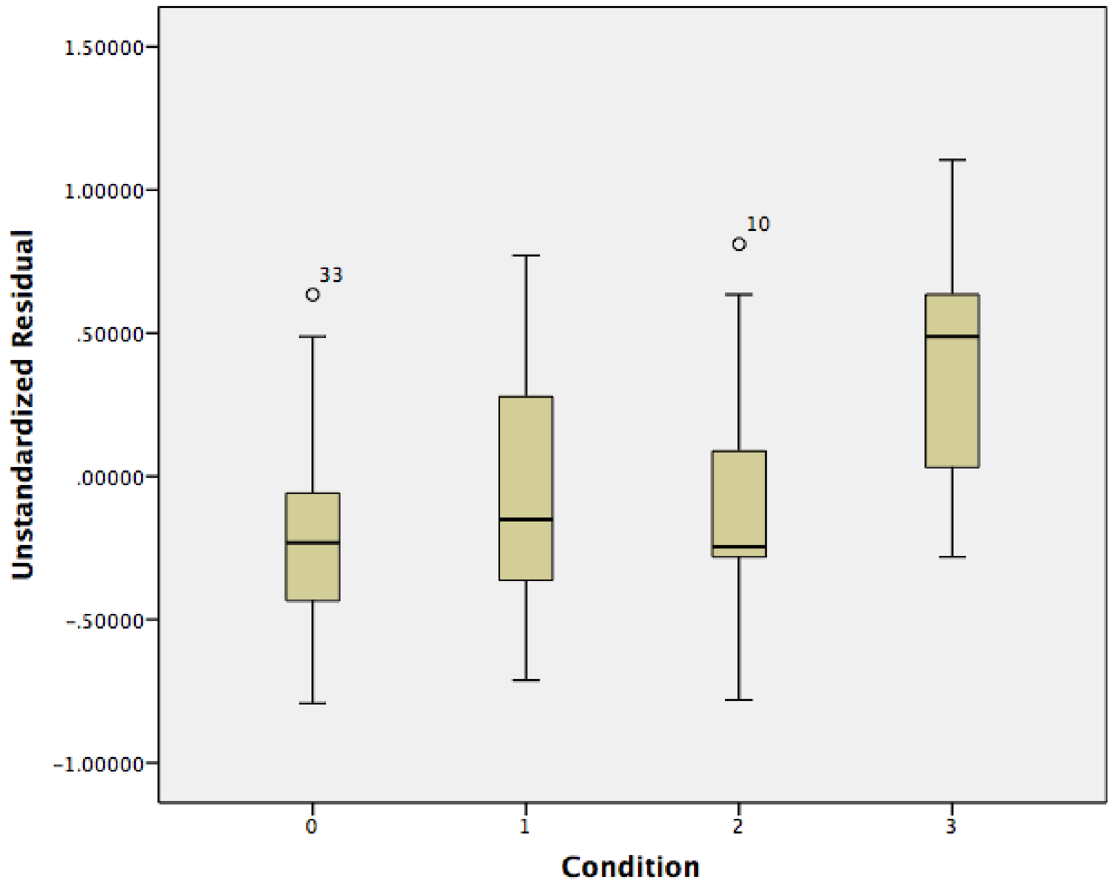


Figure 4. Pilot Study 3: Condition's effect on subject's upper bound after the systematic variation due to subject's initial position was removed. The maximum sentence was 35, 50, and 70 years in Group 1, Group 2, and Group 3, respectively. Sample size by condition: $n = 27$ (control), $n = 24$ ($U = 35$), $n = 22$ ($U = 50$), and $n = 25$ ($U = 70$).

bound. For this variable, the result showed a similar problem found in Pilot Study 2: The mean of Group 2 (the 50-year sentencing group) was not significantly different from the mean of Group 1 (the 35-year sentencing group).

In Pilot Study 2, the maximum number of years of the sentence were 49 and 68 in Group 2 and Group 3, respectively. When I compared the second column in Table 13 (for Pilot Study 3) and the second column in Table 12 (for Pilot Study 2), I found that using 50-years instead of 49-years and using 70-years instead of 68-years did increase the means of Group 2 and Group 3. For Group 2, the group mean increased from 26.88 in Pilot Study 2 to 30.50 in Pilot Study 3; for Group 3, the group mean increase from 44.36 in Pilot Study 2 to 51.60 in Pilot Study 3. However, Group 2's mean was still not large enough to be significantly greater than Group 1's mean.

Ostrom's (1970) result in his Table 1 (p. 285) provided empirical support for the assumed effectiveness of the perspective manipulation used in this dissertation. Ostrom (1970) presented a bomb threat case to his subjects, told them "the range of prison sentences provided by law for individuals convicted of this particular crime" (p. 283) was either 1 to 5 years (the narrow-perspective condition) or 1 to 30 years (the wide-perspective condition), and asked them to indicate an appropriate sentence in years for each of the nine categories from maximally lenient to maximally stern. The manipulation was successful. At the maximally-stern end, the wide-perspective subjects reported an average of 22.9 years, whereas the narrow-perspective subjects reported an average of 7.3 years. For the wide-perspective condition, the ratio between the *achieved* upper bound (22.9 years) and the *intended* upper bound (30 years) was 0.76. One could treat this ratio as a benchmark level of effectiveness of the perspective manipulation. In Pilot Study 3, using the second column in Table 13, Group 1's ratio was 31.54 / 35 = 0.90, which was greater than Ostrom's; Group 3's ratio was 0.74, which was similar to Ostrom's; Group 2's ratio was 0.61, which was the lowest among the three groups.

The puzzle was why only Group 2's manipulation was relatively not successful. The answer was still unclear. However, perhaps the problem was not in Group 2 per se. Table 14 showed that the control group had the smallest standard error and the smallest standard error of the mean. In other words, when the specific maximum sentence information was presented to the subjects in Groups 1, 2, and 3, the dispersions of subject's upper bound were greater than the dispersion of the control subjects to whom the maximum sentence information was not given. The greater dispersion for the subjects in Groups 1, 2, and 3 might be because that the maximum sentence might not have been salient enough to induce a committed change.

To elaborate, in Pilot Study 3, the maximum sentence was framed as the documented longest sentence in the U.S. history. Perhaps some subjects thought the trial from which the longest sentence was given was just an outlier. Therefore, when a subject was asked to give a *reasonable* harshest sentence, he or she might have ignored the maximum sentence information because an outlier *might not seem reasonable*. Note that Ostrom (1970) told his subjects the maximally stern sentence was given *by law*. One remedy could be to make the subjects more committed to the information regarding maximum sentence.

Table 14

Pilot Study 3: Group Ranges and Dispersions of Subject's Upper Bound of the Winsorized Data

	Range	Standard deviation	Standard error of the mean
Control $n = 27$	[12, 50]	9.86	1.90
Group 1 $n = 24$	[12, 99]	22.56	4.60
Group 2 $n = 22$	[12, 50]	15.42	3.29
Group 3 $n = 25$	[15, 99]	26.54	5.31

Note. The maximum years of sentence were 35, 49, and 68 in Group 1, Group 2, and Group 3, respectively. Data winsorized at $U = 99$ and $U = 12$.

Pilot Study 4

The first goal of Pilot Study 4 was to find three appropriate values of the upper-bound (U) manipulation in the main study. Secondly, Pilot Study 4 also examined whether the induced change in upper bound would lead to a change in psychological discrepancy (ψ). Holding positional discrepancy (D) constant, a successful manipulation of perspective ($P = U - L$) should find the following relationship: $\psi_{P_H} < \psi_{P_M} < \psi_{P_L}$ according to $\psi = kD/(U - L)$, where P_H, P_M, and P_L represent a wide, a moderate, and a narrow perspective, respectively. The greater the value of P, and the less the value of ψ.

Pilot Study 4 was a one-way five-level between-subjects design including a control group. For all subjects, the message scale value presented as a judge's sentence was fixed at 23 years. Four groups of subjects were told that the maximum years of the sentence for armed bank robbery from allegedly previous trials in the U.S. had been 12, 18, 30, and 50 years, respectively, whereas the control subjects received the same stimulus message (see below) as the others, except that they were not given the maximum sentence. Because Pilot Study 1 found that without being given the maximum sentence the subject's upper bound was about 15 years, the 12-year group should decrease the subject's upper bound, whereas the other three non-control groups should increase it.

In contrast to Pilot Study 2 and Pilot Study 3, the specific details of an armed bank robbery case were given together with the 10-year sentencing guideline. In other words, the suggested 10-year sentence was presented as a guideline that applied to *a specific armed bank robbery case* in which the convict took more than \$50,000, brandished a firearm, and was a repeat offender, whereas previously the guideline was only provided for *the more general crime category of armed bank robbery*. To let the participants commit more to a change in upper bound, the next part of the stimulus message began with a paragraph stating that judges were allowed to make their own decisions regardless of what the sentencing guidelines said (see Dillehay & Berger, 1969, on the enlarging effect of *permissive introduction* on message judgment). Then the subjects were presented with different maximum years of the sentence for similar cases of armed bank robbery from allegedly previous trials in the U.S.

Procedure

The survey was launched via TurkPrime (Litman et al., 2017) at 07:32 p.m., Eastern Standard Time, May 31, 2019. To be eligible to participate in the study, an MTurk worker must have had 5,000 or more approved MTurk tasks and a 98% or above approval rate of the completed tasks. The intended sample size for the third pilot study was 125. The data collection process was completed at 02:24 a.m., Eastern Standard Time, June 2, 2019. By then, 125 completed responses had been recorded on Qualtrics (2018). Each of the 125 subjects was paid $0.80. After debriefing, no subjects refused to allow their responses to be used. For the 125 subjects, the average time to complete the survey on Qualtrics (2018) was 720.79 seconds (SD = 411.75), which equals 12.01 minutes. The shortest and the longest completion times were 272.00 seconds (4.53 minutes) and 3,416.00 seconds (56.93 minutes), respectively. The median was 619.00 seconds, which equals 10.32 minutes.

Results

The subject's U. The subjects were asked, "What do you believe is a *maximally harshest* sentence (in number of years) for a convict of armed bank robbery?" Across all subjects, the subject's U ranged from 2 to 50 years ($M = 21.16$, $Mdn = 20.00$, $SD = 10.64$, $N = 125$). The distribution of this variable had a skewness of 1.28 ($SE = 0.22$) and a kurtosis of 1.64 ($SE = 0.43$). The 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles were 8.60, 10.00, 15.00, 20.00, 25.00, 32.00, and 50.00. Next, this variable was transformed by taking the cube root. The distribution of the transformed variable had a skewness of 0.18 ($SE = 0.22$) and a kurtosis of 0.79 ($SE = 0.43$). According to the transformed data, the group means indicated that the manipulation of subject's upper bound was successful (Figure 5). Based on Pilot Study 1, the control group's mean should have been around 15 years. However, the control group's mean (untransformed) in the Pilot Study 4 was 19.78 years, which indicated that both the 12-year maximum sentence and the 18-year maximum sentence in effect contracted the subject's upper bound. The means of these two groups indeed were less than the control group. The reason why the 18-year group was lower than the 12-year group is unknown. It is possible that a maximum sentence of 12 years seemed too lenient, and the subjects found it not credible enough and discounted the information.

To examine the possible order effect of the measures of the maximum and the minimum sentence, a two-way ANOVA on the subject's upper bound with both condition and order as the predictors was conducted. The only significant effect was the main effect for condition.

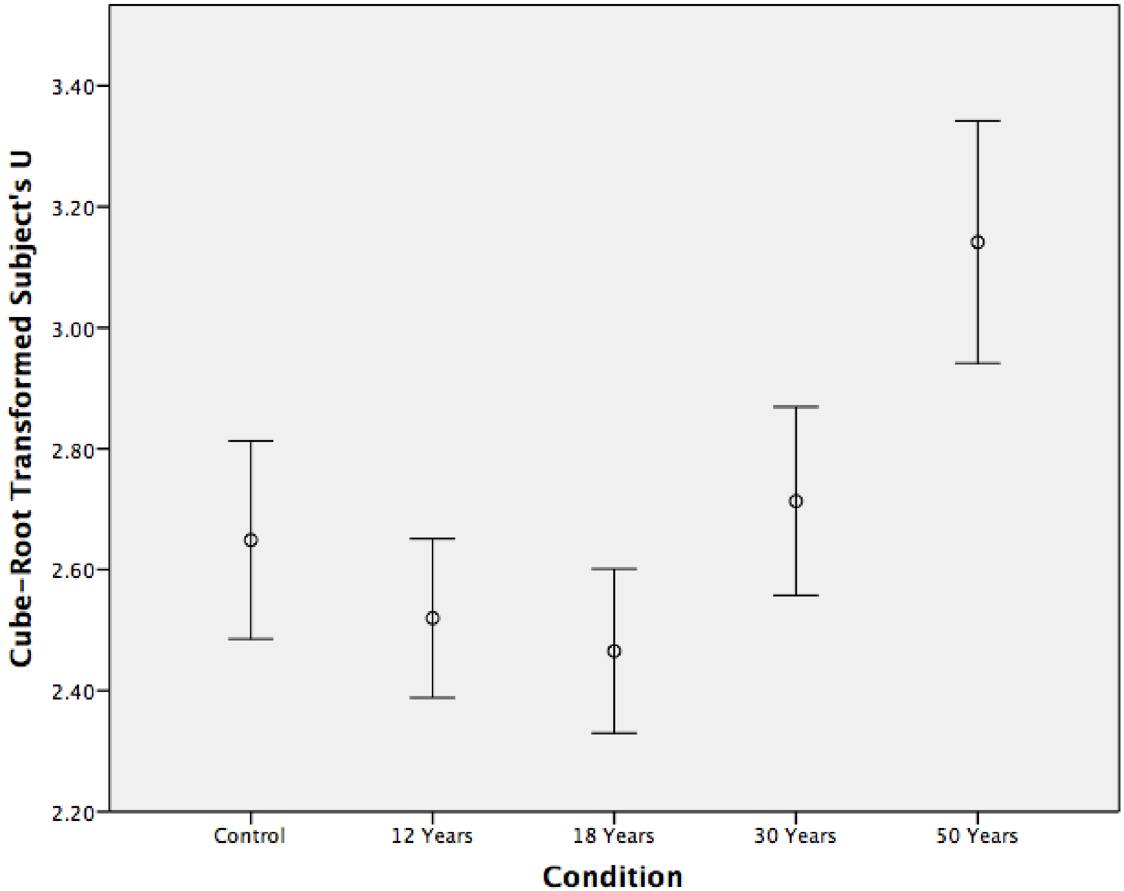


Figure 5. Pilot Study 4: The result of the manipulation check on subject's upper bound. The error bars are the 95% confidence intervals of the group means.

An omnibus one-way ANOVA with condition revealed that the control (M = 2.65, SD = 0.41, n = 27), 12-year (M = 2.52, SD = 0.33, n = 26), 18-year (M = 2.47, SD = 0.32, n = 24), 30-year (M = 2.71, SD = 0.37, n = 24), and 50-year (M = 3.14, SD = 0.47, n = 24) groups differed significantly, $F(4, 120) = 11.67$, $p < .001$, $\eta^2 = .280$. The Levene's test for equality of variances was not significant, $F(3, 94) = 2.38$, $p = .056$. Planned contrasts indicated a significant linear increase in subject's upper bound, $F(1, 120) = 23.74$, $p < .001$, $\eta^2 = .165$, and a significant quadratic effect, $F(1, 120) = 24.13$, $p < .001$, $\eta^2 = .167$. The cubic effect was not significant, $F(1, 120) = 0.19$, $p = .665$.

Because one goal of Pilot Study 4 was to find three appropriate lengths of maximum sentence that would work in the main study, further analysis was done by excluding the control group and the 12-year group. As expected, the three remaining groups differed significantly, $F(2, 69) = 18.14$, $p < .001$, $\eta^2 = .345$. The linear increase was significant, $F(1, 69) = 35.44$, $p < .001$, $\eta^2 = .339$, whereas the quadratic effect was not significant, $F(1, 69) = 0.84$, $p = .362$. To further examine which pair of differences was significant, a post hoc Tukey test comparing all pairs of group means was conducted. Among the three pairs of group mean differences, the two involving the 50-year group (i.e., the difference between the 50-year group and the 30-year group and the difference between the 50-year group and the 18-year group) were significant ($p < .002$ for both), whereas the difference between the 18-year group and the 30-year group was not (p = .081).

The subject's L. Across all subjects, the subject's L ranged from 3 to 30 years (M = 6.28, Mdn = 5.00, SD = 4.65, N = 125). The distribution of this variable had a skewness of 2.03 (*SE* = 0.22) and a kurtosis of 6.68 (*SE* = 0.43). This variable was winsorized at the 95th percentile (15 years). Next, the winsorized variable was transformed by taking the square root. The distribution of the transformed variable had a skewness of -0.16 (*SE* = 0.22) and a kurtosis of 0.02 (*SE* = 0.43). Based on the transformed winsorized values, a two-way ANOVA on the subject's lower bound with both condition and order as the predictors was conducted. None of the effects was significant. This result was expected, because the subject's lower bound was not intended to be affected by the manipulation.

Psychological discrepancy. Across all subjects, this variable ranged from 0 to 950 (*M* = 124.84, *Mdn* = 100.00, *SD* = 134.89, *N* = 125). The distribution of this variable had a skewness of 2.63 (*SE* = 0.22) and a kurtosis of 11.20 (*SE* = 0.43). This variable was winsorized at the 95th percentile (400). Next, the winsorized variable was transformed by taking the square root. The distribution of the transformed variable had a skewness of 0.14 (*SE* = 0.22) and a kurtosis of -0.56 (*SE* = 0.43). Figure 6 presents the group means with the 95% confidence intervals based on the transformed winsorized values. According to the assumption that as the subject's upper bound increases, the subject's psychological discrepancy decreases, the pattern in Figure 6 corresponded well to the pattern in Figure 5.

An omnibus one-way ANOVA with condition revealed that the control (*M* = 9.33, *SD* = 4.17, *n* = 27), 12-year (*M* = 10.21, *SD* = 5.14, *n* = 26), 18-year (*M* = 11.05, *SD* = 5.61, *n* = 24), 30-year (*M* = 10.01, *SD* = 5.32, *n* = 24), and 50-year (*M* = 7.06, *SD* = 5.42, *n* = 24) groups did not differ significantly, $F(4, 120) = 2.09$, $p = .086$, $\eta^2 = .065$. The Levene's test for equality of variances was not significant, $F(4, 120) = 0.57$, $p = .684$.

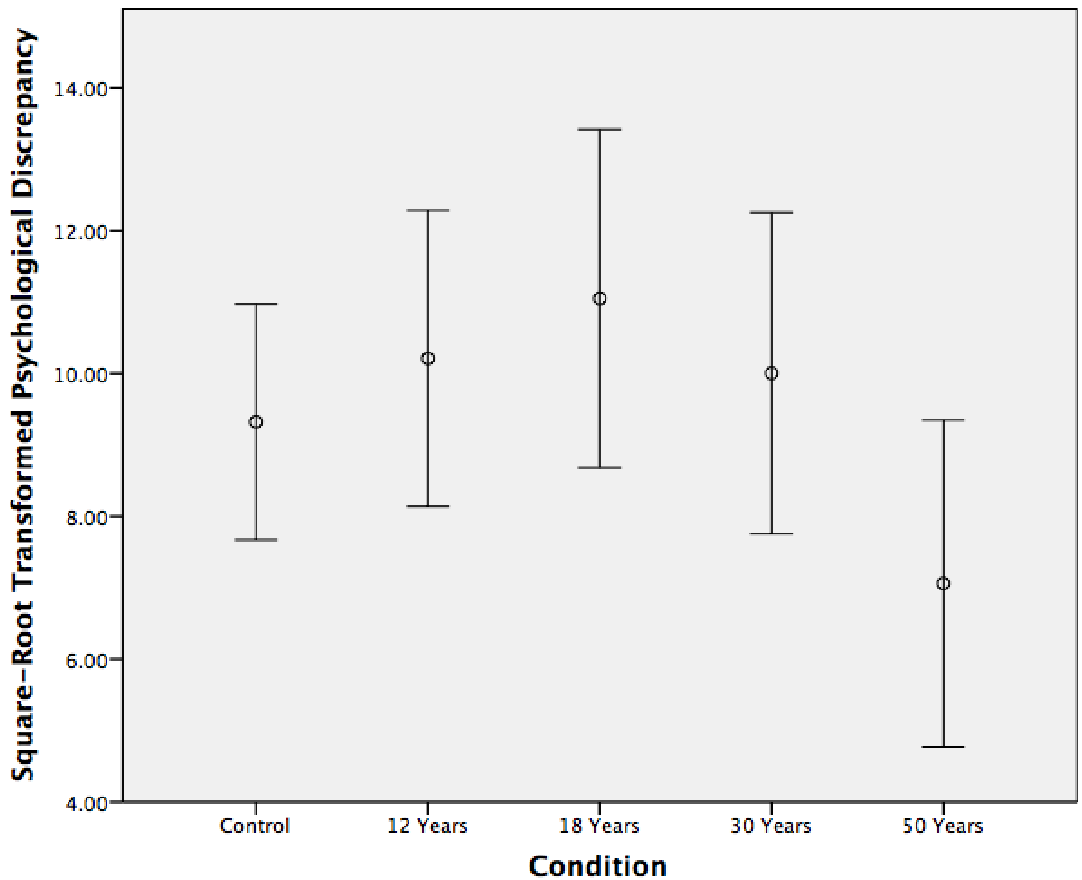


Figure 6. Pilot Study 4: The effect of the manipulation on psychological discrepancy. The error bars are the 95% confidence intervals of the group means. Skewness coefficients by condition: -0.22, 0.13, 0.02, 0.22, 0.48.

By excluding the control and the 12-year group, the remaining three groups did differ significantly, $F(2, 69) = 3.46$, $p = .037$, $\eta^2 = .091$. Planned contrasts indicated a significant linear decrease, $F(1, 69) = 6.43$, $p = .014$, $\eta^2 = .085$. The quadratic effect was not significant, $F(1, 69) = 0.49$, $p = .488$.

A post hoc Tukey test comparing all pairs of group means revealed that only the difference between the 18-year group and the 50-year group was significant ($p = .036$). Neither the difference between the 18-year and the 30-year group ($p = .785$), nor the difference between the 30-year and the 50-year group ($p = .155$) was significant.

According to $\psi = kD/(U - L)$, as the positional discrepancy increases, the psychological discrepancy increases. In Pilot Study 4, the message scale value was kept constant, whereas the subject's initial position was fixed at 10 years. However, only 47% (59 out of 125) of subjects reported 10 years as their initial positions. To have a more stringent test of the effect of the manipulation on psychological discrepancy, one should account for the subject's initial position.

Therefore, a one-way ANCOVA with condition being the predictor and initial position being the covariate was conducted for the 18-year, 30-year, and 50-year groups. The unstandardized residuals of the specified model followed an approximate normal distribution with a skewness of -0.16 ($SE = 0.28$). The Levene's test for equality of variances of the residuals was not significant, $F(2, 69) = 0.22$, $p = .805$. Adjusted for the subject's initial position, the group means did not differ significantly, $F(2, 68) = 1.97$, $p = .148$, $\eta^2 = .055$. The adjusted group means were 10.67 ($SE = 1.07$), 9.77 ($SE = 1.06$), and

7.68 (SE = 1.08), respectively. The linear trend was only marginally significant (p = .057, η^2 = .052), and the quadratic trend was not significant (p = .651).

Discussion

The first goal of Pilot Study 4 was to find three numbers of maximum sentence that would work. By refining some parts of the questionnaire, based on the direct check on the subject's upper bound, I found that 18 years, 30 years, and 50 years worked (Figure 5). The second goal of Pilot Study 4 was to see whether the manipulation of the upper bound would induce the predicted differences in psychological discrepancy. Although the group means of psychological discrepancy showed an expected pattern (Figure 6), after controlling for the effect of the subject's initial position on psychological discrepancy, the inferential tests did not show significant differences among the 18-year, 30-year, and 50-year groups.

What were the implications for the main study? First, with regard to the manipulation of upper bound, to ensure sufficient group differences in psychological discrepancy, it was possible that using 15 years instead of 18 years would generate greater differences among the three groups.

Second, the obtained size of the effect of the manipulation of the upper bound on psychological discrepancy adjusted for the subject's initial position can be used to determine the intended sample size for the main study. The major concern was how large a sample size should be to detect a probably *small* effect of the manipulation of the upper bound on psychological discrepancy given an acceptable power level (i.e., $1 - \beta$, where β is the probability of Type II error of failing to reject a false H_0). Based on the results reported above about psychological discrepancy, adjusted for the subject's initial position, the linear trend of the manipulation's effect on psychological discrepancy had a partial η^2 of .052. Using Steiger's (2004, p. 173) method, the 95% CI of the population effect size was [0, .184]. By using G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007), for a priori fixed-effects ANOVA with main effects and interactions, a partial η^2 of .052 (the point estimate of the population effect size) translated to Cohen's (1988) f = 0.23, and a partial η^2 of .184 (the upper endpoint of the 95% confidence interval of the population effect size) translated to Cohen's f = 0.47.[6]

With regard to the convention of defining the magnitude of effect size for an F test, Cohen (1988, pp. 285-288) considered f = 0.10 a small effect, f = 0.25 a medium effect, and f = 0.40 a large effect. Because the 95% confidence interval of the linear trend's effect size reported above spanned from no effect to large effect, it was difficult to decide *one* effect size to be used in the a priori power analysis. Therefore, different f values were used. Also, I chose three levels of power (i.e., .80, .90, and .95),[7] chose an alpha level of .05, and set the data analytic context as a 3 × 3 factorial between-subjects design consistent with the main study. Table 15 shows the required *total* sample sizes returned by G*Power 3.1 (Faul et al., 2007) to detect a significant main effect given the specified parameters. Table 16 is the interaction-effect counterpart of Table 15.

In consideration of feasibility and cost, I proposed recruiting 50 subjects per cell; thus, a balanced 3 × 3 design would have a total of 50 × 9 = 450 subjects. This sample size would detect a significant interaction effect of a size about 0.20 with a 5% chance of falsely retaining a false H_0 (see Table 16), which seemed reasonable.

Table 15

A Priori Power Analysis: The Required Total Sample Sizes to Detect a Significant Main Effect in a 3 × 3 Factorial Between-Subjects Design as a Function of Effect Size and Power

Power	f [a]					
	0.10	0.15	0.20	0.25	0.30	0.40
.80	967	432	244	158	111	64
.90	1269	566	320	206	144	83
.95	1548	690	390	251	175	100

Note. The results were returned by G*Power 3.1 (Faul et al., 2007). The alpha level was set at .05. G*Power settings: The test family was "*F* tests." The statistical test was "ANOVA: Fixed effects, special, main effects and interactions." Number of groups was 9. Numerator *df* was 2 for a main effect in a 3 × 3 factorial design.

[a]Cohen's (1988) effect size index *f* (see Note 5).

Table 16

A Priori Power Analysis: The Required Total Sample Sizes to Detect a Significant Interaction Effect in a 3 × 3 Factorial Between-Subjects Design as a Function of Effect Size and Power

Power	f[a]					
	0.10	0.15	0.20	0.25	0.30	0.40
.80	1199	536	304	196	138	80
.90	1546	690	390	252	177	102
.95	1862	831	470	302	212	122

Note. The results were returned by G*Power 3.1 (Faul et al., 2007). The alpha level was set at .05. G*Power settings: The test family was "*F* tests." The statistical test was "ANOVA: Fixed effects, special, main effects and interactions." Number of groups was 9. Numerator *df* was 4 for the interaction effect in a 3 × 3 factorial design.

[a]Cohen's (1988) effect size index *f* (see Note 5).

If an effect size of 0.20 was still not conservative enough, I could aim for 0.15. If one could bear with a .10 probability Type II error, the required sample size would be 690 (see Table 16), which meant more than 75 subjects per cell. I was dealing with an effect size whose confidence interval was large based on the results of Pilot Study 4. With a high level of uncertainty, it would be more appropriate to conclude that a large sample size was needed to detect a probably small effect with sufficient power.

Pilot Study 5

The first goal of Pilot Study 5 was to find three appropriate values of the message scale value (s) for manipulating positional discrepancy in the main study. More specifically, Pilot Study 5 examined whether the induced change in positional discrepancy would lead to change in psychological discrepancy (ψ). Holding perspective ($P = U - L$) constant, a successful manipulation of positional discrepancy should find the following relationship: $\psi_{D_H} > \psi_{D_M} > \psi_{D_L}$, given $\psi = kD/(U - L)$, where D_H, D_M, and D_L represent a high, a moderate, and a low positional discrepancy, respectively. Ceteris paribus, as D increases, ψ increases.

Secondly, based on previous studies of positional discrepancy's effect on belief change (Aronson et al., 1963; Bochner & Insko, 1966; Chung et al., 2008; Kaplowitz & Fink, 1991; Laroche, 1977; see also Fink & Cai, 2012; Kaplowitz & Fink, 1997), it was predicted that a successful manipulation of positional discrepancy should produce significantly different group means of the amount of belief change. Or, equivalently, holding subject's initial position constant, the manipulation of message scale value should produce significantly different group means for final belief position. Whether the values used for the message scale value (see below) were able to generate the predicted differences in the amount of belief change was tested in Pilot Study 5.

For all subjects, the maximum years of the sentence for an armed bank robbery from allegedly previous trials in the U.S. was fixed at 30 years. With regard to the manipulation of the message scale value, three groups of subjects were told that a judge's sentence was 15, 23, and 35 years, respectively. The questionnaires were exactly the same as those used in the fourth pilot study, except for the change of numbers noted above.

Procedure

The survey was launched via TurkPrime (Litman et al., 2017) at 01:06 p.m., Eastern Standard Time, June 12, 2019. To be eligible to participate in the study, an MTurk worker must have had 5,000 or more approved MTurk tasks and a 98% or above approval rate of the completed tasks. The intended sample size for Pilot Study 5 was 75. The data collection process was completed at 05:00 p.m., Eastern Standard Time, June 13, 2019. By then, 75 completed responses had been recorded on Qualtrics (2018). Each of the 75 subjects was paid $0.80. After debriefing, no subjects refused to allow their responses to be used. For the 75 subjects, the average time to complete the survey on Qualtrics (2018) was 595.96 seconds (*SD* = 216.49), which equals 9.93 minutes. The shortest and the longest completion times were 268.00 seconds (4.47 minutes) and 1,469.00 seconds (24.48 minutes), respectively. The median was 560.00 seconds, which equals 9.33 minutes.

Results

Psychological discrepancy. Across all subjects, this variable ranged from 0 to 5,000 (M = 177.08, Mdn = 60.00, SD = 585.32, N = 75). This variable was winsorized at the 95th percentile (440). Next, the winsorized variable was transformed by taking the cube root. The distribution of the transformed variable had a skewness of -0.50 (SE = 0.28) and a kurtosis of 0.04 (SE = 0.55). Figure 7 presents the group means with the 95% confidence intervals based on the transformed winsorized values. The pattern in Figure 7 met the expectation that a greater message scale value should lead to a greater level of psychological discrepancy.

An omnibus one-way ANOVA with condition revealed that the 15-year (M = 3.42, SD = 1.49, n = 26), 23-year (M = 3.80, SD = 1.80, n = 24) and 35-year (M = 4.76, SD = 2.25, n = 25) groups differed significantly, $F(2, 72) = 3.50$, $p = .035$, $\eta^2 = .089$. The Levene's test for equality of variances was not significant, $F(2, 72) = 1.76$, $p = .179$. Planned contrasts indicated a significant linear increase, $F(1, 72) = 6.65$, $p = .012$, $\eta^2 = .085$. The quadratic effect was not significant, $F(1, 72) = 0.38$, $p = .538$.

A post hoc Tukey test comparing all pairs of group means revealed that only the difference between the 15-year group and the 35-year group was significant ($p = .032$). Neither the difference between the 15-year and the 23-year group ($p = .744$), nor the difference between the 23-year and the 35-year group ($p = .177$) was significant.

To have a more stringent test of the effect of the manipulation of the message scale value on psychological discrepancy, one should account for the subject's initial position, just like what was done in Pilot Study 4. Therefore, a one-way ANCOVA with

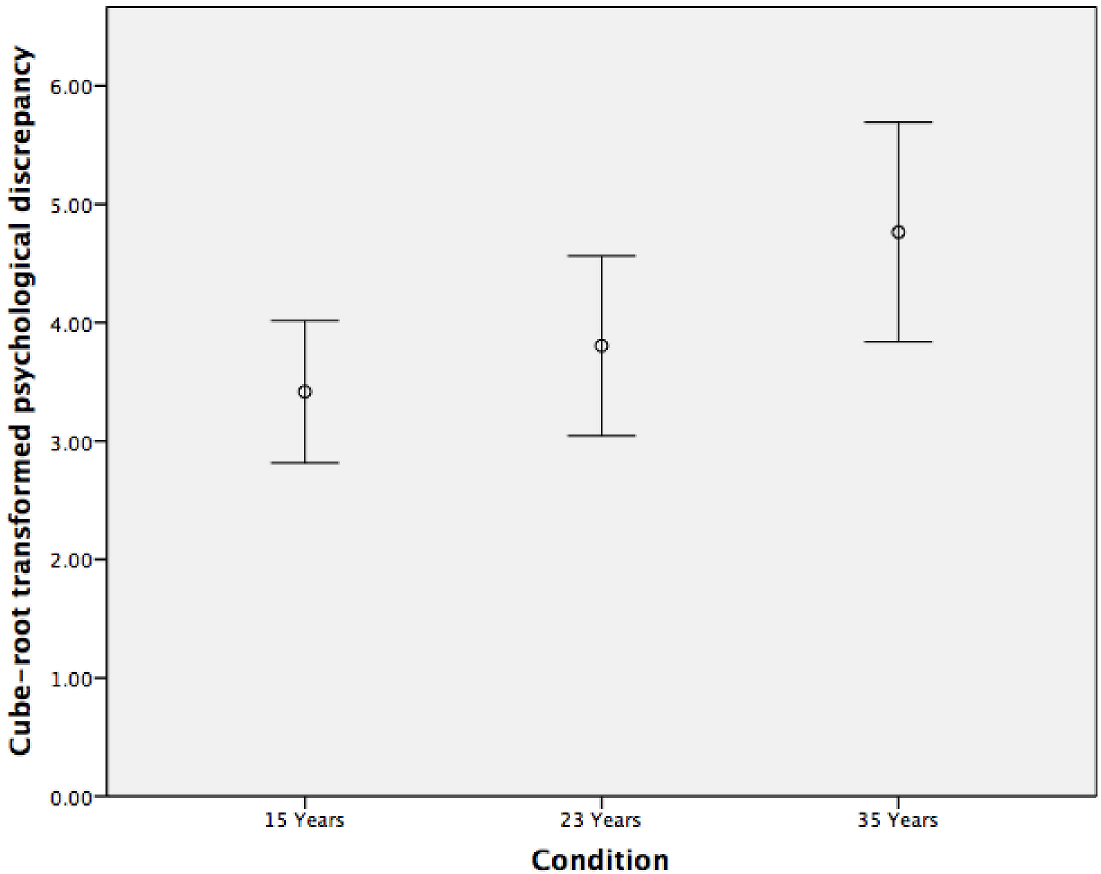


Figure 7. Pilot Study 5: The effect of the manipulation on psychological discrepancy. The error bars are the 95% confidence intervals of the group means. Skewness coefficients and their standard errors by condition: -0.64 (*SE* = 0.46), -0.95 (*SE* = 0.47), -0.97 (*SE* = 0.46).

condition being the predictor and initial position being the covariate was conducted. The unstandardized residuals of the specified model followed an approximate normal distribution with a skewness of -0.37 (SE = 0.28). The Levene's test for equality of variances of the residuals was not significant, $F(2, 72) = 0.95$, $p = .391$. Adjusted for the subject's initial position, the group means of psychological discrepancy still differed significantly, $F(2, 71) = 3.94$, $p = .024$, $\eta^2 = .100$. The adjusted group means were 3.46 (SE = 0.33), 3.77 (SE = 0.35), and 4.75 (SE = 0.34), respectively. The linear trend was significant ($p = .009$, $\eta^2 = .093$), and the quadratic trend was not significant ($p = .433$).

The subject's final position. The subjects were asked, "To how many years in prison do you think Convict X should have been sentenced?" Across all subjects, the subject's final position (R) ranged from 4 to 35 years (M = 14.81, Mdn = 15.00, SD = 7.44, N = 75). The distribution of this variable had a skewness of 1.11 (SE = 0.28) and a kurtosis of 0.98 (SE = 0.55). After taking the square root, the distribution of the transformed variable had a skewness of 0.48 (SE = 0.28) and a kurtosis of 0.11 (SE = 0.55). An omnibus one-way ANOVA with condition revealed that the 15-year (M = 3.68, SD = 0.85, n = 26), 23-year (M = 3.77, SD = 0.91, n = 24) and 35-year (M = 3.76, SD = 1.06, n = 25) groups did not differ significantly, $F(2, 72) = 0.07$, $p = .929$, $\eta^2 = .002$. A one-way ANCOVA with condition being the predictor and initial position being the covariate was conducted. Adjusted for subject's initial position, the group means did not differ significantly, $F(2, 71) = 0.58$, $p = .562$, $\eta^2 = .016$. The expected effect of the manipulation of the message scale value on the subject's final position was not found.

Discussion

The first goal of the fifth pilot study was to find three appropriate values of the message scale value (*s*) for manipulating positional discrepancy in the main study. If the manipulation works, the induced change in positional discrepancy would lead to change in psychological discrepancy (ψ). I found that 15 years, 23 years, and 35 years worked (Figure 7). The effect size was larger than the effect size of the manipulation of upper bound (see Pilot Study 4). Thus, the estimated sample size for the main study would not change from what was proposed in the previous section regarding Pilot Study 4.

The second goal of the fifth pilot study was to examine whether the manipulation of positional discrepancy would induce significant group difference in subject's final position in light of relevant previous studies (Aronson et al., 1963; Bochner & Insko, 1966; Chung et al., 2008; Kaplowitz & Fink, 1991; Laroche, 1977; see also Fink & Cai, 2012; Kaplowitz & Fink, 1997). The answer was no. This result was unexpected. If a discrepancy model in belief change (Fink & Cai, 2012) is plausible, whether it be a linear or a nonlinear one, if the manipulation of positional discrepancy worked, I should have observed significant group differences in the subject's final position. There were two implications: (1) Something was wrong: A well-known discrepancy effect could not be replicated; (2) there might be another way to account for the absence of a significant *total* effect of manipulation on final position.

First, it might be the case that the intervals between the message scale values were not great enough. In Chung et al.'s (2008) study and Kaplowitz & Fink's (1991) study, the judge's decisions were 17, 30, and 50 years, respectively. It was possible that using these years would have generated greater differences in the subject's final position among three groups.

With regard to the absence of a significant total effect of the manipulation of the message scale value on final position, it could be that there existed a *mediator* from *manipulation* to *final position*. It was possible that the *indirect* effect and the *direct* effect were both significant but of opposite signs. For this dissertation, psychological discrepancy was a theoretically important mediator in the belief change process. In Pilot Study 5, the effect of the manipulation on psychological discrepancy was positive and significant, and the Pearson correlation between psychological discrepancy and the subject's final position was negative and significant ($r = -.51$, $p < .001$), but the effect of the manipulation on the subject's final position was nonsignificant. Therefore, it seemed that manipulation's indirect effect on subject's final position via psychological discrepancy might be negative and significant. This hypothesis was consistent with the psychological-discrepancy-weight-discounting model proposed and tested in Fink et al.'s (1983) study: As positional discrepancy increases, psychological discrepancy increases; the increased psychological discrepancy discounts the weight of the judge's decision, which in turn reduces the effect of positional discrepancy on final position. This hypothesis was also consistent with the psychological-discrepancy-scale-value-pullback model I proposed: As psychological discrepancy increases due to an increased positional discrepancy, the increased psychological discrepancy pulls the scale value of the judge's decision towards one's initial position, which also reduces the effect of positional discrepancy on final position.

A post hoc mediation analysis was conducted by using the PROCESS v2 macro (Hayes, 2013) in SPSS. A simple mediation model with subject's initial position as a statistical control was specified (Figure 8). The analysis showed that the indirect effect of message scale value (i.e., the manipulation in the fifth pilot study) on final position through psychological discrepancy was significant (unstandardized $ab = -0.01$) with initial position as a statistical control. A percentile bootstrap confidence interval (95%) for the indirect effect based on 10,000 bootstrap samples (Hayes, 2013, pp. 105-111) was entirely below zero [-0.017, -0.002]. Therefore, I could claim with 95% confidence that the indirect effect was negative. The direct effect of message scale value on final position was not significant (unstandardized $c' = 0.01$, $p = 0.07$).

The significant indirect effect of positional discrepancy on final position via psychological discrepancy in the absence of a significant direct effect was an important finding that supported psychological discrepancy as a *mediator* in discrepancy models of belief change. No studies on discrepancy models that I was aware of had reported a mediation analysis with psychological discrepancy as a mediator in a linear structural equation model. However, the results of the mediation analysis should be viewed with caution in two aspects.

First, there could be other unmeasured mediators that covaried with final position. For example, the perceived credibility, fairness, and convincingness of the judge's decision could be such variables. Other authors (Bullock, Green, & Ha, 2008, 2010; Bullock & Ha, 2011) have cautioned that the estimated indirect effect has to be biased with the existence of any unmeasured mediators covarying with the dependent variable in

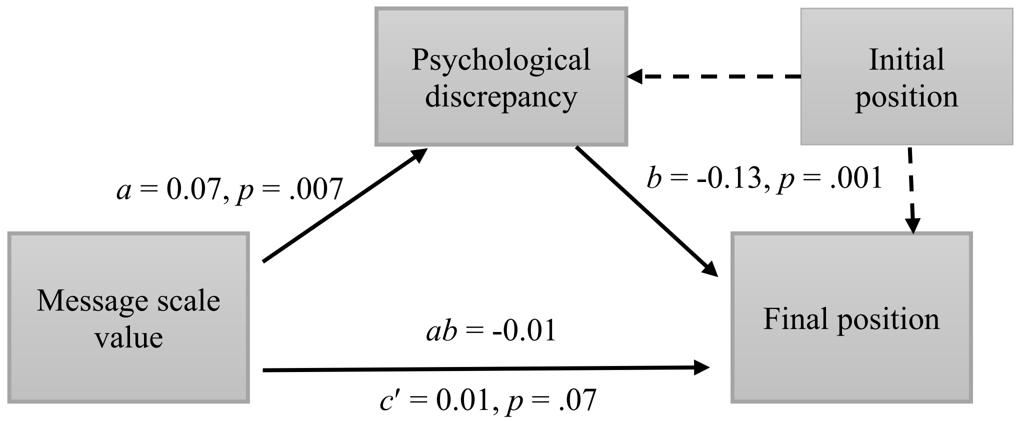


Figure 8. Pilot Study 5: A post hoc mediation analysis. The variable of initial position is a statistical control. The notation of the ordinary least squares (OLS) regression unstandardized estimates (*a*, *b*, and *c′*) follows Hayes (2013, p. 91). The statistical inference test for the indirect effect, *ab*, is reported in the main text. The *p* values for *a*, *b*, and *c′* were calculated in the same way that a *p* value is calculated in an OLS multiple regression (see Hanushek & Jackson, 1977, pp. 122-124; Hayes, 2013, p. 101). *N* = 75.

a mediation analysis. For Pilot Study 5, it was plausible that the indirect-effect estimate could indeed be biased. Secondly, note that the direct effect of the manipulation on final position was not significant. If a discrepancy model with psychological discrepancy as a mediator was plausible, the message scale value should have had a significant positive effect on final position after the path through psychological discrepancy was accounted for. To be consistent with the findings of previous studies (Aronson et al., 1963; Bochner & Insko, 1966; Chung et al., 2008; Kaplowitz & Fink, 1991; Laroche, 1977; see also Fink & Cai, 2012; Kaplowitz & Fink, 1997), the total effect of the manipulation on final position should have been positive and significant, which means that the positive direct effect should have been greater than the negative indirect effect in magnitude.

Therefore, another pilot study was needed. First, was there a greater interval in the judge's decision across the three conditions that would lead to a significant total effect of manipulation on final position? Second, was there a significant indirect effect in the mediation analysis reported above that could be replicated?

Pilot Study 6

This pilot study used 17, 30, and 50 years as the judge's sentence. These values were exactly the same as the values used in two previous similar studies (Chung et al., 2008; Kaplowitz & Fink, 1991). Pilot Study 5 found that the effect of the message-scale-value manipulation on the final position via psychological discrepancy was significant. Thus, Pilot Study 6 was conducted to check whether this indirect effect could be replicated. The design of Pilot Study 6 was the same as Pilot Study 5 except for the change of the message scale values noted above.

Procedure

The survey was launched via TurkPrime (Litman et al., 2017) at 6:34 p.m., Eastern Standard Time, June 14th, 2019. To be eligible to participate in the study, an MTurk worker must have had 5,000 or more approved MTurk tasks and a 98% or above approval rate of completed tasks. The intended sample size for Pilot Study 6 was 75 (25 for each of the three conditions). The data collection process was completed at 12:38 p.m., Eastern Standard Time, June 15th, 2019. By then, 75 completed responses had been recorded on Qualtrics (2018). Each of the 75 subjects was paid $0.80. After debriefing, no subjects refused to allow their responses to be used. For the 75 subjects, the average time to complete the survey on Qualtrics (2018) was 676.37 seconds (SD = 347.64), which equals 11.27 minutes. The shortest and the longest completion times were 297.00 seconds (4.95 minutes) and 2,277.00 seconds (37.95 minutes), respectively. The median was 627.00 seconds, which equals 10.45 minutes.

Results

The subject's final position. The subjects were asked, "To how many years in prison do you think Convict X should have been sentenced?" Across all subjects, the subject's final position (R) ranged from 1 to 50 years (M = 16.25, Mdn = 15.00, SD = 8.74, N = 75). This variable had a skewness of 1.06 (SE = 0.28) and a kurtosis of 1.97 (SE = 0.55). After winsorizing this variable at the 95th percentile (31 years), the winsorized variable had a skewness of 0.41 (SE = 0.28) and a kurtosis of -0.54 (SE = 0.55). After winsorization, the skewness was not significantly different from zero ($0.41/0.28 = 1.46 <$

$z_{.025} = 1.96$); likewise, the kurtosis was not significantly different from zero (0.54/0.55 = 0.98 < $z_{.025} = 1.96$). Table 17 shows that after winsorization, *in each group*, the skewness and the kurtosis were not significantly different from zero either.

A one-way ANCOVA on the subject's final position with condition being the predictor and initial position being the covariate was conducted. The unadjusted group means were 15.08 (SD = 4.90, n = 24), 15.85 (SD = 9.39, n = 27), and 16.75 (SD = 8.27, n = 24) for the 17-, 30-, and 50-year group, respectively. The adjusted group means controlling for the subject's initial position were 14.62 (SE = 1.26), 17.18 (SE = 1.20), and 15.73 (SE = 1.26), respectively. The pattern of the adjusted group means was in an inverted U-shape. Adjusted for subject's initial position, the group means did not differ significantly, $F(2, 71) = 1.09$, $p = .343$, $\eta^2 = .030$. The linear trend was not significant, $F(1, 71) = 0.39$, $p = .533$, $\eta^2 = .006$. The quadratic trend was not significant either, $F(1, 71) = 1.77$, $p = .188$, $\eta^2 = .024$.

Were the results improved from the last pilot study? In Pilot Study 5, adjusted for initial position, the condition's effect was not significant, $F(2, 71) = 0.58$, $p = .562$, $\eta^2 = .016$. The linear trend was not significant, $F(1, 71) = 0.71$, $p = .403$, $\eta^2 = .010$, and the quadratic trend was not significant either, $F(1, 71) = 0.44$, $p = .507$, $\eta^2 = .006$. Comparing the two sets of results revealed that the manipulation used in the sixth pilot study did improve with respect to the effect size of condition ($\eta^2 = .030$ vs. $\eta^2 = .016$). However, the inferential test was still not significant.

Table 17

*Pilot Study 6: Group Means (*SE*), Skewness Coefficients (*SE*), Kurtosis Coefficients (*SE*) of the Subject's Final Position With the Raw Data and the Winsorized Data for the 17-Year, 30-Year, and 50-Year Groups*

		Raw data	Winsorized data
Group 1 17 years *n* = 24	Mean	15.08 (1.00)	15.08 (1.00)
	Skewness	-0.02 (0.47)	-0.02 (0.47)
	Kurtosis	0.81 (0.92)	0.81 (0.92)
Group 2 30 years *n* = 27	Mean	16.15 (1.91)	15.85 (1.81)
	Skewness	0.40 (0.45)	0.24 (0.45)
	Kurtosis	-0.92 (0.87)	-1.25 (0.87)
Group 3 50 years *n* = 24	Mean	17.54 (2.11)	16.75 (1.69)
	Skewness	1.48 (0.47)	0.55 (0.47)
	Kurtosis	2.91 (0.92)	-0.75 (0.92)

Note. The values are the point estimates of the statistics with their standard errors in the parentheses. Data winsorized at R = 31.

Further analysis of the condition's effect on the final position. The above results indicated that using 17, 30, and 50 years as the message scale values was not able to replicate the message discrepancy effect, although the effect size increased. To further analyze the data, I considered the six conditions (the 15-, 23-, 35-, 17-, 30-, and 50-year groups) from Pilot Study 5 and Pilot Study 6 together. Doing this could be justified because the six conditions only differed in the message scale value, which can be corroborated by examining whether the two sets of data differed on other variables that were not supposed to be affected by the message-scale-value manipulation (see Appendix G). I could treat any three conditions as if I randomized and administered the three questionnaires in one attempt. If I could find three conditions that differed significantly in the final position, the message scale values of these conditions would be good candidates for the main study.

Because I got better results in Pilot Study 6 than in Pilot Study 5, I started with the 17-, 30-, and 50-year groups. Next, I examined the means of the final position of the six groups based on the raw data. The means were 14.23 (n = 26), 15.08 (n = 24), 15.04 (n = 24), 16.15 (n = 27), 15.20 (n = 25), and 17.54 (n = 24) for the 15-, 17-, 23-, 30-, 35-, and 50-year group, respectively. I found that the 15-year group's mean was lower than the 17-year group. Therefore, I might obtain even better results than the ones of Pilot Study 6 if I could run an analysis on the 15-, 30-, and 50-year groups.

Across the 15-, 30-, and 50-year conditions, the subject's final position ranged from 1 to 50 years (M = 15.94, Mdn = 15.00, SD = 9.07, N = 77). This variable had a skewness of 1.08 (SE = 0.27) and a kurtosis of 1.54 (SE = 0.54). After winsorizing this variable at the 95th percentile (34 years), the winsorized variable had a skewness of 0.63 (*SE* = 0.27) and a kurtosis of -0.39 (*SE* = 0.54). The winsorized variable was transformed via a square root. Now the distribution of the square root of the winsorized variable had a skewness of 0.05 (*SE* = 0.27) and a kurtosis of -0.42 (*SE* = 0.54). This skewness was not significantly different from zero (0.05/0.27 = 0.19 < $z_{.025}$ = 1.96); likewise, the kurtosis was not significantly different from zero (0.42/0.54 = 0.78 < $z_{.025}$ = 1.96). Table 18 shows that after winsorization and transformation, *in each group*, the skewness and the kurtosis were not significantly different from zero either.

A one-way ANCOVA on the subject's final position with the condition being the predictor and the initial position being the covariate was conducted. The unadjusted group means were 3.68 (*SD* = 0.85, *n* = 26), 3.80 (*SD* = 1.30, *n* = 27), and 3.98 (*SD* = 1.04, *n* = 24) for the 15-, 30-, and 50-year group, respectively. The adjusted group means accounting for the subject's initial position were 3.58 (*SE* = 0.15), 4.05 (*SE* = 0.15), and 3.81 (*SE* = 0.16), respectively. Again, the pattern of these adjusted means reflected an inverted-*U* shape. The Levene's test of equality of error variances was not significant, $F(2, 74) = 1.63$, $p = .203$. Adjusted for the subject's initial position, the group means did not differ significantly, $F(2, 73) = 2.46$, $p = .093$, $\eta^2 = .063$. However, notice the greater effect size based on the 15-, 30-, and 50-year groups than the effect size based on the 17-, 30-, and 50-year groups ($\eta^2 = .063$ vs. $\eta^2 = .030$). The linear trend was still not significant, $F(1, 73) = 1.09$, $p = .300$, $\eta^2 = .015$. The quadratic trend was *marginally* significant, $F(1, 73) = 3.69$, $p = .059$, $\eta^2 = .048$. Again, notice the improved effect size

Table 18

*Pilot Study 5 and Pilot Study 6: Group Means (*SE*), Skewness Coefficients (*SE*), Kurtosis Coefficients (*SE*) of the Subject's Final Position With the Raw Data, the Winsorized Data, and the Square Root of the Winsorized Data for the 15-Year, 30-Year, and 50-Year Groups*

		Raw data	Winsorized data	Square root of winsorized data
Group 1 15 years *n* = 26	Mean	14.23 (1.29)	14.23 (1.29)	3.68 (0.17)
	Skewness	1.14 (0.46)	1.14 (0.46)	0.41 (0.46)
	Kurtosis	1.94 (0.89)	1.94 (0.89)	0.63 (0.89)
Group 2 30 years *n* = 27	Mean	16.15 (1.91)	16.07 (1.88)	3.80 (0.25)
	Skewness	0.40 (0.45)	0.26 (0.45)	-0.17 (0.45)
	Kurtosis	-0.92 (0.87)	-1.03 (0.87)	-0.85 (0.87)
Group 3 50 years *n* = 24	Mean	17.54 (2.11)	16.88 (1.69)	3.98 (0.21)
	Skewness	1.48 (0.47)	0.63 (0.47)	0.19 (0.47)
	Kurtosis	2.91 (0.92)	-0.58 (0.92)	-0.69 (0.92)

Note. The values are the point estimates of the statistics with their standard errors in the parentheses. Data winsorized at $R = 34$.

(η^2 = .048 vs. η^2 = .024). By analyzing the 15-, 30-, and 50-year groups I indeed obtained better results.

A mediation analysis. Based on the 15-, 30-, and 50-year groups, the same simple mediation model as in Pilot Study 5 (Figure 9) was run by using the PROCESS v2 macro (Hayes, 2013) in SPSS. As expected, the analysis showed that the indirect effect of message scale value on final position through psychological discrepancy was significant (ab = -0.01), with initial position as a statistical control. A percentile bootstrap confidence interval (95%) for the indirect effect based on 10,000 bootstrap samples (Hayes, 2013, pp. 105-111) was entirely below zero [-0.022, -0.006]. Also, as expected, the manipulation's effect on psychological discrepancy was highly significant (a = 0.18, p < 0.001), which showed the manipulation was successful. In contrast to Pilot Study 5, the direct effect of message scale value on the final position was significant (c' = 0.02, p = 0.004). This positive direct effect meant that as message scale value increased, the subject's final position increased after the path through psychological discrepancy was accounted for. Therefore, analyzing the data of the 15-, 30-, and 50-year groups was able to replicate the message discrepancy effect.

Main Study

Participants

Like the pilot studies, the main study recruited its participants from MTurk via TurkPrime (Litman et al., 2017). To be eligible to participate in the study, an MTurk worker must have had 5,000 or more approved MTurk tasks and a 98% or above approval rate of the completed tasks. The intended sample size for the main study was 450 based

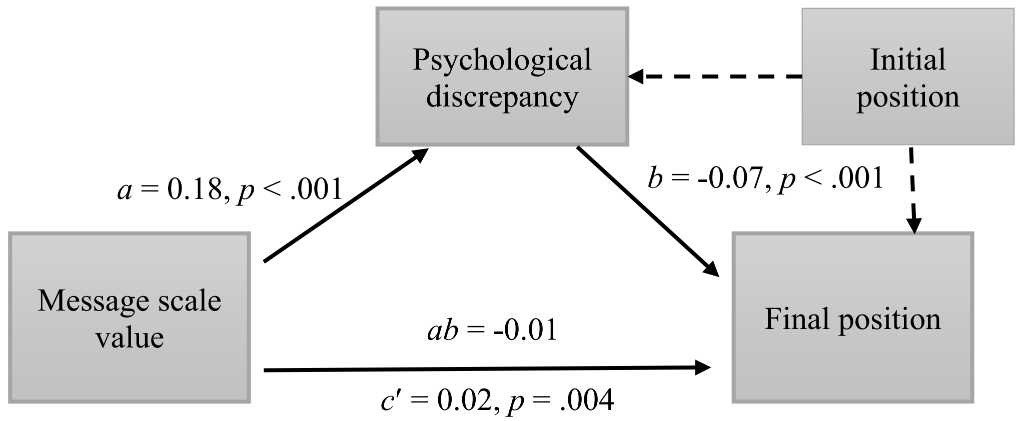


Figure 9. Pilot Study 6: A mediation analysis. The variable of initial position is a statistical control. The notation of the ordinary least squares (OLS) regression unstandardized estimates (*a*, *b*, and *c′*) follows Hayes (2013, p. 91). The statistical inference test for the indirect effect, *ab*, is reported in the main text. The *p* values for *a*, *b*, and *c′* were calculated in the same way that a *p* value is calculated in an OLS multiple regression (see Hanushek & Jackson, 1977, pp. 122-124; Hayes, 2013, p. 101). $N = 77$.

on the power analysis reported in the section on Pilot Study 4. The survey was launch at 10:21 a.m., Eastern Standard Time, November 9, 2019. The data collection process was completed at 10:53 p.m., Eastern Standard Time, November 14, 2019. By then, 450 completed responses had been recorded on Qualtrics (2018). All 450 subjects' responses were approved for compensation. Each of the 450 subjects was paid \$2.20. This payment was decided according to the U.S. federal minimum wage that was effective in November 2019. The minimum wage was \$7.25 (U.S. Department of Labor, 2009). Given that the estimated duration for a subject to complete the survey for the main study on Qualtrics (2018) was 18 minutes based on the pilot studies, $(\$7.25/60)\times 18 = \$2.175 \approx \$2.20$. After debriefing, two subjects refused to allow their responses to be used. Therefore, the remaining 448 subjects' responses were used for data analysis.

Among the 448 subjects, 240 were males (53.57%), 204 were females (45.54%), and 4 reported other (0.89%). The average age was 37.60 years (Mdn = 35.00, SD = 11.00, Min = 19.00, Max = 72.00, Sk = 0.88 [SE = 0.12], Ku = 0.26 [SE = 0.23]). The number of self-identified Democrats was 244 (54.46%), and the number of self-identified Republicans was 137 (30.58%).

For the 448 subjects, the average time to complete the survey on Qualtrics (2018) was 1,197.83 seconds (SD = 853.93), which equals 19.96 minutes. The shortest and the longest completion times were 429.00 seconds (7.15 minutes) and 9,217.00 seconds (153.62 minutes), respectively. The median was 950.50 seconds, which equals 15.84 minutes.

Experiment Stimulus and Procedures

The main study was a between-subjects 3 (perspective) × 3 (message scale value) design. Following Ostrom's (1970) manipulation of perspective, defined as the difference between subject's upper bound (U) and lower bound (L) of belief position on some topic (i.e., $U - L$), perspective was varied by only changing subject's U across the three levels of perspective. Message scale value was manipulated by changing the length of a sentence in years given by a judge (Chung et al., 2008; Kaplowitz & Fink, 1991). According to the results of the pilot studies, the main study used the values presented in Table 19 for the two factors. A subject was randomly assigned to one of the nine conditions.

Table 19

*Manipulation Levels in the Main Study: The Advocated Years of Sentence (*s*) and the Upper Bound (*U*)*

	Low U	Moderate U	High U
Low s	15 / 15	15 / 30	15 / 50
Moderate s	30 / 15	30 / 30	30 / 50
High s	50 / 15	50 / 30	50 / 50

Note. In each cell, the value before the slash is the advocated number of years of the sentence given by the judge (s), and the value after the slash is the harshest number of years of the sentence in previous trials (U).

A stimulus message was composed of three parts (see Appendix F for the questionnaire used for the main study). The first part presented a (bogus) U.S. federal sentencing guideline for armed bank robbery that explicitly states that an appropriate sentence was 10 years (see also Chung et al., 2008; Kaplowitz & Fink, 1991). The length of this sentence was determined based on the results of Pilot Study 1. The second part of the stimulus message presented different (bogus) values of maximum years of the sentence for armed bank robbery from allegedly previous trials in the U.S. (U), by which perspective was manipulated. The third part presented a (bogus) decision for Convict X (including a speech that stated the length of imprisonment in years) made by Judge Walters (a fictitious person who was described as a judge in a U.S. district court), in which the message scale value was presented (s). To keep the message weight constant, the messages across different conditions were the same except for the value of the previous maximum sentence. The questions for measuring the dependent variables (see below) were inserted either between the three components in the stimulus message (see above) or after the stimulus message.

Measures

Initial Belief Position

After the sentencing guideline was presented, the subjects were asked, "What do you believe is the most appropriate sentence (in number of years) for the crime of armed bank robbery?"

Subjects' U and L

After reading the maximum sentence for armed robbery in the U.S., subjects were asked, "What do you believe is a maximally harsh or lenient sentence (in number of years) for a convict of armed bank robbery?" For about half of the subjects, they were asked about the maximally harsh sentence first and then asked about the maximally lenient sentence; for the other half of the subjects, the order was reversed. The subjects were randomly assigned to one of these two order conditions.

Self-Report Measure of Psychological Discrepancy

After reading the stimulus message, the degree of psychological discrepancy was measured using Fink et al.'s (1983, p. 422) method. The subjects were asked, "How different is Judge Walters' decision about Convict X from your own view?" The subjects were told that a moderate degree of difference corresponded to a rating of 100. A rating of zero indicated no difference, and there was no upper bound to the scores.

To train the participants to properly use the magnitude scale, before the measure of psychological discrepancy the participants completed two practice tasks that were irrelevant to the topic of the study. In the first task, the participants were told that a moderate psychological distance between two colors was 100. Then, the participants were asked to indicate the distances between the colors of red and pink, between red and green, and between red and black with a positive number. The participants were told that zero represented "Not at all different," and there was no upper bound (see Fink, Monahan, & Kaplowitz, 1989). In the second task, the participants were told that a moderate hostility level between two countries was 100. Then, the participants were asked to indicate the hostility levels between the U.S. and Canada, between the U.S. and India, and between the U.S. and Russia with a positive number. The participants were told that zero represented "No hostility at all," and there was no upper bound (see Sulfaro & Crislip, 1997).

Final Belief Position

For subject's final belief position after reading the judge's sentence, subjects were asked, "To how many years in prison do you think Convict X should have been sentenced?"

Perceived Message Weight

Conceptually, weight represents the importance (Himmelfarb, 1975) or amount of information (Anderson, 2008, p. 43; Saltiel & Woelfel, 1975) of an incoming message or of one's prior beliefs. For an incoming message, the psychological meaning of weight can be its salience, relevance (Anderson, 1981), or informativeness (Fiske, 1980). To operationalize perceived message weight based on the conceptualization above, this variable has been tied to three variables in previous studies: message evaluation (Cacioppo, Petty, & Morris, 1983; Eagly & Telaak, 1972), attention (Fiske, 1980; Meffert, Chung, Joiner, Waks, & Garst, 2006), and perceived importance (Anderson & Alexander, 1971; Zalinski & Anderson, 1989).

The perceived message weight was measured by an importance rating (Anderson & Alexander, 1971; Zalinski & Anderson, 1989). The subjects were asked, "How important was Judge Walters' sentence when you decided what an appropriate sentence should be for Convict X?" This measure used a magnitude scale, where a moderate level was 100 with a lower bound of zero and no upper bound. The main study used two other ways to measure perceived message weight.

First, message evaluation is a composition of measures on specific features of the source and the content of a message (Cacioppo et al., 1983; Eagly & Telaak, 1972). It is assumed that the more positively a subject evaluates a feature of a message, the greater psychological weight the subject places on the message. For message content, the features include the effectiveness, the argument quality, and the fairness (Eagly & Telaak, 1972, p. 391) of a message, how compelling a message is, and how convincing a message is (Cacioppo et al., 1983, p. 808). For message source, the features include the unbiasedness, the credibility, the trustworthiness, and the expertness of the source. The measures of the above nine features all used magnitude scales. First, after the measure of psychological discrepancy, the subjects were told that the researcher wanted to know how they evaluated the judge's sentence. Then, the participants were asked, "How effectively did Judge Walters' statement make its point?"; "How well written was Judge Walters' statement?"; "How fair was Judge Walters' sentence of Convict X?"; "How compelling was Judge Walters' statement?"; "How convincing was Judge Walters' statement?"; "How unbiased was Judge Walters?"; "How credible was Judge Walters?"; "How expert was Judge Walters?"; "How trustworthy was Judge Walters?". For these nine magnitude scale measures, a moderate level is 100, with a lower bound of zero and no upper bound.

Second, following Kahneman (1973), attention can be defined as the allocation of cognitive effort. This dissertation assumes that the more attention a subject pays to a message, the more cognitive effort is allocated to the message, and the greater psychological weight the subject places on the message. Two previous studies operationalized attention to a message as the time a subject spent reading the message (Fiske, 1980; Meffert et al., 2006). The main study used the timing function in Qualtrics (2018) to measure how much time a subject spent on reading Judge Walters' sentence.

Perceived Message Scale Value

Three items were used for this dependent variable. The first item was similar to the measure of psychological discrepancy. The question asked,

> Think about your own view about the appropriate sentence for armed robbery after you read the Sentencing Guideline but before reading Judge Walters' decision. How different in years was Judge Walters' sentence of Convict X from your own view of the proper sentence? You have answered a similar question a moment ago with the scale where 100 represented a moderate difference. For the current question, please enter a positive number that represents a difference in years.

Note that whereas the psychological discrepancy measure used a magnitude scale (Fink et al., 1983), this measure used the belief position in years. So, the perceived message scale value could later be calculated based on subject's initial position.

The other two items used the same magnitude scale described above. The questions asked, "How harsh was Judge Walters' decision?" and "How punitive was Judge Walters' decision?" The rationale for these two measures was drawn from perspective theory (Ostrom & Upshaw, 1968, pp. 223-230). Perspective theory proposes that $R' = f(\frac{C - L}{U - L})$, where R' is the rating of a subject's attitude on a unidimensional scale (here, a leniency-harshness scale), C is the attitude content's position on another unidimensional scale (here, sentence in years), and R' is assumed to be a linear function of $\frac{C - L}{U - L}$ (1968, p. 224). Whereas perspective theory focuses on a subject's own attitude, this dissertation adapts perspective theory's logic and proposes $s_{\mathrm{p}} = \alpha R_{\mathrm{M}} = \alpha f(\frac{s - L}{U - L})$, where s_{p} is the perceived message scale value by a subject, α is a positive constant, and R_{M} is the subject's evaluation of the attitude rating of a message (e.g., how harsh a subject believes the judge's decision is). The critical assumption of this indirect measure is $s_{\mathrm{p}} = \alpha R_{\mathrm{M}}$, which indicates that the harsher a subject's evaluation of a judge's decision is (i.e., the greater the value of R_{M}), the greater the sentence the subject perceives the judge's decision (i.e., the greater the value of s_{p}).

CHAPTER 7

RESULTS

I used SPSS (2011) Version 20.0.0 and R (R Core Team, 2017) for data analysis. The logarithm and the anti-logarithm in the following text are to the base of the constant, e (i.e., the natural logarithm). The significance level (alpha) is .05 for hypothesis testing (two-tailed).

Missing Value Analysis

There were 14 variables that had missing data. For all of these 14 variables, the number of missing values was no more than two (i.e., no more than 2/448 = 0.45 percent). Little's (1988) missing completely at random (MCAR) test was conducted to examine the pattern of missing data. The null hypothesis of missing completely at random could not be rejected, χ^2 (314, N = 448) = 207.33, p = 1.00. Therefore, given the MCAR mechanism in the data, listwise deletion was used in the following analyses.

Subject's Initial Position

Subject's initial position, s_0, ranged from 0 to 100 years (M = 10.17, Mdn = 10.00, SD = 7.52, N = 448). The distribution of this variable had a skewness of 6.53 (SE = 0.12) and a kurtosis of 64.56 (SE = 0.23). The 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles were 3.45, 5.00, 7.00, 10.00, 10.00, 15.00, and 20.00. There were 388 subjects (86.61%) who reported an initial position between five and 15 years (inclusive). This percentage was close to the 88 percent reported in Kaplowitz and Fink (1991, p. 197). Therefore, as expected, the use of the sentencing guideline that recommended a 10-year sentence for an armed bank robbery did fix most subjects' initial positions at around 10 years.

Subject's initial position was winsorized at the 5th percentile and the 95th percentile. After winsorization, subject's s_0 ranged from 3.45 to 20 years ($M = 9.65$, $Mdn = 10.00$, $SD = 4.04$, $N = 448$). The distribution of this variable had a skewness of 0.80 ($SE = 0.12$) and a kurtosis of 0.65 ($SE = 0.23$). The winsorized variable was further transformed by taking the square root. The mean of the winsorized transformed variable was 3.04 ($SD = 0.64$). The skewness of the winsorized transformed variable was not significantly different from zero ($0.22/0.12 = 1.83 < z_{.025} = 1.96$); likewise, the absolute value of the kurtosis was not significantly different from zero ($|-0.09/0.23| = 0.39 < z_{.025} = 1.96$).

Manipulation Check

Subject's Upper Bound

The subject's upper bound, U, ranged from 1 to 110 years ($M = 23.83$, $Mdn = 20.00$, $SD = 12.68$, $N = 448$). The distribution of this variable had a skewness of 2.26 ($SE = 0.12$) and a kurtosis of 6.94 ($SE = 0.23$). The 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles were 10.00, 10.00, 15.00, 20.00, 30.00, 50.00, and 50.00. This variable was winsorized at the 95th percentile. After winsorization, subject's U ranged from 1 to 50 years ($M = 22.85$, $Mdn = 20.00$, $SD = 12.68$, $N = 448$). The distribution of this variable had a skewness of 1.04 ($SE = 0.12$) and a kurtosis of 0.16 ($SE = 0.23$). The winsorized variable was further transformed by taking the cube root. The mean of the winsorized transformed variable was 2.75 ($SD = 0.51$). The skewness of the winsorized transformed variable was not significantly different from zero ($0.21/0.12 = 1.75 < z_{.025} = 1.96$); Likewise, the kurtosis was not significantly different from zero ($0.12/0.23 = 0.52 < z_{.025} = 1.96$). Table 20 shows that after winsorization and transformation, in each one of the three upper bound conditions, the skewness and the kurtosis were greatly reduced.

To check the manipulation of *U*, I conducted a two-way analysis of covariance (ANCOVA) with *U* and the order of the questions asking the subject's *U* and *L* (*U*-then-*L* versus *L*-then-*U*) as independent variables, subject's initial position as a covariate, and subject's *U* as the dependent variable. The ANCOVA revealed a significant main effect of *U*, $F(2, 441) = 61.38$, $p < .001$, $\eta^2 = .218$, and a significant $U \times$ order interaction, $F(2, 441) = 3.73$, $p = .025$, $\eta^2 = .017$ (Figure 10). As expected, the linear increase was significant, $F(1, 441) = 122.54$, $p < .001$, $\eta^2 = .217$, whereas the quadratic effect was not significant, $F(1, 441) = 0.84$, $p = .362$. Examining subject's upper bound only, the manipulation of *U* was successful.

Subject's Lower Bound

The subject's lower bound, *L*, ranged from 0 to 100 years ($M = 9.52$, $Mdn = 6.00$, $SD = 10.80$, $N = 448$). The distribution of this variable had a skewness of 4.33 ($SE = 0.12$) and a kurtosis of 26.91 ($SE = 0.23$). This variable was winsorized at the 95th percentile (25 years). After winsorization, subject's *L* ranged from 1 to 25 years ($M = 8.50$, $Mdn = 6.00$, $SD = 6.36$, $N = 448$). The distribution of this variable had a skewness of 1.21 ($SE = 0.12$) and a kurtosis of 0.81 ($SE = 0.23$). The winsorized variable was further transformed by taking the square root. The mean of the winsorized transformed variable was 2.71 ($SD = 1.06$). The skewness of the winsorized transformed variable was

Table 20

*Main Study: Group Means (*SE*), Skewness Coefficients (*SE*), and Kurtosis Coefficients (*SE*) of Subject's Upper Bound With the Raw Data, the Winsorized Data, and the Transformed Winsorized Data*

		Raw data	Winsorized data	Cube root of winsorized data
U = 15 years *n* = 149	Mean	15.85 (0.47)	15.85 (0.47)	2.48 (0.02)
	Skewness	1.85 (0.20)	1.85 (0.20)	0.37 (0.20)
	Kurtosis	8.29 (0.40)	8.29 (0.40)	2.21 (0.40)
	SD	5.69	5.69	0.28
U = 30 years *n* = 151	Mean	24.22 (1.28)	23.03 (0.91)	2.76 (0.04)
	Skewness	2.75 (0.20)	0.70 (0.20)	-0.45 (0.20)
	Kurtosis	11.43 (0.39)	0.40 (0.39)	0.77 (0.39)
	SD	15.75	11.21	0.49
U = 50 years *n* = 148	Mean	31.48 (1.58)	29.72 (1.25)	3.00 (0.05)
	Skewness	1.24 (0.20)	0.29 (0.20)	-0.33 (0.20)
	Kurtosis	1.92 (0.40)	-1.43 (0.40)	-0.24 (0.40)
	SD	19.22	15.18	0.57

Note. The values are the point estimates of the statistics with their standard errors in the parentheses. Data winsorized at *U* = 50.

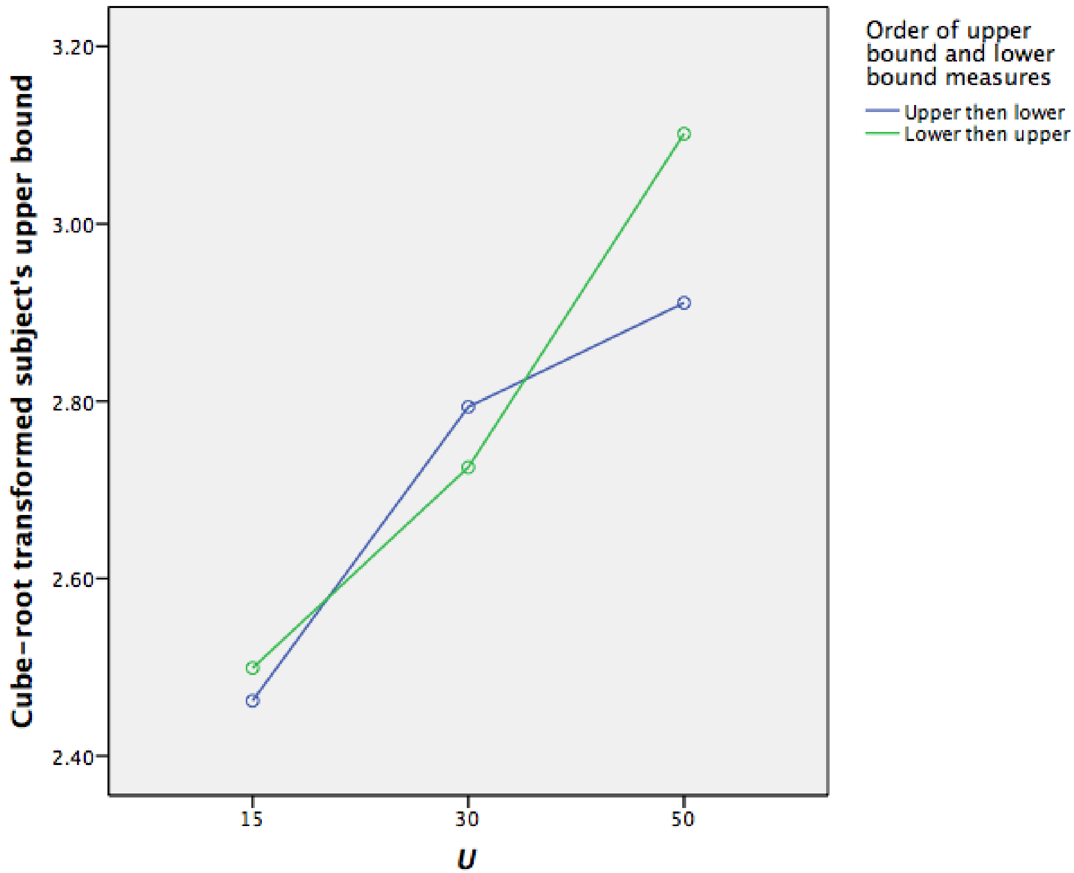


Figure 10. Two-way ANCOVA: Manipulation check of upper bound in the main study. Covariate: subject's initial position. $N = 448$.

greatly reduced to 0.36; Likewise, the absolute value of the kurtosis was not significantly different from zero ($|-0.10/0.23| = 0.43 < z_{.025} = 1.96$).

The manipulation of upper bound was not expected to affect subject's lower bound. This was not the case. The same two-way ANCOVA as the one described above but with lower bound as the dependent variable revealed a significant main effect of *U*, $F(2, 441) = 3.93$, $p = .020$, $\eta^2 = .018$, and a significant order effect, $F(2, 441) = 60.25$, $p < .001$, $\eta^2 = .120$ (Figure 11). The linear increase was significant, $F(1, 441) = 6.34$, $p < .012$, $\eta^2 = .014$, whereas the quadratic effect was not significant, $F(1, 441) = 1.48$, $p = .225$. However, with respect to effect size, the main effect of the upper bound manipulation on subject's upper bound ($\eta^2 = .218$) was much greater than its main effect on subject's lower bound ($\eta^2 = .018$). The effect size of the linear trend in the subject's upper bound ($\eta^2 = .217$) was also greater than the linear trend in subject's lower bound ($\eta^2 = .014$). Thus, whereas the manipulation of upper bound also increased subject's lower bound, the magnitude of the increase was less than the increase in subject's upper bound due to the manipulation of upper bound.

Subject's Perspective

To examine whether subject's perspective ($U - L$) increased as a result of the manipulation of *U*, the subject's lower bound was subtracted from the subject's upper bound to create the subject's perspective. There were 12 subjects who had a negative value for perspective. Neither the manipulation of upper bound nor the measure order seemed to result in a negative value in a systematic way. Among these 12 subjects, two were in the 15 years and *U*-then-*L* group; two were in the 15 years and *L*-then-*U* group;

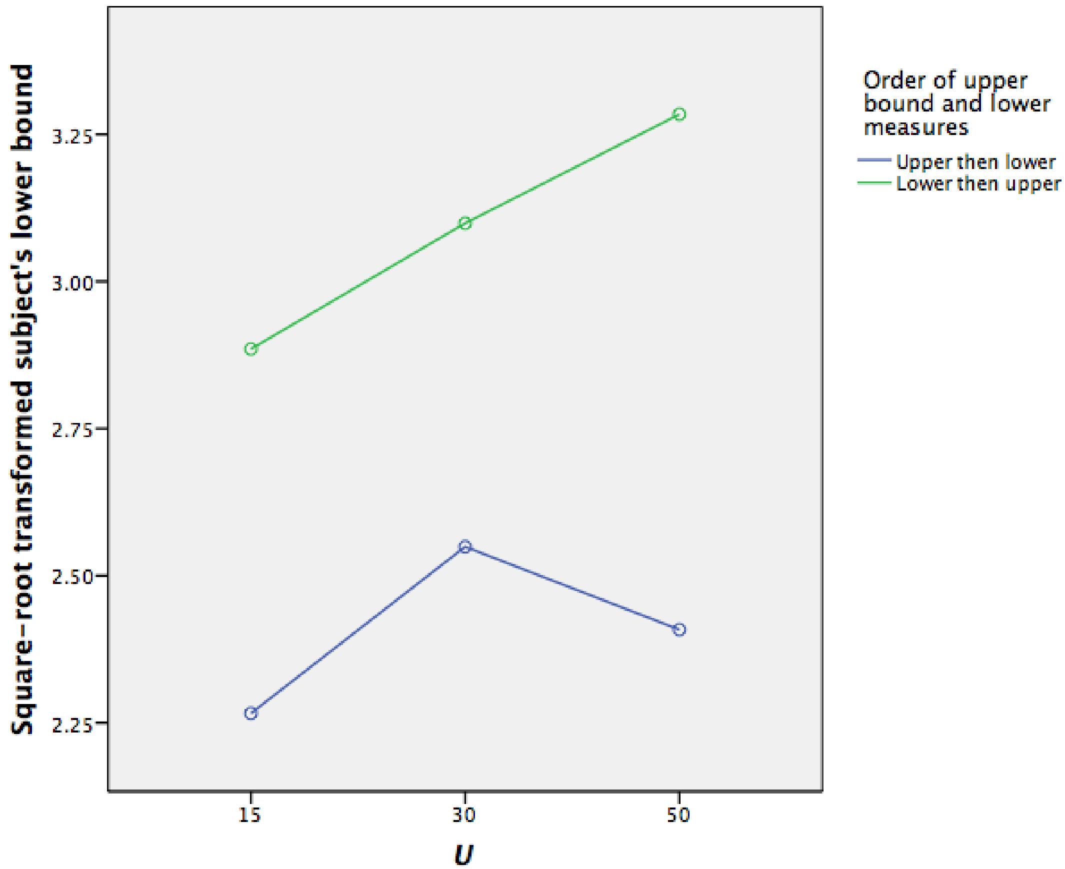


Figure 11. Two-way ANCOVA: The effect of the manipulation of upper bound on subject's lower bound in the main study. Covariate: subject's initial position. $N = 448$.

three were in the 30 years and *U*-then-*L* group; three were in the 30 years and *L*-then-*U* group; one was in the 50 years and *U*-then-*L* group; one was in the 50 years and *L*-then-*U* group. These 12 subjects were excluded in the following analysis. On the remaining 436 subjects, standard data winsorization and transformation were performed on the subject's perspective to approximate a normal distribution. The square root of the winsorized variable was used (M = 3.48, SD = 1.57, Sk = 0.07 [SE = 0.12], Ku = 0.20 [SE = 0.23], n = 436). The same two-way ANCOVA as the one described above but with perspective as the dependent revealed a significant main effect of U, $F(2, 429) = 33.14$, $p < .001$, η^2 = .134, and a significant main order effect, $F(2, 429) = 5.83$, $p = .016$, $\eta^2 = .013$ (Figure 12). The linear increase was significant, $F(1, 429) = 66.27$, $p < .001$, $\eta^2 = .134$, whereas the quadratic effect was not significant, $F(1, 429) = 0.003$, $p = .959$.

Psychological Discrepancy

Psychological discrepancy, ψ, ranged from 0 to 100,000 (M = 385.47, Mdn = 120.00, SD = 4,720.37, N = 448). The distribution of this variable had a skewness of 21.12 (SE = 0.12) and a kurtosis of 446.67 (SE = 0.23). This variable was winsorized at the 95th percentile (500). After winsorization, ψ ranged from 0 to 500 (M = 149.31, Mdn = 120.00, SD = 130.62, N = 448). The distribution of this variable had a skewness of 1.12 (SE = 0.12) and a kurtosis of 0.79 (SE = 0.23). The winsorized variable was further transformed by taking the square root. The mean of the transformed and winsorized variable was 10.73 (SD = 5.86). The absolute value of the kurtosis of the transformed winsorized variable was reduced to 0.46; the absolute value of the skewness was not significantly different from zero ($|-0.06/0.12| = 0.50 < z_{.025} = 1.96$). Table 21 shows that

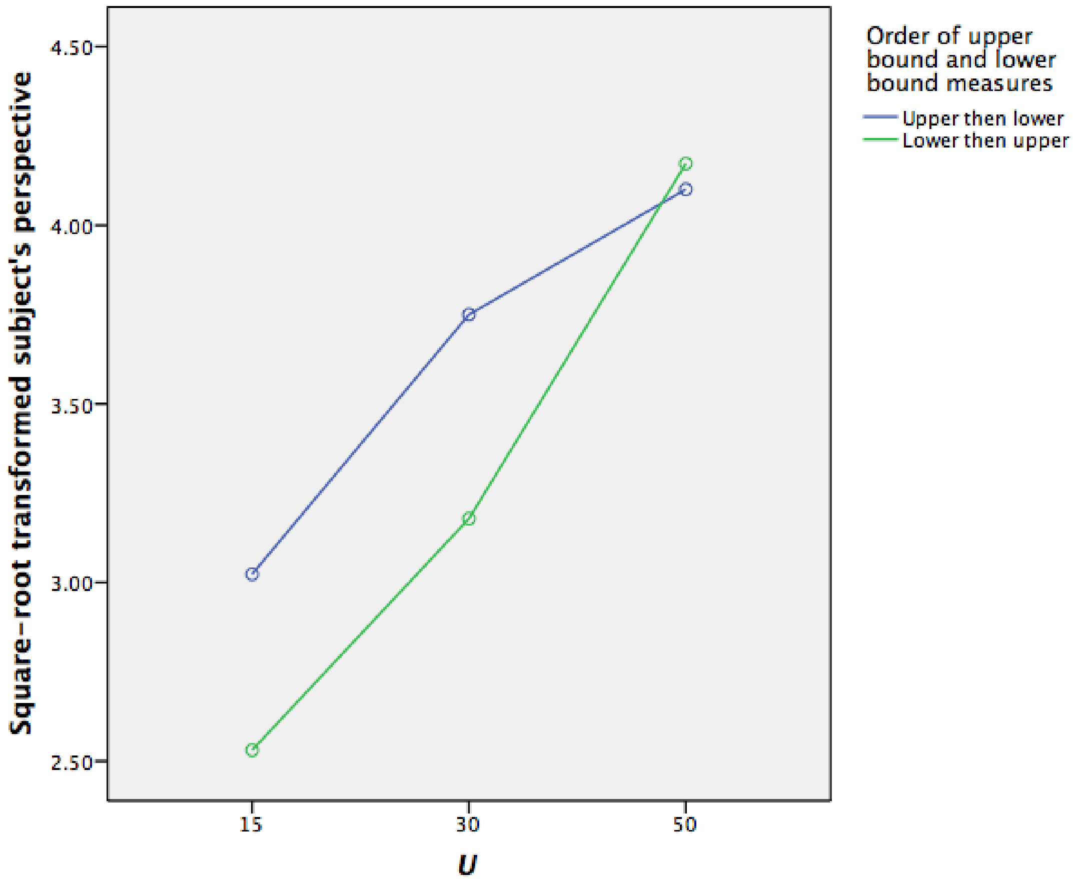


Figure 12. Two-way ANCOVA: The effect of the manipulation of upper bound on subject's perspective in the main study. Covariate: subject's initial position. n = 436.

Table 21

*Main Study: Group Means (*SE*), Skewness Coefficients (*SE*), and Kurtosis Coefficients (*SE*) of Psychological Discrepancy With the Raw Data, the Winsorized Data, and the Transformed Winsorized Data*

Condition		Raw data	Winsorized data	Square root of winsorized data
		U = 15 years		
s = 15 years *n* = 49	Mean	55.20 (8.50)	55.20 (8.50)	5.64 (0.70)
	Skewness	0.93 (0.34)	0.93 (0.34)	0.06 (0.34)
	Kurtosis	0.07 (0.67)	0.07 (0.67)	-1.42 (0.67)
	SD	59.47	59.47	4.89
s = 30 years *n* = 50	Mean	198.90 (27.95)	178.90 (17.98)	12.47 (0.69)
	Skewness	2.87 (0.34)	1.20 (0.34)	0.01 (0.34)
	Kurtosis	9.66 (0.66)	1.27 (0.66)	0.46 (0.66)
	SD	197.66	127.15	4.89
s = 50 years *n* = 50	Mean	212.92 (24.05)	202.92 (18.89)	13.36 (0.71)
	Skewness	2.23 (0.34)	0.70 (0.34)	-0.14 (0.34)
	Kurtosis	8.37 (0.66)	0.02 (0.66)	-0.42 (0.66)
	SD	170.03	133.59	5.00
		U = 30 years		
s = 15 years *n* = 50	Mean	89.42 (15.13)	89.42 (15.13)	7.93 (0.74)
	Skewness	2.17 (0.34)	2.17 (0.34)	0.67 (0.34)
	Kurtosis	4.87 (0.66)	4.87 (0.66)	0.61 (0.66)
	SD	106.97	106.97	5.21
s = 30 years *n* = 51	Mean	201.27 (28.67)	177.75 (17.83)	12.42 (0.69)
	Skewness	2.71 (0.33)	1.12 (0.33)	0.04 (0.33)
	Kurtosis	8.25 (0.66)	0.86 (0.66)	0.23 (0.66)
	SD	204.75	127.36	4.90
s = 50 years *n* = 50	Mean	249.20 (35.73)	211.20 (22.61)	13.08 (0.90)
	Skewness	1.78 (0.34)	0.51 (0.34)	-0.35 (0.34)
	Kurtosis	3.05 (0.66)	-0.88 (0.66)	-0.60 (0.66)
	SD	252.66	159.89	6.39

Table 21

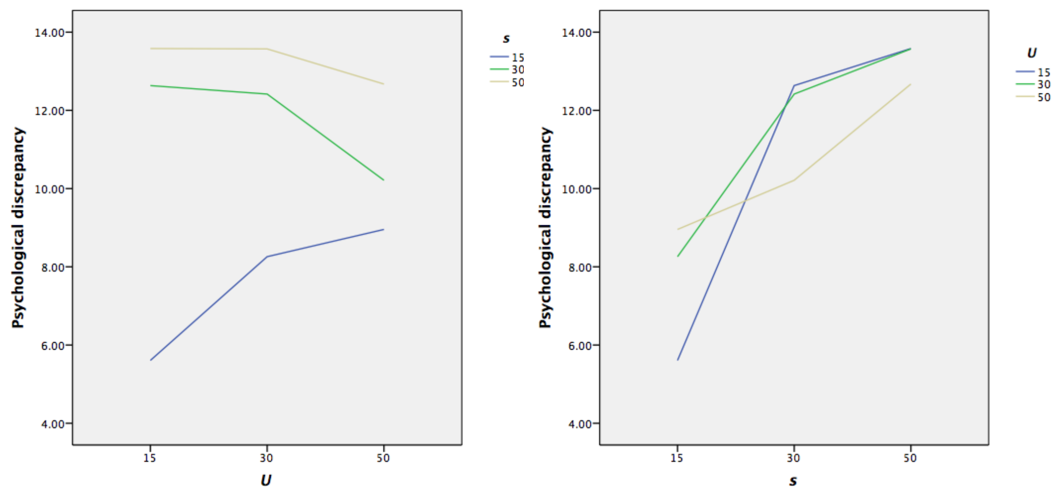
(continued)

Condition		Raw data	Winsorized data	Square root of winsorized data
		U = 50 years		
s = 15 years n = 49	Mean	139.80 (31.76)	113.27 (18.58)	8.79 (0.87)
	Skewness	3.05 (0.34)	1.83 (0.34)	0.43 (0.34)
	Kurtosis	9.32 (0.67)	3.15 (0.67)	-0.09 (0.67)
	SD	222.34	130.05	6.06
s = 30 years n = 49	Mean	2171.98 (2038.19)	133.20 (16.85)	10.00 (0.83)
	Skewness	7.00 (0.34)	1.21 (0.34)	-0.17 (0.34)
	Kurtosis	48.99 (0.67)	1.97 (0.67)	-0.48 (0.67)
	SD	14267.31	117.95	5.81
s = 50 years n = 50	Mean	178.40 (14.76)	178.40 (14.76)	12.67 (0.60)
	Skewness	0.69 (0.34)	0.69 (0.34)	-0.41 (0.34)
	Kurtosis	0.61 (0.66)	0.61 (0.66)	0.63 (0.66)
	SD	104.37	104.37	4.26

Note. N = 448. The values are the point estimates of the statistics with their standard errors in the parentheses. Data winsorized at ψ = 500.

after winsorization and transformation, in each group, the skewness and the kurtosis were greatly reduced as compared to the raw data.

To check the manipulation of message scale value, s, I conducted a two-way ANCOVA with U and s as independent variables, subject's initial position as a covariate, and ψ as the dependent variable. Because a negative value of perspective indicated a subject's failure in understanding the questions asking about the most lenient sentence and the harshest sentence, the analysis excluded 12 cases that had a negative value of perspective (one subject in the $U = 15$ years and $s = 15$ years group, two in the $U = 15$ years and $s = 30$ years group, one in the $U = 15$ years and $s = 50$ years group, two in the $U = 30$ years and $s = 15$ years group, four in the $U = 30$ years and $s = 50$ years group, one in the $U = 50$ years and $s = 15$ years group, and one in the $U = 50$ years and $s = 30$ years group). Including these 12 cases returned similar results as those reported below. There was a significant main effect of s, $F(2, 426) = 57.19$, $p < .001$, $\eta^2 = .212$, and a marginally significant $U \times s$ interaction, $F(4, 426) = 2.39$, $p = .051$, $\eta^2 = .022$ (see Panel b in Figure 13). The linear increase of ψ as a function of s was significant, $F(1, 426) = 107.26$, $p < .001$, $\eta^2 = .201$, and the quadratic effect of s was also significant, $F(1, 426) = 7.23$, $p = .007$, $\eta^2 = .017$. The significant quadratic effect of s was not due to a downturn in ψ as s increased (see Panel a in Figure 13). The manipulation of message scale value was successful.



(a) (b)

Figure 13. Two-way ANCOVA: The effect of the manipulation of message scale value on psychological discrepancy in the main study. The y-axis represents the predicted value of the ANCOVA model based on the square root of the transformed psychological discrepancy. Covariate: subject's initial position. $n = 436$. In Panel a, the x-axis represents the upper bound (U), and each line represents one of the message scale value (s) conditions; in Panel b, the x-axis represents the message scale value (s), and each line represents one of the upper bound (U) conditions.

Reliability of the Measures for Perceived Weight and Perceived Scale Value

Perceived Message Weight

There were 11 items that measured perceived message weight. Because the 11 items were all positively skewed, I winsorized all of them at the 95th percentile. After winsorization, except for the time spent of reading the decision, the magnitudes of skewness and kurtosis appeared acceptable. For the time spent of reading the decision, I further took the logarithm of the winsorized variable (see Table 22 for the descriptive statistics for the 11 items after data transformation). The five-item measure for content evaluation was reliable (Cronbach's α = .91). The four-items for source evaluation was reliable (Cronbach's α = .86). By combining the measures for source and content evaluation as well as the importance measure, the 10-item Cronbach's α was .92. By including the measure for attention, the 11-item Cronbach's α was .91. A further check on the correlations among the 11 items revealed that except for attention, all 10 items correlated positively with each other with a p value less than .001. The greatest Pearson correlation was between compelling and convincing, $r = .89$, whereas the smallest Pearson correlation was between important and unbiased, $r = .25$. Attention significantly correlated with expert, $r = .11$, $p = .018$, trustworthy, $r = .13$, $p = .004$, compelling, $r = .10$, $p = .045$, convincing, $r = .13$, $p = .005$, and important, $r = .10$, $p = .028$. By examining the magnitude or r, it seemed that attention differed from the other 10 items measuring perceived message weight.

Table 22

Main Study: Descriptive Statistics for the 11 Items Measuring the Perceived Message Weight

		M	*SD*	Skewness	Kurtosis	Number of missing cases
Content evaluation (Cronbach's α = .91, *n* = 444)	Fair	79.62	60.52	0.62 (0.12)	-0.44 (0.23)	0
	Effective	135.16	75.74	0.43 (0.12)	-0.35 (0.23)	2
	Well written	130.79	61.84	0.14 (0.12)	-0.82 (0.23)	1
	Compelling	114.45	65.29	0.34 (0.12)	-0.72 (0.23)	1
	Convincing	113.48	66.36	0.35 (0.12)	-0.71 (0.23)	1
Content evaluation (Cronbach's α = .86, *n* = 447)	Unbiased	86.81	58.43	0.47 (0.12)	-0.50 (0.23)	0
	Credible	108.46	59.22	0.17 (0.12)	-0.86 (0.23)	0
	Expert	120.33	62.88	0.27 (0.12)	-0.63 (0.23)	0
	Trustworthy	112.13	62.75	0.20 (0.12)	-0.82 (0.23)	1
	Important	96.08	63.43	0.27 (0.12)	-0.87 (0.23)	0
	Time spent*	3.56	0.47	0.96 (0.12)	0.24 (0.23)	0

Note. All items were winsorized at the 95th percentile first. *N* = 448 (if without missing cases).

*Item was winsorized and then transformed logarithmically (base *e*).

To corroborate the observation that the item for attention differed from the other 10 items of the perceived weight measure, exploratory factor analysis (EFA) was conducted. A preliminary principal components analysis (PCA) with 11 components revealed that the first component explained 55% of the total variance with an eigenvalue of 6.06. The second principal component explained 9.7% of the total variance with an eigenvalue of 1.07. The third principal component explained 8.6% of the total variance with an eigenvalue of 0.95. Next, an EFA using the maximum likelihood extraction and the direct-oblimin rotation with *two* factors was conducted. I used this oblimin rotation, because I assumed that the two factors (e.g., the content evaluation and the source evaluation on a judge's decision) would be correlated. The two-factor EFA did not indicate a good fit, $\chi^2(34, 443) = 154.92$, $p < .001$. Neither did the three-factor EFA indicate a good fit, $\chi^2(25, 443) = 72.95$, $p < .001$, nor the four-factor EFA, $\chi^2(17, 443) = 29.48$, $p = .030$.

Table 23 shows the factor loadings controlled for factor correlations (see Table 24) based on the four-factor solution. The loadings on the first two factors showed a neat structure. Except for fairness, the four items measuring content evaluation loaded highly on the first factor, whereas the four items measuring source evaluation, together with fairness, loaded highly on the second factor. Fairness loaded highly by itself on the third factor, whereas expertness and trustworthiness loaded highly on the fourth factor. Importance loaded moderately on the first factor and the second factor. The overall loading pattern reflected the intended measurement scheme for the perceived message weight, except that the item for attention (i.e., the *time spent* item in Table 23) did not

Table 23

Main Study: Factor Loadings Controlled for Factor Correlation Based on an EFA With Maximum Likelihood Extraction and Direct-Oblimin Rotation for the 11 Items Measuring the Perceived Message Weight

Item	Content evaluation	Source evaluation I	Fairness	Source evaluation II
Unbiased	.01	**.66**	-.03	.17
Credible	.13	**.74**	-.01	-.11
Expert	.29	**.44**	.13	**-.40**
Trustworthy	.02	**.64**	-.001	**-.52**
Well written	**.80**	.02	.09	-.07
Effective	**.81**	.06	.15	.06
Compelling	**.93**	-.07	-.14	-.02
Convincing	**.82**	.01	-.26	-.05
Fair	.13	**.42**	**-.43**	-.22
Important	.25	.25	-.10	-.05
Time spent	.02	-.03	-.02	-.20

Note. Absolute factor loadings ≥ 0.40 are in boldface. $n = 443$.

Table 24

Main Study: Factor Correlation Matrix Based on an EFA With Maximum Likelihood Extraction and Direct-Oblimin Rotation for the 11 Items Measuring the Perceived Message Weight

Factor	Content evaluation	Source evaluation I	Fairness	Source evaluation II
Content evaluation	—			
Source evaluation I	.62	—		
Fairness	-.25	-.22	—	
Source evaluation II	-.53	-.29	.25	—

Note. n = 443.

load highly on either of the factors. Therefore, the item for attention was excluded from the following analysis.

Excluding the item for attention, an EFA based on the remaining 10 items using the maximum likelihood extraction and the direct-oblimin rotation with four factors was conducted. The EFA indicated a good fit, $\chi^2(11, 443) = 18.45$, p = .072. Table 25 shows the factor loadings controlled for factor correlations (see Table 26) based on the four-factor solution excluding the item for attention. To be consistent with the above EFA that included the attention item for the sake of comparison and interpretation, I labeled the factor that had the two high loadings from expertness and trustworthiness as *source evaluation II*. By comparing Table 24 and Table 26, I found that the only great discrepancy regarding factor correlation was between *source evaluation I* and

Table 25

Main Study: Factor Loadings Controlled for Factor Correlation Based on an EFA With Maximum Likelihood Extraction and Direct-Oblimin Rotation for the 10 Items Measuring the Perceived Message Weight

Item	Content evaluation	Source evaluation II	Source evaluation I	Fairness
Unbiased	-.03	.01	**.61**	-.004
Credible	.10	-.18	**-.70**	.01
Expert	.24	**-.63**	.08	.08
Trustworthy	-.06	**-.97**	.05	-.05
Well written	**.75**	-.18	-.05	.10
Effective	**.75**	.03	.13	.17
Compelling	**.92**	-.004	-.03	-.13
Convincing	**.83**	-.03	.02	-.26
Fair	.18	-.26	.25	**-.44**
Important	.26	-.08	.20	-.09

Note. Absolute factor loadings ≥ 0.40 are in boldface. n = 443.

Table 26

Main Study: Factor Correlation Matrix Based on an EFA With Maximum Likelihood Extraction and Direct-Oblimin Rotation for the 10 Items Measuring the Perceived Message Weight

Factor	Content evaluation	Source evaluation II	Source evaluation I	Fairness
Content evaluation	—			
Source evaluation II	-.72	—		
Source evaluation I	.58	-.76	—	
Fairness	-.22	.29	-.25	—

Note. n = 443.

source evaluation II (i.e., -.29 in Table 24 but -.76 in Table 26). The factor loadings in Table 25 revealed a similar pattern to those in Table 23. Again, the overall loading pattern reflected the intended measurement scheme for the perceived message weight.

Perceived Message Scale Value

Based on the perceived difference (in years) between the judge's sentence and a subject's initial position (i.e., perceived difference = perceived judge's sentence – initial position), I could calculate the perceived message scale value by summing the perceived difference and the initial position (i.e., perceived judge's sentence = perceived difference + initial position). This calculated variable was winsorized at the 95th percentile and then transformed by taking the logarithm (M = 3.65, SD = 0.90; Mdn = 3.40; Sk = 0.51 [SE = 0.12], Ku = -0.48 [SE = 0.23]). The harshness item and the punitiveness item were winsorized at the 95th percentile and then transformed by taking the square root (harshness: *M* = 12.07, *SD* = 5.51; *Mdn* = 12.25; *Sk* = -0.08 [*SE* = 0.12], *Ku* = -0.07 [*SE* = 0.23]; punitiveness: *M* = 12.36, *SD* = 4.98; *Mdn* = 12.25; *Sk* = 0.12 [*SE* = 0.12], *Ku* = 0.21 [*SE* = 0.23]). The three-item measure was not sufficiently reliable, Cronbach's α = .68. By excluding the item for the perceived judge's sentence in years, the two-item reliability was high, with a Spearman-Brown statistic of .86 (see Eisinga, te Grotenhuis, & Pelzer, 2013 for the rationale to use the Spearman-Brown statistic for assessing the reliability of a two-item scale). Therefore, the item for the perceived judge's sentence in years was excluded from the following analysis.

Hypothesis Testing

Final Position

Final position, *R*, ranged from 0 to 100 (*M* = 14.30, *Mdn* = 10.00, *SD* = 10.97, *N* = 448). The distribution of this variable had a skewness of 3.58 (*SE* = 0.12) and a kurtosis of 19.87 (*SE* = 0.23). This variable was winsorized at the 95th percentile (30 years). After winsorization, *R* ranged from 0 to 30 (*M* = 13.30, *Mdn* = 10.00, *SD* = 7.05, *N* = 448). The distribution of this variable had a skewness of 0.94 (*SE* = 0.12) and a kurtosis of 0.36 (*SE* = 0.23). The winsorized variable was further transformed by taking the square root. The mean of the winsorized transformed variable was 3.52 (*SD* = 0.96). The kurtosis of the winsorized transformed variable was not significantly different from zero (0.40/0.23 = 1.74 < $z_{.025}$ = 1.96); the skewness was not significantly different from zero (0.13/0.12 < $z_{.025}$ = 1.96). Table 27 shows that after winsorization and transformation, in each group, the skewness and the kurtosis were greatly reduced as compared to the raw data.

Table 27

*Main Study: Group Means (*SE*), Skewness Coefficients (*SE*), and Kurtosis Coefficients (*SE*) of Final Position With the Raw Data, the Winsorized Data, and the Transformed Winsorized Data*

Condition		Raw data	Winsorized data	Square root of winsorized data
U = 15 years				
s = 15 years *n* = 49	Mean	12.35 (0.79)	12.14 (0.65)	3.43 (0.09)
	Skewness	2.62 (0.34)	1.18 (0.34)	0.18 (0.34)
	Kurtosis	12.54 (0.67)	3.88 (0.67)	1.77 (0.67)
	SD	5.51	4.57	0.65
s = 30 years *n* = 50	Mean	12.61 (1.01)	12.61 (1.01)	3.42 (0.14)
	Skewness	1.10 (0.34)	1.10 (0.34)	0.41 (0.34)
	Kurtosis	0.87 (0.66)	0.87 (0.66)	0.04 (0.66)
	SD	7.12	7.12	0.97
s = 50 years *n* = 50	Mean	12.66 (1.22)	12.06 (0.92)	3.33 (0.14)
	Skewness	2.34 (0.34)	0.93 (0.34)	-0.42 (0.34)
	Kurtosis	7.82 (0.66)	1.07 (0.66)	1.82 (0.66)
	SD	8.62	6.51	1.00
U = 30 years				
s = 15 years *n* = 50	Mean	10.92 (0.61)	10.92 (0.61)	3.24 (0.09)
	Skewness	0.59 (0.34)	0.59 (0.34)	-0.16 (0.34)
	Kurtosis	1.17 (0.66)	1.17 (0.66)	0.56 (0.66)
	SD	4.30	4.30	0.67
s = 30 years *n* = 51	Mean	16.43 (2.12)	14.51 (1.13)	3.65 (0.15)
	Skewness	3.70 (0.33)	0.51 (0.33)	0.05 (0.33)
	Kurtosis	18.49 (0.66)	-0.76 (0.66)	-0.89 (0.66)
	SD	15.15	8.07	1.08
s = 50 years *n* = 50	Mean	15.02 (1.49)	13.92 (1.07)	3.61 (0.13)
	Skewness	1.99 (0.34)	1.15 (0.34)	0.80 (0.34)
	Kurtosis	3.78 (0.66)	0.14 (0.66)	-0.25 (0.66)
	SD	10.54	7.58	0.95

Table 27

(continued)

Condition		Raw data	Winsorized data	Square root of winsorized data
		U = 50 years		
s = 15 years n = 49	Mean	10.37 (0.71)	10.37 (0.71)	3.12 (0.12)
	Skewness	0.24 (0.34)	0.24 (0.34)	-0.20 (0.34)
	Kurtosis	-0.81 (0.67)	-0.81 (0.67)	-0.84 (0.67)
	SD	4.97	4.97	0.81
s = 30 years n = 49	Mean	18.43 (1.67)	17.20 (1.15)	4.01 (0.15)
	Skewness	2.20 (0.34)	0.32 (0.34)	-0.82 (0.34)
	Kurtosis	7.50 (0.67)	-0.82 (0.67)	2.45 (0.67)
	SD	11.68	8.02	1.08
s = 50 years n = 50	Mean	19.82 (2.57)	15.92 (1.17)	3.85 (0.15)
	Skewness	2.55 (0.34)	0.53 (0.34)	-0.05 (0.34)
	Kurtosis	7.70 (0.66)	-0.85 (0.66)	-0.28 (0.66)
	SD	18.19	8.28	1.07

Note. N = 448. The values are the point estimates of the statistics with their standard errors in the parentheses. Data winsorized at R = 30.

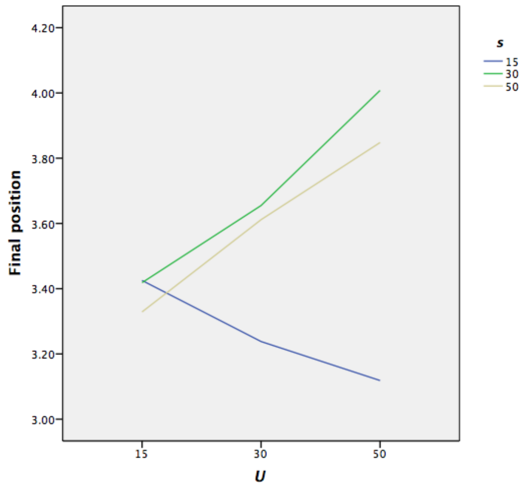
To test H1 and H2, I conducted a two-way ANCOVA with U and s as independent variables, subject's initial position as a covariate, and R as the dependent variable (Figure 14). There was a significant main effect of s, $F(2, 438) = 12.53$, $p < .001$, $\eta^2 = .054$, and a significant main effect of U, $F(2, 438) = 2.40$, $p = .004$, $\eta^2 = .025$. The $U \times s$ interaction effect was not significant, $F(4, 438) = 1.95$, $p = .101$, $\eta^2 = .017$, which did not support H2b.

The linear increase of R as a function of s was significant, $F(1, 438) = 12.27$, $p < .001$, $\eta^2 = .027$, and the quadratic effect of s was also significant, $F(1, 438) = 12.90$, $p < .001$, $\eta^2 = .029$. H2a was supported. The linear increase of R as a function of U was significant, $F(1, 438) = 11.06$, $p = .001$, $\eta^2 = .025$, and the quadratic effect of U was not significant, $F(1, 438) = 0.04$, $p = .842$. Therefore, H1a was supported.

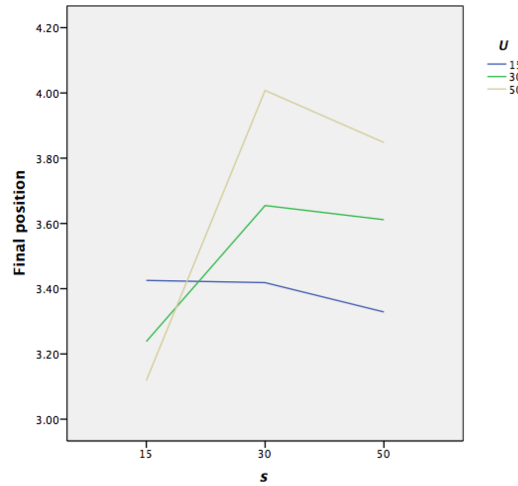
Boomerang Effect

H3a posits that the mean of the final position of a subject's belief, R, is greater than the mean of the initial position of the same belief, s_0. Out of the 448 subjects, there were 15 subjects (3.35%) who reported a final position less than an initial position (i.e., $R < s_0$). On average, there was a significant increase in the subject's position after reading the judge's decision ($M = 3.52$, $SD = 0.96$) than before reading the judge's decision ($M = 3.04$, $SD = 0.64$), $t(447) = 14.79$, $p < .001$, $d = 0.70$. Therefore, H3a was supported.

H3b posits that the proportion of cases that have a boomerang effect in the most extreme message scale value condition is greater than the proportions in the less extreme message scale value conditions. The numbers of subjects who had a boomerang effect were 7 (4.73%), 3 (2.00%), 5 (3.33%) in the $s = 15$ years ($n = 148$),



(a)



(b)

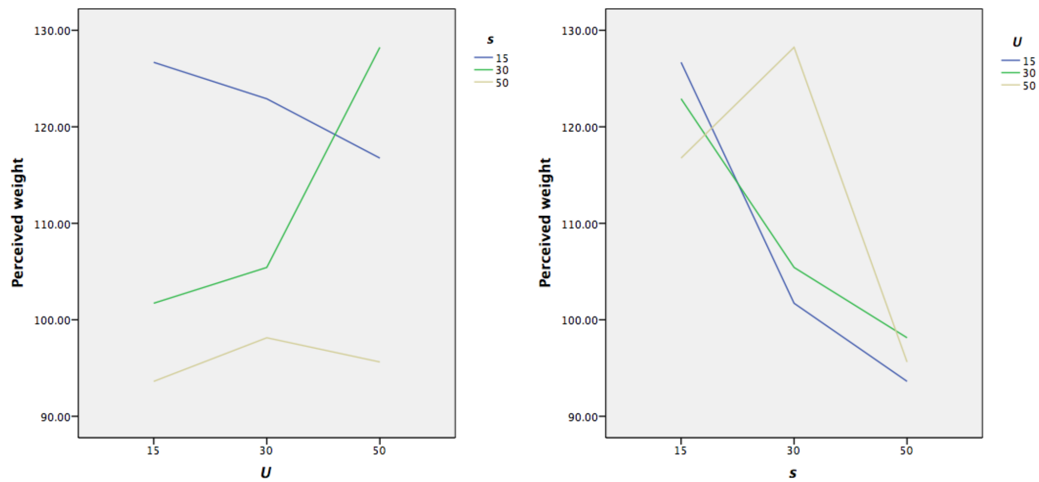
Figure 14. Two-way ANCOVA on final position in the main study. The y-axis represents the predicted value of the ANCOVA model based on the square root of the transformed final position. Covariate: subject's initial position. $N = 448$. In Panel a, the x-axis represents the upper bound (U), and each line represents one of the message scale value (s) conditions; in Panel b, the x-axis represents the message scale value (s), and each line represents one of the upper bound (U) conditions.

s = 30 years (n = 150), and s = 50 years (n = 150) groups, respectively. The proportion of subjects who reported a final position less than an initial position did not differ by message scale value, $\chi^2(2, N = 448) = 1.72$, $p = .424$. H3b was not supported.

Perceived Message Weight

Before formally testing H4, I conducted a two-way ANCOVA with U and s as independent variables, subject's initial position as a covariate, and the 10-item average of perceived weight (M = 109.77, SD = 49.08, Sk = 0.15 [SE = 0.12], Ku = -0.67 [SE = 0.23], n = 443) as the dependent variable (Figure 15). Listwise deletion was applied given the MCAR mechanism (n = 443). The only significant effect was the main effect of s, $F(2, 433) = 13.44$, $p < .001$, $\eta^2 = .058$. The linear decrease of perceived weight as a function of s was significant, $F(1, 433) = 26.58$, $p < .001$, $\eta^2 = .058$, and the quadratic effect of s was not significant, $F(1, 433) = 0.26$, $p = .612$.

The equation, $w_p = w\Delta(\psi) = we^{-\gamma\psi}$, where $\psi = kD/P$ and w_p represents perceived message weight, was predicted by the weight-discounting model and the complex model. I tested this equation directly in order to test H4. First, by taking the natural logarithm on both sides of the equation to linearize the relationship, the equation in question became $\ln(w_p) = \ln(w) - \gamma\psi$. Because $\psi = kD/P$, a test of the linearized equation could be done by testing this linear regression model: $\ln(w_p) = a + b_1D + b_2(1/P) + b_3(D/\mathrm{P}) + \epsilon$ (see Blanton & Jaccard, 2006, for a recommended way to test this multiplicative model). If $w_p = w\Delta(\psi) = we^{-\gamma\psi}$ and $\psi = kD/P$ are plausible, the estimated b_1, b_2, and b_3 would all be negative, and b_3 would be significantly different from zero.



(a) (b)

Figure 15. Two-way ANCOVA on perceived weight in the main study. The y-axis represents the predicted value of the ANCOVA model based on the 10-item average of perceived weight. Covariate: subject's initial position. *n* = 443. In Panel a, the x-axis represents the upper bound (*U*), and each line represents one of the message scale value (*s*) conditions; in Panel b, the x-axis represents the message scale value (*s*), and each line represents one of the upper bound (*U*) conditions.

Thirty-six subjects with a nonpositive value of perspective, twelve subjects with a negative value of message discrepancy, and five subjects with a missing value in one of the perceived-weight items were excluded. There were 398 subjects left for analysis. Next, perspective was transformed to its inverse (i.e., $1/P$) and mean centered. Message discrepancy was also mean centered. Then, the natural logarithm of the 10-item average of perceived weight was regressed on the centered inverse of perspective, the centered message discrepancy, and the product of these two centered variables. The results of the regression indicated that the three predictors explained 13.83% of the variance (Adjusted $R^2 = .132$, $F[3, 394] = 21.09$, $p < .001$). Although both estimates of b_1 and b_2 were negative and significant, the estimated coefficient of the multiplicative term was not significant (see Table 28), which indicated that the functional form in $\psi = kD/P$ was implausible. Without including the multiplicative term, the results of the regression indicated that the two predictors explained 13.83% of the variance, Adjusted $R^2 = .134$, $F(2, 395) = 31.69$, $p < .001$, which indicated that a more parsimonious additive model of psychological discrepancy would have been more plausible than the assumed multiplicative model of psychological discrepancy. In sum, the linear regression results supported H4a, because the logarithm of perceived weight was found to be a decreasing function of the inverse of perspective. However, the support for the weight-discounting model and the complex model was inconclusive, because the process (i.e., the multiplicative model) hypothesized to lead to the positive relationship between perceived weight and perspective was not found to be the case (i.e., H4b was not supported).

Table 28

Main Study: Summary of Linear Regression Analysis for Variables Predicting Perceived Weight

	Unstandardized coefficient	*SE*	Standardized coefficient	Semipartial correlation
Intercept	4.61***	0.03		
Message discrepancy	-0.01***	0.002	-.25***	-.25
Inverse of perspective	-1.97***	0.34	-.28***	-.28
Multiplicative term	-.01	0.02	-.01	-.01

Note. $n = 398$. Model fit: $R^2 = .138$, Adjusted $R^2 = .132$, $F(3, 394) = 21.09$, $p < .001$.

*$p < .05$. **$p < .01$. ***$p < .001$.

H4c predicts that the inverse of perspective has no effect on perceived message weight, and H4d predicts that the multiplicative term of message discrepancy and the inverse of perspective has no effect on the perceived message weight. Because when testing H4a the null hypothesis H_0: $b_2 \neq 0$ was rejected, H4c was rejected. However, when testing H4b the failure to reject H_0: $b_3 \neq 0$ did not indicate the acceptance of H4d, equivalence testing (Levine, Weber, Park, & Hullett, 2008; Weber & Popova, 2012) was conducted to test H4d.

In this equivalence test, first, the measure of effect size of a regression coefficient was chosen to be the semipartial correlation, r_{sp}, between the predictor of concern and the dependent variable (Levine et al., 2008). The semipartial correlation is the increase in R^2

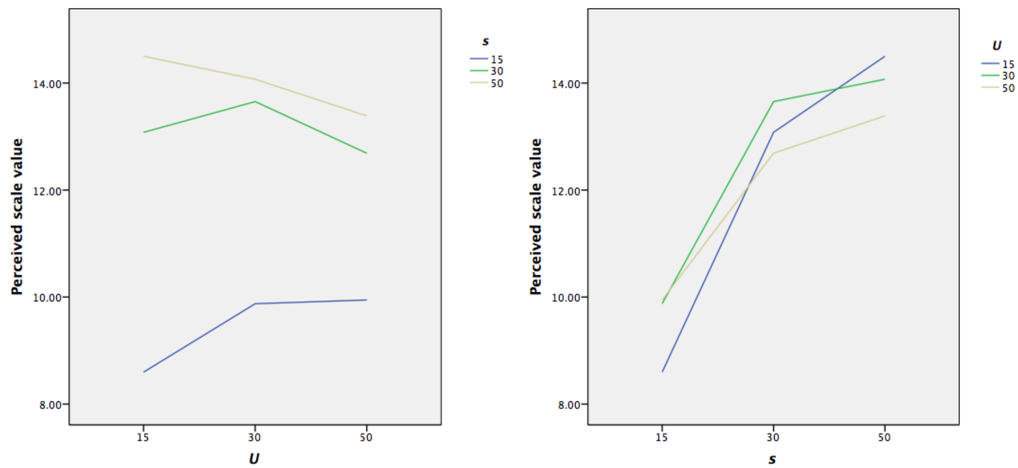
from the model without the predictor of concern to the model with this predictor included. The r_{sp} between the multiplicative term and the logarithm of perceived weight was -.009. Next, a minimum substantial effect, Δ, was defined as $(r^2/2)^{1/2}$ (Weber & Popova, 2012, p. 195), where r was the average effect size determined based on a meta-analysis of a group of studies in a certain field (Weber & Popova, 2012). For the current equivalence test, $|r| = .114$ was chosen. According to a summary of percentiles of communication meta-analyses by topic area (Weber & Popova, 2012, Table 3, p. 198), $|r| = .114$ was the 50th percentile among the studies on persuasion effects. The less the value of $|r|$, the more conservative an equivalence test is. Based on the equation of $\Delta = (r^2/2)^{1/2}$, the value of Δ in this equivalence test could be calculated as .08. Now the null hypothesis and the alternative hypothesis for the equivalence test could be defined, H_0: $|\rho_{sp}| \geq .08$, H_a: $|\rho_{sp}| < .08$. Next, the noncentrality parameter, λ, and the empirical t value were calculated based on the formulas provided in Weber and Popova's article (2012, p. 203). Finally, in a noncentral t distribution with $df = n - 2 = 396$ and $\lambda = 1.61$, a p value of .08 was calculated with $|t| \leq .18$. Therefore, this equivalence test failed to reject its null hypothesis that the population semipartial correlation between the multiplicative term and the logarithm of perceived message weight were equal to or greater than a minimum substantial effect size of .08. In conclusion, H4d was not supported.

Perceived Message Scale Value

I conducted a two-way ANCOVA with U and s as independent variables, subject's initial position as a covariate, and the two-item average of perceived scale value ($M = 12.21$, $SD = 4.91$, $Sk = 0.16$ [$SE = 0.12$], $Ku = -0.11$ [$SE = 0.23$], $N = 448$) as the dependent variable (Figure 16). Consistent with the results when examining perceived weight, the only significant effect was the main effect of *s*, $F(2, 438) = 47.28$, $p < .001$, $\eta^2 = .178$. The linear increase of perceived scale value as a function of *s* was significant, $F(1, 438) = 83.68$, $p < .001$, $\eta^2 = .160$, and the quadratic effect of *s* was also significant but with a much smaller effect size compared with the linear effect, $F(1, 438) = 11.17$, $p = .001$, $\eta^2 = .025$.

The equation, $s_p = s\Delta(\psi) = se^{-\gamma\psi}$, where $\psi = kD/P$ and s_p represents perceived message scale value, was predicted by the scale-value-pullback model. Again, I tested this equation directly. The strategy and procedure were the same as the one reported above for perceived weight. By taking the natural logarithm on both sides of the equation to linearize the relationship, the equation in question became $\ln(s_p) = \ln(s) - \gamma\psi$. Because $\psi = kD/P$, a test of the linearized equation could be done by testing this linear regression model: $\ln(s_p) = a + b_1D + b_2(1/P) + b_3(D/\mathrm{P}) + \epsilon$. The two-item average of perceived scale value was used. Thirty-six subjects with a nonpositive value of perspective and twelve subjects with a negative value of message discrepancy were excluded. There were 403 subjects left for analysis.

The results of the regression indicated the three predictors explained 19.51% of the variance, Adjusted $R^2 = .189$, $F(3, 399) = 32.23$, $p < .001$. Only b_1 was positive and significant (see Table 29), which again indicated that the functional form in $\psi = kD/P$ was implausible. The estimated b_2 and b_3 were positive and not significant, which did not support H5c or H5d predicted by the scale-value-pullback model. Without including the multiplicative term, the results of the regression indicated the two predictors explained



(a) (b)

Figure 16. Two-way ANCOVA on perceived scale value in the main study. The y-axis represents the predicted value of the ANCOVA model based on the two-item average of perceived scale value. Covariate: subject's initial position. $N = 448$. In Panel a, the x-axis represents the upper bound (U), and each line represents one of the message scale value (s) conditions; in Panel b, the x-axis represents the message scale value (s), and each line represents one of the upper bound (U) conditions.

Table 29

Main Study: Summary of Linear Regression Analysis for Variables Predicting Perceived Scale Value

	Unstandardized coefficient	*SE*	Standardized coefficient	Semipartial correlation
Intercept	2.52***	0.02		
Message discrepancy	0.01***	0.001	.44***	.44
Inverse of perspective	0.04	0.28	.01	.01
Multiplicative term	0.01	0.02	.01	.01

Note. $n = 403$. Model fit: $R^2 = .195$, Adjusted $R^2 = .189$, $F(3, 399) = 32.23$, $p < .001$.

*$p < .05$. **$p < .01$. ***$p < .001$.

19.49% of the variance, Adjusted $R^2 = .191$, $F(2, 402) = 48.41$, $p < .001$, which did not support H5c.

Because failing to reject the null hypotheses when testing H5c and H5d did not indicate the acceptance of H5a or H5b, equivalence testing was conducted to evaluate H5a and H5b. The procedure of equivalence testing was the same as the one reported above for the equivalence testing of H4d. For H5a, $t(401) = 0.12$, the noncentrality parameter $\lambda = 1.62$, $p = .07$; for H5b, $t(401) = 0.28$, the noncentrality parameter $\lambda = 1.62$, $p = .09$. Both of the equivalence tests were nonsignificant. Therefore, H5a and H5b were not supported.

The equation $s_p = s\Delta(\psi) + s_0 = sk'(1 - e^{-\gamma\psi}) + s_0$, where $\psi = kD/P$ and s_p represents perceived message scale value, was predicted by the complex model. I used nonlinear regression to fit this equation directly using the same 403 subjects. The specified equation was: $s_p = b_0 s\{1 - \exp[b_1 D/P + b_3 D + b_4(1/P)]\} + b_2 s_0 + b_5 + \epsilon$. The model explained 27.85% of the variance, Adjusted $R^2 = .270$, $F(5, 397) = 30.72$, $p < .001$. AIC, AIC_c, and BIC were 2,301.87, 2,302.16, 2,329.87, respectively. The standard error of the regression was 4.17. The skewness of the residuals was -0.01 ($SE = 0.12$). The coefficients b_1, b_3, and b_4 were negative, as expected, and all had a 95% CI excluding zero (see Table 30).

Without including the multiplicative term, the model explained 26.50% of the variance, Adjusted $R^2 = .258$, $F(4, 398) = 35.87$, $p < .001$. AIC, AIC_c, and BIC were 2,307.33, 2,307.54, 2,331.33, respectively. The standard error of the regression was 4.20. The skewness of the residuals was 0.04 ($SE = 0.12$). The model including the multiplicative term fit the data only slightly better than the model excluding the multiplicative term. Therefore, the more parsimonious model without the multiplicative term was more plausible. In the model without the multiplicative term, the coefficient for the inverse of perspective was not significant (see Table 31). In conclusion, when the additive model of psychological discrepancy was assumed, perceived message scale value was not found to be an increasing function of perspective, which did not support H5e or H5f; when the multiplicative model of psychological discrepancy was assumed, perceived message scale value was an increasing function of the inverse of perspective, which supported both H5e and H5f. However, because the nonlinear regression model with the multiplicative assumption only fit the data slightly better than the nonlinear regression model with the additive assumption, the latter was favored. Therefore, neither H5e nor H5f was supported.

Table 30

Main Study: Summary of Nonlinear Regression Analysis for Variables Predicting Perceived Scale Value: Multiplicative Term Included

	Estimated coefficient	*SE*	95% Confidence Interval
b_0	0.03	0.01	[0.01, 0.05]
b_1 (The multiplicative term)	-0.69	0.23	[-1.14, -0.25]
b_2	-0.29	0.39	[-1.05, 0.48]
b_3 (Message discrepancy)	-0.13	0.02	[-0.16, -0.10]
b_4 (Inverse of perspective)	-10.58	3.91	[-18.27, -2.89]
b_5	14.53	1.13	[12.32, 16.75]

Note. n = 403. AIC, AIC_c, and BIC were 2,301.87, 2,302.16, and 2,329.87, respectively. AIC is the Akaike information criterion; AIC_c is the AIC corrected for a small sample size; BIC is the Bayesian information criterion.

Table 31

Main Study: Summary of Nonlinear Regression Analysis for Variables Predicting Perceived Scale Value: Multiplicative Term Excluded

	Estimated coefficient	*SE*	95% Confidence Interval
b_0	0.04	0.01	[0.02, 0.07]
b_2	-0.66	0.37	[-1.39, 0.07]
b_3 (Message discrepancy)	-0.11	0.02	[-0.14, -0.08]
b_4 (Inverse of perspective)	0.00	1.11	[-2.17, 2.17]
b_5	15.18	1.11	[13.00, 17.36]

Note. n = 403. AIC, AIC_c, and BIC were 2,307.33, 2,307.54, and 2,331.33, respectively. AIC is the Akaike information criterion; AIC_c is the AIC corrected for a small sample size; BIC is the Bayesian information criterion.

Direct Fitting of Model Equations

By using the same 403 subjects noted in the above section, a series of nonlinear regression model were fitted, in which the dependent variable was subject's final position (see Table 32). The model fit statistics are reported in Table 33, and the estimated coefficients of the first three models in Table 32 are reported in Tables 34 to 36. For the full complex model, although the fit statistics indicate a good fit, none of the coefficients had a 95% CI that excluded zero. Therefore, the restricted complex model was the more plausible one between the two variations of the complex model.

The more parsimonious weight-discounting model and scale-value-pullback model explained the variance in final position more than the complex model did. Regarding the estimated coefficient of the multiplicative term, this coefficient was negative and significant in all three models (see Tables 34 to 36). Also, the coefficients for message discrepancy and the inverse of perspective were all negative, as expected, in all three models. The results about the estimated coefficients indicated that the multiplicative model of psychological discrepancy was indeed plausible with final position being the dependent variable.

Because the nonlinear regression models in Table 32 were not nested, I used AIC, AIC_c, and BIC as the criteria for model selection (Burnham & Anderson, 2004). The weight-discounting model had the lowest AIC, AIC_c, and BIC among the four nonlinear regression models (see Table 33). Therefore, the weight-discounting model was the most plausible model based on the nonlinear regression analysis, although the scale-value-pullback model was a close second.[8]

Table 32

Direct Fitting of Model Equations in the Main Study: Nonlinear Regression Model Equations

Psychological-discrepancy-weight-discounting model	$\widehat{R} = \dfrac{B_1 s_0 + s^*\exp[B_2 D + B_3(1/P) + B_4 D(1/P)]}{B_1 + \exp[B_2 D + B_3(1/P) + B_4 D(1/P)]}$
Psychological-discrepancy-scale-value-pullback model	$\widehat{R} = \dfrac{B_1 s_0 + s^*\exp[B_2 D + B_3(1/P) + B_4 D(1/P)]}{B_1 + 1}$
Complex model (restricted; $k' = 1$)	$\widehat{R} = \dfrac{B_1 s_0 + \exp[B_2 D + B_3(1/P) + B_4 D(1/P)]^*\{s^*[1\text{-} \exp(B_2 D + B_3(1/P) + B_4 D(1/P))] + B_5 s_0\}}{B_1 + \exp[B_2 D + B_3(1/P) + B_4 D(1/P)]}$
Complex model (full)	$\widehat{R} = \dfrac{B_1 s_0 + \exp[B_2 D + B_3(1/P) + B_4 D(1/P)]^*\{sB_6{}^*[1\text{-}\exp(B_2 D + B_3(1/P) + B_4 D(1/P))] + B_5 s_0\}}{B_1 + \exp[B_2 D + B_3(1/P) + B_4 D(1/P)]}$

Note. A symbol with a carat (^) is for a value without the error terms.

Table 33

Direct Fitting of Model Equations in the Main Study: Model Fit Statistics

	Sum of squares	*df*	*F*	*p*	R^2_{adj}	SE_{est}[a]	*Sk*[b]	AIC	AICc	BIC
Total	343.283	402								
Explained by weight-discounting model	196.28	3	177.75	< .001	.569	0.61	0.91	747.25	747.40	767.24
Explained by scale-value-pullback model	195.91	3	177.02	< .001	.568	0.61	0.92	748.25	748.40	768.25
Explained by complex model ($k' = 1$)	182.36	4	112.65	< .001	.526	0.64	1.06	785.70	785.91	809.69
Explained by the full complex model	188.38	5	96.65	< .001	.543	0.62	1.03	772.35	772.63	800.34

Note. $n = 403$. AIC is the Akaike information criterion; AIC_c is the AIC corrected for a small sample size; BIC is the Bayesian information criterion.

[a]SE_{est} is the standard error of the estimate. See the note under Table 4.

[b]*Sk* is the skewness of residuals. The standard errors of *Sk* all equaled to 0.12.

Table 34

Summary of Nonlinear Regression Analysis for Variables Predicting Final Position in the Main Study: Weight-Discounting Model

	Estimated coefficient	*SE*	95% Confidence Interval
B_1	62.03	5.57	[51.13, 73.01]
B_2 (Message discrepancy)	-0.03	0.01	[-0.04, -0.02]
B_3 (Inverse of perspective)	-5.76	1.43	[-8.58, -2.94]
B_4 (Multiplicative term)	-0.23	0.10	[-0.43, -0.03]

Note. *n* = 403. Model fit statistics reported in Table 33.

Table 35

Summary of Nonlinear Regression Analysis for Variables Predicting Final Position in the Main Study: Scale-Value-Pullback Model

	Estimated coefficient	*SE*	95% Confidence Interval
B_1	62.53	5.50	[51.71, 73.35]
B_2 (Message discrepancy)	-0.03	0.01	[-0.04, -0.02]
B_3 (Inverse of perspective)	-5.05	1.27	[-7.54, -2.57]
B_4 (Multiplicative term)	-0.20	0.09	[-0.37, -0.03]

Note. *n* = 403. Model fit statistics reported in Table 33.

Table 36

Summary of Nonlinear Regression Analysis for Variables Predicting Final Position in the Main Study: Complex Model (Restricted With k' = 1*)*

	Estimated coefficient	*SE*	95% Confidence Interval
B_1	45.94	7.02	[32.15, 59.74]
B_2 (Message discrepancy)	-0.04	0.01	[-0.05, -0.03]
B_3 (Inverse of perspective)	-1.73	1.62	[-4.91, 1.45]
B_4 (Multiplicative term)	-0.20	0.09	[-0.38, -0.01]
B_5	8.62	1.13	[6.39, 10.84]

Note. n = 403. Model fit statistics reported in Table 33.

Structural Equation Modeling

To account for measurement error of perceived weight and to examine the mediating role of psychological discrepancy in the belief change process, a structural equation model (SEM hereafter) was fit (Figure 17). The SEM analysis was exploratory, because the fitted model did not take the functional forms of the four models into account (Equations 8, 10, and 11). The SEM analysis was useful in that it tested a model that conceptually represented what was supposed to be going on in a belief change process.

Table 37 shows the sample covariance matrix of the observed variables. First, each of the two factors in the experiment was coded into dummy variables. For both

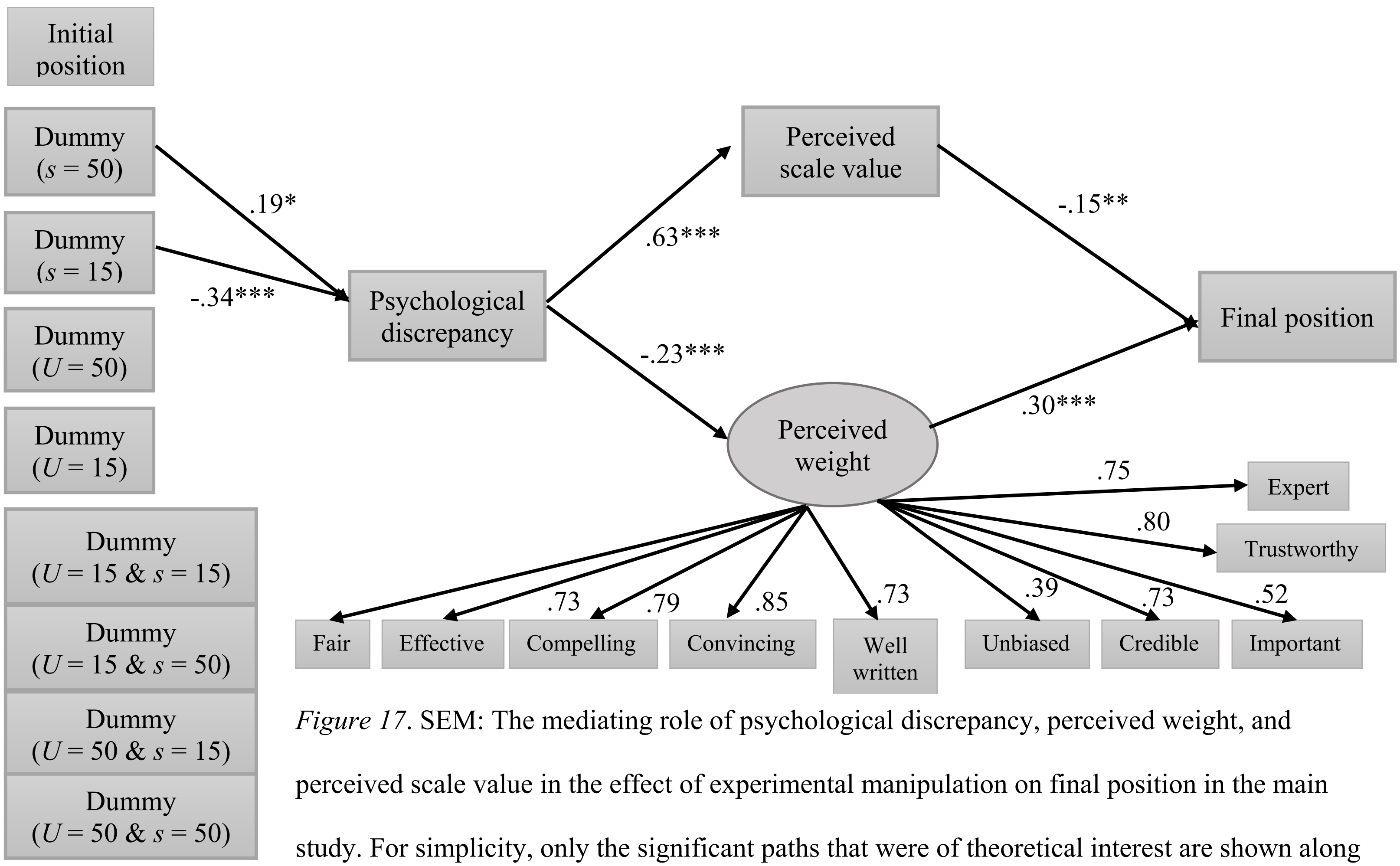


Figure 17. SEM: The mediating role of psychological discrepancy, perceived weight, and perceived scale value in the effect of experimental manipulation on final position in the main study. For simplicity, only the significant paths that were of theoretical interest are shown along with the standardized coefficients. *$p < .05$. **$p < .01$. ***$p < .001$.

Table 37

Main Study: Sample Covariance Matrix of the Observed Variables for SEM

	Fair	Eff	Wrt	Compel	Convinc	Import	Unbia	Cred	Expert	Trust	ψ	s_p	R	U_50	U_15	s_50	s_15	sU_50_50	sU_50_15	sU_15_50	sU_15_15	s_0
Fair	3688.93																					
Eff	1909.77	5911.74																				
Wrt	1815.67	2975.39	3802.17																			
Compel	2472.98	3487.35	3106.28	4374.56																		
Convinc	2772.76	3400.33	3027.72	3959.10	4479.36																	
Import	1690.85	1907.28	1452.88	1705.30	1845.69	4127.19																
Unbia	1272.69	1170.76	952.16	1112.69	1273.72	875.08	3544.23															
Cred	2316.06	2330.54	1914.64	2301.14	2450.08	1513.16	1828.70	3572.23														
Expert	2263.19	2813.56	2367.33	2749.47	2830.58	1616.26	1272.19	2622.01	4022.14													

Table 37

(continued)

	Fair	Eff	Wrt	Compel	Convinc	Import	Unbia	Cred	Expert	Trust	ψ	s_p	R	U_50	U_15	s_50	s_15	sU_50_50	sU_50_15	sU_15_50	sU_15_15	s_0
Trust	2721.86	2417.99	2288.69	2636.35	2864.55	1681.90	1629.70	2905.13	3228.56	4028.90												
ψ	-180.72	-32.72	-49.04	-102.84	-136.52	-57.07	-45.38	-93.79	-66.43	-122.09	32.79											
s_p	-140.63	19.63	11.06	-26.98	-53.53	-20.73	-36.64	-62.03	-20.56	-72.61	20.23	23.66										
R	27.14	20.57	18.34	23.42	25.01	14.75	7.08	17.96	18.64	24.62	-1.84	-1.23	0.88									
U_50	0.91	2.61	-0.14	0.82	0.97	-0.04	1.44	2.64	1.51	1.67	-0.11	-0.11	0.04	0.22								
U_15	-2.11	0.49	0.76	-0.85	-0.65	-0.55	-0.88	-3.00	-3.87	-2.41	-0.08	-0.05	-0.05	-0.11	0.22							
s_50	-8.87	-0.59	-3.07	-4.30	-5.70	-2.24	-2.37	-5.60	-4.47	-5.97	0.79	0.71	0.04	0.00	0.00	0.22						
s_15	8.89	0.29	0.82	2.46	4.35	3.04	4.44	6.22	2.93	4.74	-1.09	-1.01	-0.08	0.00	-0.00	-0.11	0.22					
sU_50_50	-2.73	-0.46	-2.83	-2.19	-3.01	-1.06	-0.40	-0.69	-1.62	-2.15	0.18	0.15	0.04	0.08	-0.04	0.08	-0.04	0.10				

Table 37

(continued)

	Fair	Eff	Wrt	Compel	Convinc	Import	Unbia	Cred	Expert	Trust	ψ	s_p	R	U_50	U_15	s_50	s_15	sU_50_50	sU_50_15	sU_15_50	sU_15_15	s_0
sU_50_15	-4.46	0.39	-0.56	-1.75	-1.82	-0.10	-1.01	-3.07	-2.90	-3.09	0.31	0.29	-0.02	-0.04	0.07	0.07	-0.04	-0.01	0.10			
sU_15_50	2.30	-0.16	0.01	-0.73	0.81	0.01	1.81	1.49	0.18	1.23	-0.23	-0.29	-0.04	0.07	-0.04	-0.04	0.08	-0.01	-0.01	0.10		
sU_15_15	3.94	0.26	1.15	2.00	2.22	0.77	1.38	2.26	0.83	2.42	-0.58	-0.44	-0.01	-0.04	0.07	-0.04	0.07	-0.01	-0.01	-0.01	0.09	
s_0	12.94	9.30	10.10	10.92	11.93	5.37	3.25	7.96	7.48	10.94	-1.20	-0.76	0.43	-0.00	-0.00	0.02	-0.02	-0.00	-0.00	-0.02	0.01	0.39

Note. n = 407. Fair, Eff, Wrt, Compel, Convinc, Import, Unbia, Cred, Expert, and Trust are fairness, effectiveness, the quality of being well written, the quality of being compelling, the quality of being convincing, importance, unbiasedness, credibility, expertness, and trustworthiness, respectively; R, ψ, s_p, s_0 are for final position, psychological discrepancy, perceived scale value, and initial position, respectively; U_50, U_15, s_50, s_15, sU_50_50, sU_50_15, sU_15_50, sU_15_15 are the dummy variables for $U = 50$, $U = 15$, $s = 50$, $s = 15$, $s = 50$ and $U = 50$, $s = 50$ and $U = 15$, $s = 15$ and $U = 50$, and $s = 15$ and $U = 15$, respectively. All variables were winsorized and transformed as described in the main text.

factors, the 30 years group was used as the reference group. Next, four dummy variables for the interaction effects could be created (see Figure 17). All paths from the exogenous variables to psychological discrepancy, perceived weight, perceived scale value, and final position were estimated. The dummy variables were allowed to covary with each other. The initial position as an exogenous covariate was allowed to covary with the dummy variables. The residual covariance between perceived weight and perceived scale value was estimated (see the recommendation for doing this in Preacher & Hayes, 2008, pp. 882-883). A total of 407 subjects were used after excluding 36 subjects who had nonpositive perspective values and applying listwise deletion for missing data.

For all the observed variables excluding the exogenous variables, Mardia's skewness was 1,455.02 ($p < .001$), and Mardia's kurtosis was 22.11 ($p < .001$). The results indicated that the multivariate normality assumption was severely violated. Thus, a robust ML estimation should be used (Kline, 2012, p. 122). Given a sample size of 407 and a large number of free parameters (113) in the model, the Yuan-Bentler variation of the robust ML estimation was used (Bentler & Yuan, 1999; see also Ullman, 2006, p. 43) to generate the robust test statistic and the robust standard errors of the estimated coefficients. I used the *lavaan* package in R (Rosseel, 2012) to run the SEM model. The initial model fit did not allow the indicators of the perceived weight latent variable to covary. Then some items' residuals for the perceived weight latent variable were allowed to covary based on the modification indices. Table 38 shows the estimated residual covariances among the 10 items for the perceived weight latent variable in the respecified model. The respecified model's global fit statistics are reported in Table 39.

Table 38

Estimated Standardized Residual Covariances Among the 10 Items for the Perceived Weight Latent Variable in SEM

	Fair	Eff	Wrt	Compel	Convinc	Import	Unbia	Cred	Expert	Trust
Fair	.30***									
Eff	-.51***	.47***								
Wrt	-.27**	.16**	.47***							
Compel	—	.17**	.41***	.37***						
Convinc	—	—	.30***	.68***	.27***					
Import	—	—	—	—	—	.73***				
Unbia	—	—	—	—	—	—	.85***			
Cred	—	—	—	—	—	—	.34***	.46***		
Expert	—	—	—	.07*	—	—	—	.29***	.43***	
Trust	.14*	-.21**	—	—	—	—	.16**	.42***	.51***	.36***

Note. n = 407. Fair is fairness; Eff is effectiveness; Wrt is the quality of being well written; Compel is the quality of being compelling; Convinc is the quality of being convincing; Import is importance; Unbia is unbiasedness; Cred is credibility; Expert is expertness; Trust is trustworthiness. A cell with — means that the covariance was not estimated. *$p < .05$. **$p < .01$. ***$p < .001$

Table 39

SEM Model Fit Statistics: Examining Final Position in the Main Study

df	χ^2 (p)	*CFI*	*RMSEA* 90% CI	*SRMR*	R^2 for structural equations	Adjusted R^2 for structural equations
128	330.68 (< .001)	.97	.06 [.06, .07]	.05	.330, .240, .561, .681	.315, .550

Note. n = 407. Number of free parameters: 113. Estimation method: Maximum likelihood. *CFI*: Comparative fit index; *RMSEA*: Root mean square error of approximation; *SRMR*: Standardized root mean square residual. The first four model fit statistics (χ^2, *CFI, RMSEA*, and *SRMR*) were the robust versions (Yuan-Bentler; Bentler & Yuan, 1999) that adjusted for multivariate nonnormality in the data. From left to right, the R^2s were for psychological discrepancy, perceived message weight, perceived message scale value, and final position, respectively. The Adjusted R^2s were for psychological discrepancy and perceived message scale value, respectively. The adjusted R^2s for perceived message weight and final position were unknown, because the regression models for these two variables include a latent variable.

Except the chi-square test, the remaining statistics indicated an acceptable global model fit. The respecified model's local fit was assessed by examining the difference between the sample covariance matrix and the model-implied covariance matrix for each pair of the observed variables (see Table 40). The value in each cell of the matrix in Table 40 is in the correlation metric. An absolute correlation greater than .1 indicates a poor local fit between a pair of variables (see Goodboy & Kline, 2017). Examining Table 40 revealed that 16 out of 253 (6%) values indicated a poor local fit. All of the poor local fits involved one of the 10 items for the perceived weight. No further model respecification involving the items in the measurement component of the model seemed reasonable at this stage.

All loadings of the latent variable were positive and had p values less than .001. The fairness loading was fixed to one. A check on the estimated path coefficients revealed that none of the paths from the upper bound dummy variables (the $U = 15$ dummy variable and the $U = 50$ dummy variable) had a significant coefficient. Regarding the dummy variables representing an interaction effect, none of them were significant.

In Figure 17, the significant standardized path coefficients that were of interest are displayed. Without the presence of a significant interaction term, the signs of the coefficients of the $s = 50$ dummy ($\hat{\beta} = .19$, $z = 2.14$, $p = .032$) and the $s = 15$ dummy ($\hat{\beta} = -.34$, $z = -4.54$, $p < .001$) indicated that ψ increased as s increased. Perceived scale value was an increasing function of ψ, $\hat{\beta} = .63$, $z = 14.62$, $p < .001$, whereas perceived weight was a decreasing function of ψ, $\hat{\beta} = -.23$, $z = -3.59$, $p < .001$, which provided some support for the complex model if the equation $\psi = kD/P$ was plausible.

Table 40

Matrix of Residuals in the Correlation Metric for SEM

	Fair	Eff	Wrt	Compel	Convinc	Import	Unbia	Cred	Expert	Trust	ψ	s_p	R	U_50	U_15	s_50	s_15	sU_50_50	sU_50_15	sU_15_50	sU_15_15	s_0
Fair	0.00																					
Eff	-0.01	0.00																				
Wrt	-0.02	0.02	-0.00																			
Compel	-0.05	0.04	0.01	0.00																		
Convinc	-0.03	0.04	0.01	0.00	-0.00																	
Import	0.00	0.01	-0.01	-0.01	-0.01	0.00																
Unbia	0.02	-0.03	-0.03	-0.03	-0.02	0.02	0.00															
Cred	0.03	-0.03	-0.02	0.00	-0.01	0.01	0.01	0.00														

Table 40

(continued)

	Fair	Eff	Wrt	Compel	Convinc	Import	Unbia	Cred	Expert	Trust	ψ	s_p	*R*	U_50	U_15	s_50	s_15	sU_50_50	sU_50_15	sU_15_50	sU_15_15	s_0
Expert	-0.04	0.03	0.06	0.03	0.02	0.01	0.04	0.01	0.00													
Trust	-0.01	-0.00	0.00	-0.01	-0.01	-0.00	0.03	0.01	-0.00	-0.00												
ψ	**-0.19**	**0.21**	**0.15**	0.04	-0.02	0.05	0.02	0.01	**0.11**	-0.02	0.00											
s_p	**-0.26**	**0.24**	**0.22**	**0.12**	0.05	0.07	-0.03	-0.02	**0.13**	-0.03	-0.00	0.00										
R	0.07	-0.07	-0.04	-0.01	-0.01	-0.01	-0.06	-0.03	-0.05	0.03	0.00	0.00	0.00									
U_50	-0.01	0.03	-0.04	-0.02	-0.02	-0.03	0.03	0.05	0.01	0.01	-0.00	-0.00	-0.00	0.00								
U_15	-0.03	0.05	0.06	0.01	0.02	0.01	-0.01	-0.07	-0.09	-0.04	0.00	0.00	0.00	-0.00	0.00							
s_50	**-0.12**	**0.15**	0.06	0.04	0.01	0.04	0.00	-0.03	0.02	-0.02	0.00	-0.00	-0.00	0.00	0.00	0.00						
s_15	**0.14**	**-0.14**	**-0.12**	-0.08	-0.04	-0.01	0.08	0.07	-0.06	-0.00	-0.00	0.00	0.00	0.00	0.00	-0.00	0.00					

Table 40

(continued)

	Fair	Eff	Wrt	Compel	Convinc	Import	Unbia	Cred	Expert	Trust	ψ	s_p	R	U_50	U_15	s_50	s_15	sU_50_50	sU_50_15	sU_15_50	sU_15_15	s_0
sU_50_50	-0.03	0.08	-0.05	0.00	-0.03	0.02	0.03	0.06	0.02	-0.00	0.00	0.00	0.00	-0.00	0.00	0.00	0.00	0.00				
sU_50_15	**-0.11**	**0.12**	0.08	0.03	0.04	0.07	0.00	-0.06	-0.04	-0.04	-0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.00			
sU_15_50	0.06	-0.06	-0.05	-0.09	-0.03	-0.04	0.07	0.02	-0.05	0.00	-0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00		
sU_15_15	0.09	-0.09	-0.04	-0.01	-0.01	-0.03	0.02	0.02	-0.06	0.01	0.00	0.00	0.00	0.00	0.00	-0.00	0.00	0.00	-0.00	0.00	0.00	
s_0	0.06	-0.06	0.01	-0.01	-0.01	-0.04	-0.05	-0.04	-0.07	0.00	0.00	0.00	-0.00	-0.00	0.00	0.00	-0.00	-0.00	0.00	0.00	-0.00	0.00

Note. $n = 407$. Fair, Eff, Wrt, Compel, Convinc, Import, Unbia, Cred, Expert, and Trust are fairness, effectiveness, the quality of being well written, the quality of being compelling, the quality of being convincing, importance, unbiasedness, credibility, expertness, and trustworthiness, respectively; R, ψ, s_p, s_0 are for final position, psychological discrepancy, perceived scale value, and initial position, respectively; U_50, U_15, s_50, s_15, sU_50_50, sU_50_15, sU_15_50, sU_15_15 are the dummy variables for $U = 50$, $U = 15$, $s = 50$, $s = 15$, $s = 50$ and $U = 50$, $s = 50$ and $U = 15$, $s = 15$ and $U = 50$, and $s = 15$ and $U = 15$, respectively. The absolute correlations that are greater than .1 are in bold.

A number of 12 three-path indirect effects from message scale value's dummy variables to final position were estimated (Table 41). The indirect effects were more plausible for the $U = 30$ group than the other two upper bound groups (i.e., the $U = 15$ group and the $U = 50$ group). Across the three upper bound groups (i.e., the $U = 15$ group, the $U = 30$ group, and the $U = 50$ group), the indirect effects starting from $s = 15$ were more plausible than the indirect effects starting from $s = 50$, because the effect of $s = 15$ on ψ was greater than the effect of $s = 50$ on ψ. One caution regarding the indirect effects was that the negative path from s_p to final position went against the prediction. All models predicted that the relationship would be positive. The positive path from w_p to final position was consistent with the prediction.

Lastly, a number of six total effects of message scale value's dummy variables on final position were estimated (Table 42). The total effect of $s = 15$ on R when $U = 30$, and the total effect of $s = 15$ on R when $U = 50$ was significant. The negative signs for both $s = 15$ and $s = 50$ in all upper bound groups to some extent corroborated the significant quadratic trend of s found in the ANCOVA on final position reported earlier.

Table 41

Three-Path Indirect Effects by Upper Bound in the Main Study: From Message Scale Value's Dummy Variables to Final Position

	Indirect effect	Standardized $\hat{\beta}$	z value	p
$U = 15$	s50 -> ψ -> w_p -> R	-.01	-1.23	.220
	s15 -> ψ -> w_p -> R	.03	3.12	.002
	s50 -> ψ -> s_p -> R	-.01	-1.13	.257
	s15 -> ψ -> s_p -> R	.04	2.69	.007
$U = 30$	s50 -> ψ -> w_p -> R	-.01	-1.87	.061
	s15 -> ψ -> w_p -> R	.02	2.85	.004
	s50 -> ψ -> s_p -> R	-.02	-1.70	.090
	s15 -> ψ -> s_p -> R	.03	2.44	.015
$U = 50$	s50 -> ψ -> w_p -> R	-.01	-1.57	.116
	s15 -> ψ -> w_p -> R	.02	2.01	.045
	s50 -> ψ -> s_p -> R	-.01	-1.54	.125
	s15 -> ψ -> s_p -> R	.02	1.74	.082

Note. s50 and s15 are the $s = 50$ and the $s = 15$ dummy variables, respectively. w_p stands for perceived weight; s_p stands for perceived scale value. The Sobel test (1982) was used for testing the statistical significance of each indirect effect.

Table 42

Total Effects of Message Scale Value's Dummy Variables on Final Position by Upper Bound in the Main Study

	Total effect	Standardized $\hat{\beta}$	*z* value	*p*
$U = 15$	s50 on *R*	-.06	-0.99	.323
	s15 on *R*	-.12	-1.73	.083
$U = 30$	s50 on *R*	-.06	-0.87	.385
	s15 on *R*	-.15	-2.94	.003
$U = 50$	s50 on *R*	-.01	0.15	.885
	s15 on *R*	-.21	-3.71	< .001

Note. s50 and s15 are the $s = 50$ and the $s = 15$ dummy variables, respectively. The Sobel test (1982) was used for testing the statistical significance of each total effect.

CHAPTER 8

DISCUSSION

This dissertation found an unresolved problem in the role of psychological discrepancy in the psychological-discrepancy-weight-discounting model of belief change (Fink et al., 1983). The puzzle is: Does psychological discrepancy affect the weight of a message only, affect the scale value of the message only, affect neither, or affect both? For the four mathematical models that represent the above four possibilities, competing hypotheses were posited either through an analytic derivation or a computational approximation. These hypotheses were tested in an experiment. The results showed that the weight-discounting model fit the data the best, which indicated that psychological discrepancy affected the message weight only.

Model Evaluation

Table 43 shows the overall results of hypothesis testing. H1 and H2 tested the relationship between the experimental manipulation (i.e., message scale value and perspective) and final belief position. H1a posited that the wider the perspective, the greater the final position (based on the weight-discounting model, the scale-value-pullback model, and the independent-psychological-discrepancy model), whereas H1b posited an inverted U-shape relationship between perspective and final position (based on the complex model). The data showed that the linear increase of final position as a function of perspective was indeed significant, whereas the quadratic effect of perspective was nonsignificant. H2a posited an inverted U-shape relationship between

Table 43

Main Study: Hypothesis Testing Results by the Four Models

Competing hypotheses	Weight-discounting model	Scale-value-pullback model	Independent model	Complex model
H1a	×	×	×	
H1b				×
H2a	×			×
H2b		×	×	
H3a	×			×
H3b		×	×	
H4a	×			×
H4c		×	×	
H4b	×			×
H4d		×	×	
H5a	×		×	
H5c		×		
H5e				×
H5b	×		×	
H5d		×		
H5f				×

Note. A verbal statement of each hypothesis can be found in the main text. The symbol of × means that a specific hypothesis is predicted by a certain model. A shaded cell means that a hypothesis was supported in the main study. Each block separated by horizontal lines is composed of competing hypotheses.

message scale value and final position (based on the weight-discounting model and the complex model), whereas H2b posited an interaction effect between message scale value and perspective upon final position (based on the scale-value-pullback model and the independent-psychological-discrepancy model). The data showed a significant linear increase of final position as a function of message scale value, a significant quadratic effect with an inverted U-shape of message scale value, and a nonsignificant interaction between message scale value and perspective. The testing results of H1 and H2 were in favor of the weight-discounting model.

However, there was also a discrepancy in the data pattern from what was expected. Panel a of Figure 14 showed that for the s = 15 years condition, final position was a decreasing function of upper bound (i.e., a decreasing function of perspective assuming a constant lower bound; see Figure 12 for the manipulation check), whereas an increasing function would have been found for all three message scale value conditions given the acceptance of H1a. The irregular pattern for the s = 15 years condition was probably due to the irregular pattern of psychological discrepancy for the same condition (see Panel a of Figure 13). A discussion of psychological discrepancy is presented later in this chapter.

H3 tested the existence of a boomerang effect. H3a posited no boomerang effect (based on the weight-discounting model and the complex model), whereas H3b posited that the proportion of boomerang subjects in the s = 50 years condition was greater than the proportions in the other two message scale value conditions (based on the scale-value-pullback model and the independent-psychological-discrepancy model). The dependent-samples *t* test found that the subjects reported their final positions significantly greater than their initial positions after reading the judge's decision (H3a supported. The proportion of boomerang subjects did not differ by message scale value conditions (H3b not supported). The results for H3 were in favor of the weight-discounting model and the complex model.

H4 tested the relationship between perspective and perceived message weight. The gist of the H4 testing was to examine, assuming $\psi = kD/P$, whether psychological discrepancy had effects on message weight due to the manipulation of perspective. H4a posited that the inverse of perspective negatively predicted the natural logarithm of perceived message weight when message discrepancy was at its sample mean. H4b tested the assumption of $\psi = kD/P$ directly and posited that the multiplicative term of message discrepancy and the inverse of perspective negatively and significantly predicted the natural logarithm of perceived message weight. Both H4a and H4b were predicted by the weight-discounting model and the complex model. The data indeed indicated a negative relationship between the inverse of perspective and the natural logarithm of perceived message weight, which supported H4a. However, the multiplicative term was not significant, which did not support H4b. H4c and H4d were the no-effect counterparts of H4a and H4b, respectively. Because H4a was supported, H4c was not accepted. However, failing to reject the null hypothesis of no effect when testing H4b did not lead to the acceptance of the null hypothesis (i.e., H4d). The equivalence test was conducted to test the no effect explicitly, and the test was nonsignificant. That is, the estimated coefficient of the multiplicative term of message discrepancy and the inverse of perspective was not significantly different from zero, nor was its absolute value significantly smaller than the absolute value of the specified minimum substantial effect. Considering the testing results of H4, the weight-discounting model and the complex model were favored.

A few comments can be made about the nonsignificant equivalence test of H4d. First, the equivalence test adopted a conservative strategy by setting the minimum substantial effect as .08, an even smaller value than Cohen's (1988) definition of a small effect of .10. Second, a consequence of using a conservative strategy was that the equivalence test might have been underpowered (Weber & Popova, 2012, pp. 205-206). That is, there might have been a true no-effect. The equivalence test was nonsignificant, probably because the sample size was not great enough to allow the test to detect a significant smaller absolute (because the test was two-tailed) effect size than the absolute minimum substantial effect. Third, the observed r_{sp} of the multiplicative term was -.009, which is a fairly small value. The absolute value of -.009 is more than 10 times less than a conventional small effect of .10. Also, the *p* value of the equivalence test was .08, which could be considered as marginally significant.

Taken together, considering the perceived message weight as the dependent variable, I questioned the plausibility of a multiplicative model of $\psi = kD/P$. In contrast, the alternative additive model was more plausible, where perspective indeed positively predicted perceived weight. The conclusion of testing H4 regarding model preference was that the weight-discounting model and the complex model were more plausible than the other two models, but the assumed model regarding psychological discrepancy should have been additive.

H5 tested the relationship between perspective and perceived message scale value. Similar to the testing of H4, the gist of the H5 testing was to examine, assuming $\psi = kD/P$, whether psychological discrepancy had effects on message scale value due to the manipulation of perspective. H5a posited that the inverse of perspective did not predict the natural logarithm of perceived message scale value when message discrepancy was at its sample mean. H5b tested the assumption of $\psi = kD/P$ directly and posited that the multiplicative term of message discrepancy and the inverse of perspective did not predict the natural logarithm of perceived message scale value. Testing H5a and H5b, both predicted by the weight-discounting model and the complex model, required equivalence tests. The equivalence tests were nonsignificant. However, neither H5c (the negative-effect counterpart of H5a) nor H5d (the negative-effect counterpart of H5b) was significant. Moreover, neither H5e (the positive-effect counterpart of H5a) nor H5f (the positive-effect counterpart of H5b) was significant. Taken together, again, the multiplicative assumption in $\psi = kD/P$ was implausible, because including the multiplicative term in the linear and the nonlinear regressions did not improve the model fit as compared to a model without the multiplicative term. Also, the inverse of perspective did not predict perceived message scale value. Although the equivalence tests when testing H5a and H5b were nonsignificant, the estimated values of r_{sp} for the multiplicative term and the inverse of perspective were .014 and .006, respectively, in the linear regression model. Both of these effect sizes seem small. In addition, the p values of the two equivalence tests were .07 and .09, which were marginally significant. Again, the equivalence tests might have been underpowered. The conclusion of testing H5 regarding model preference was that the weight-discounting model was more plausible than the other three, but the assumed model regarding psychological discrepancy should have been additive.

In conclusion, based on hypothesis testing, the weight-discounting model was the most plausible model. Four out of its seven hypotheses were supported. One was not supported. The remaining two were not supported based on an equivalence test that might have been underpowered. The observed effect sizes in these two hypotheses were indeed small. The complex model came second to the weight-discounting model, whereas most of the hypotheses predicted by the scale-value-pullback model and the independent-psychological-discrepancy model were not supported.

Choosing the weight-discounting model as the most plausible reflects the results of the nonlinear regression analysis that fit the model equations directly (Table 33). The weight-discounting model and the scale-value-pullback model were more parsimonious than the complex model; they explained more variance in the final position than the complex model did; and most importantly, they had the lowest AIC, AIC_c, and BIC.

A Problematic Assumption?

The results of testing some hypotheses above indicated that a multiplicative model of psychological discrepancy did not fit the data as well as an additive model. This further indicated that the assumption expressed in the equation, $\psi = kD/P$, might have been problematic. This equation predicts that, ceteris paribus, ψ is an increasing function of s (Prediction 1), but a decreasing function of U (Prediction 2). The ANCOVA conducted on ψ only supported Prediction 1 but not Prediction 2 (see Figure 13). Given that the manipulation of the upper bound was successful (see Figure 10 and Figure 12), the functional form of $\psi = kD/P$ might be implausible. To corroborate the results, a post hoc test of the functional form of the equation was conducted. The test followed Blanton and Jaccard's (2006) way of testing multiplicative models. First, $s - s_0$ (message discrepancy, D) and $U - L$ (perspective, P) were calculated for each subject. A number of 36 subjects with a nonpositive value for perspective and a number of 12 subjects with a negative value for message discrepancy were excluded. There were 403 subjects left for analysis. Next, perspective was transformed to its inverse (i.e.,$1/P$) and mean centered. Message discrepancy was also mean centered. Then, the transformed winsorized psychological discrepancy was regressed on the centered inverse of perspective, the centered message discrepancy, and the product of these two centered variables. The results of the regression indicated the three predictors explained 25.06% of the variance (Adjusted $R^2 = .245$, $F[3, 399] = 44.49$, $p < .001$). The estimated coefficient of the multiplicative term was not significant (see Table 44), which indicated that the functional form in $\psi = kD/P$ was implausible.

A similar regression model was run to corroborate the above results. This time, the actual values of the message scale value and upper bound manipulation were used. First, the message scale value and the inverse of upper bound were both mean centered. Next, by using the same 403 subjects, the transformed winsorized psychological discrepancy was regressed on initial position (as a control), the centered inverse of the

Table 44

Main Study: Summary of Linear Regression Analysis for Variables Predicting Psychological Discrepancy 1

	Unstandardized coefficient	*SE*	Standardized coefficient
Intercept	11.03***	0.25	
Message discrepancy	0.20***	0.02	.49***
Inverse of perspective	5.29	3.25	.07
Multiplicative term	0.17	0.22	.03

Note. n = 403. Model fit: R^2 =.251, Adjusted R^2 = .245, F(3, 399) = 44.49, p < .001.

*p < .05. **p < .01. ***p < .001.

upper bound, the centered message scale value, and the product of these two centered variables. The results of the regression indicated that the four predictors explained 31.03% of the variance (Adjusted R^2 = .303, F[4, 398] = 44.77, p < .001). The estimated coefficient of the multiplicative term was marginally significant, p = .070 (see Table 45). The pattern of the interaction is shown in Figure 18, where the unstandardized residuals of the above regression model (i.e., the model in Table 45) was plotted against the inverse of upper bound by message scale value. The predicted pattern should have been one where the slopes are positive for all groups of message scale value, and the magnitude of the slope increases as a function of message scale value. In contrast, the observed pattern deviated from the predicted pattern in that the slope of the s = 15 group was negative and the slope of the s = 30 group was positive but close to zero.

Table 45

Main Study: Summary of Linear Regression Analysis for Variables Predicting Psychological Discrepancy 2

	Unstandardized coefficient	*SE*	Standardized coefficient
Intercept	22.13***	1.25	
Initial position	-3.66***	0.41	-.38***
Message scale value	0.18***	0.02	.44***
Inverse of upper bound	-2.87	12.22	-.01
Multiplicative term	1.56	0.86	.08

Note. n = 403. Model fit: R^2 =.310, Adjusted R^2 = .303, $F(4, 398) = 44.77$, $p < .001$.

*$p < .05$. **$p < .01$. ***$p < .001$.

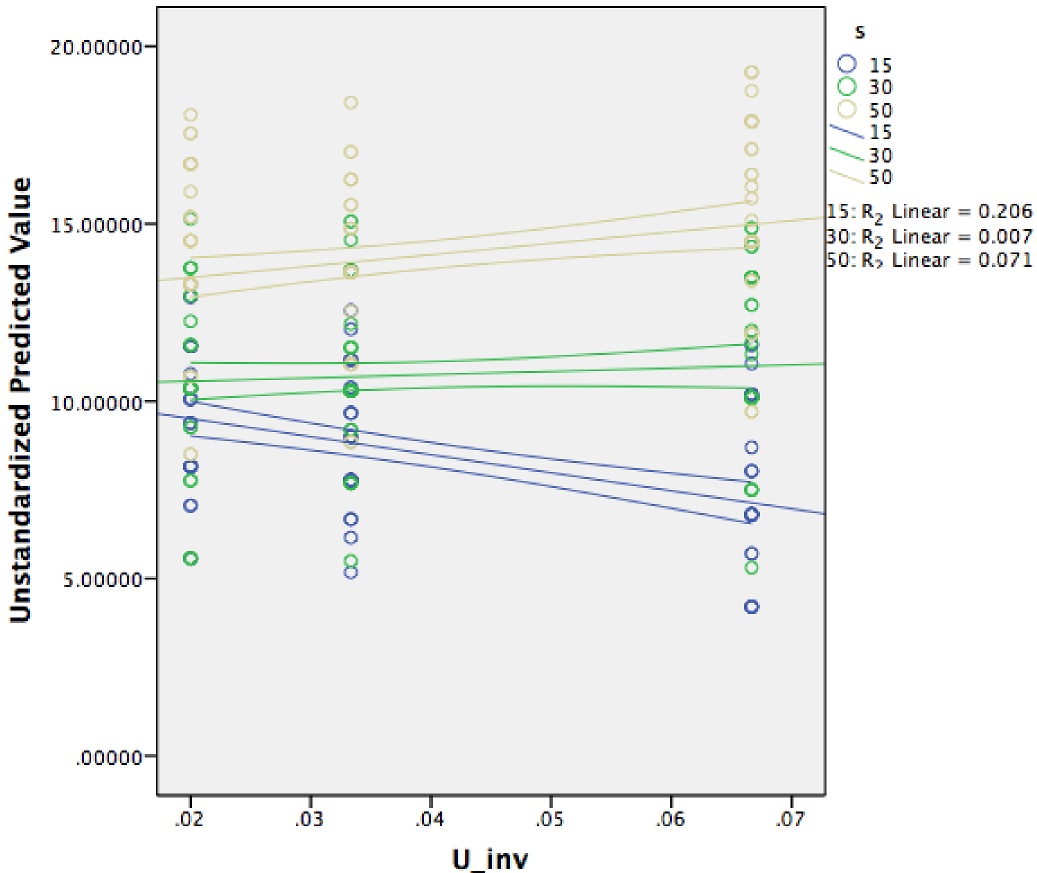


Figure 18. Predicting psychological discrepancy in the main study: Interaction of the inverse of upper bound and message scale value. In each group of *s*, the upper line and the lower line represent the 95% CI of the estimated slope in the middle. $N = 403$.

At this point, I also considered the possibility that the experimental design might be problematic, which might lead to the unique pattern of the $s = 15$ group in Panel a of Figure 13. One explanation for the positive relationship between U and s for the $s = 15$ group could be that after the subjects received the upper bound message, their belief positions changed as a function of U. Assuming that this happened, when a subject is asked to evaluate how different the judge's decision was, his or her response is based on how different the judge's decision was from his or her current belief position (i.e., a value close to the received value of U), rather than his or her initial position. However, a rebuttal of this explanation is that, if a subject's belief position changed to the received value of U, the data pattern for the $s = 30$ group in Panel a of Figure 13 should have been V-shaped, which was not the case.

If the evidence indicated the implausibility of the multiplicative model of psychological discrepancy, what could be the alternative functional forms of psychological discrepancy? I tested two alternative models: $\psi = k(D + 1/P)$ and $\psi = k(D - P)$. Both of these alternative models were consistent with the original multiplicative model in that all of them posited that the greater the perspective, the less the psychological discrepancy. The first model explained 25.00% of the variance, Adjusted $R^2 = .246$, $F(2, 400) = 66.49$, $p < .001$, whereas the second model explained 27.90% of the variance, Adjusted $R^2 = .276$, $F(2, 400) = 77.44$, $p < .001$. As expected, the term for message discrepancy was positive and significant in both models, standardized $\hat{\beta}_D = .49$, $t = 11.40$, $p < .001$ in the first model, standardized $\hat{\beta}_D = .50$, $t = 11.70$, $p < .001$ in the second model. In the first model, the term of the inverse of perspective was positive but not significant, $\hat{\beta}_{1/P} = .07$, $t = 1.72$, $p = .086$; in the second model, the term of perspective was negative and significant, $\hat{\beta}_P = -.10$, $t = -4.42$, $p < .001$. Therefore,

$\psi = k(D - P)$ seemed to be a plausible alternative model for psychological discrepancy. Also, the summation, $(D + 1/P)$ in $\psi = k(D + 1/P)$ did not make much sense regarding the units, whereas the deduction, $(D - P)$ in $\psi = k(D - P)$ did, because D and P had the same units. Although such post hoc regression analyses might rely too much on the collected data for this dissertation alone, they do suggest some directions for revising an existing model and make it possible to subject any newly proposed model to empirical testing in the future.

A piece of evidence supporting the original multiplicative model could be found in the estimated coefficients in the nonlinear regression analysis that fit the model equations directly and used final belief position as the dependent variable. The estimated coefficients for the multiplicative term were negative and significant in all three nonlinear models (see Tables 34 to 36). Also, the coefficients for message discrepancy and the inverse of perspective were all negative, as expected, in all three nonlinear models. These results about the estimated coefficients indicated that the multiplicative model of psychological discrepancy was indeed plausible with final position being the dependent variable. Therefore, this dissertation found the contradictory results about the multiplicative model's plausibility between a linear regression model predicting psychological discrepancy and several nonlinear regression models predicting final belief position. It is still a puzzle why the assumed multiplicative model of psychological discrepancy worked in an indirect test (i.e., in a model using the final belief position as the dependent variable instead of using psychological discrepancy as the dependent variable directly). Measurement errors might lead to the contradictory results. Future research should investigate this puzzle further.

Implication of the SEM Results

The SEM analysis was exploratory. The fitted SEM model did not take the functional forms in any of the four models into account. In addition to taking the measurement error of the perceived message weight into account, what the SEM analysis did was to test a model that was conceptually equivalent to what was supposed to be going on in a belief change process involving psychological discrepancy. One important conclusion based on the SEM results was that psychological discrepancy discounted the perceived message weight and enlarged the perceived message scale value. Also, when $U = 15$ or 30, the indirect effects of an $s = 15$ message (compared against an $s = 30$ message) via psychological discrepancy and perceived weight were significant (see Table 41); in addition, when $U = 30$ or 50, the indirect effects of an $s = 15$ message (compared against an $s = 30$ message) via psychological discrepancy and perceived message scale value were also significant (see Table 42). These results were consistent with the complex model and supported the mediating role of psychological discrepancy in belief change. One just need to view these results with caution due to the exploratory feature of the SEM analysis. With a forward-looking viewpoint, future research can build on the exploratory results of the SEM analysis here and propose alternative models of belief change involving psychological discrepancy.

Significance of the Study

This dissertation has theoretical, methodological, and practical significance. The theoretical significance of this dissertation is twofold. First, from the perspective of the discrepancy model in persuasion, there have been few attempts to examine the psychological process underlying the original psychological-discrepancy-discounting model. Fink et al. (1983) remarked on their study's limitation: "A mechanism explaining the discounting function needs to be explicated and tested" (p. 429). Different model assumptions imply different psychological mechanisms. The original psychological-discrepancy-discounting model was based on the weight-discounting assumption. However, this assumption was not directly tested empirically.

Second, because the psychological-discrepancy-discounting model (Fink et al., 1983) is based on Anderson's (1971a, 1981) averaging model, this dissertation also has implications for information integration theory, by directly addressing one of the basic assumptions of the information integration approach: the scale value constancy assumption. This assumption has been validated through a series of experiments (Anderson, 1971b, 1973, 1981, 2008; Anderson & Hubert, 1963; Anderson & Jacobson, 1965; Tesser, 1968; see also Fiske, 1980, p. 893; Ostrom & Davis, 1979, p. 2040), but most of these experiments examine whether a message scale value changed depending on different message combinations. This line of research had not addressed how the scale value of a message might vary as a function of participant's initial position, which is a function of discrepancy (both positional and psychological). This dissertation filled this gap by directly testing whether a subject's initial belief position before message receipt and his or her perspective had an effect on message scale value.

The results of this dissertation upheld most of hypotheses predicted by the weight-discounting model and the scale value constancy assumption. Himmelfarb (1975) made a strong argument that scale value is analogous to the concept of mass in Newton's laws. In Newton's laws, mass is constant. Himmelfarb's argument indicates that, in information integration, the only parameter that is allowed to vary is weight, not scale value. The results of this dissertation provide empirical support for this argument. More importantly, this dissertation theorized the valuation process in information integration theory (Anderson, 2008) and subjected some possibilities of the valuation process to empirical tests. The support for the weight-discounting model and the rejection of the scale-value-pullback model was consistent with a previous literature review (Kaplowitz & Fink, 1997) regarding which of the verbally-stated theories could best explain the psychological processes involved in a model that includes psychological discrepancy: Cognitive dissonance theory, which posits source derogation (i.e., a form of weight being discounted) as a function of discrepancy, was better supported than social judgment theory that posits perceived scale value as a function of the latitudes of indifference, rejection, and acceptance. Although the evidence for the valuation process included in the scale-value-pullback model was weak, the valuation process included in the complex model had some indirect support based on the SEM results. In other words, based on the four models regarding the role of psychological discrepancy in belief change tested in this dissertation, the scale value constancy assumption was upheld, and it seemed that there should be no further attempts at theorizing the valuation process. However, the four tested models were among many possibilities that could have been proposed and tested. Future studies can indeed explore other models of the valuation process. In addition, a message with a vague advocated position may tell a different story from a message with an explicit advocated position considered in this dissertation. For example, consider a message that states *please donate a lot of money* rather than *please donate $100*. Future studies can explore how the meaning of "a lot of money" varies as a function of perspective.

Another forward-looking implication of this dissertation is that this dissertation finds it necessary to further examine the relationship between psychological discrepancy, message discrepancy, and perspective. This dissertation somewhat naively assumed $\psi = kD/P$, and only retrospectively found that the assumed relationship might have been implausible. Before any further modeling of the valuation process as a function of psychological discrepancy, there is an urge to clarify the relationship between ψ, D, and P. Without an established relationship among these three constructs, there would be no firm basis to manipulate ψ in a systematic way in future studies.

A third forward-looking implication of this dissertation is to apply the best supported weight-discounting model to explain the psychological processes involved in more specific research areas and to make new predictions for future studies. For example, Peng and Nisbett (1999) found that Chinese participants used a compromise approach to reconcile two seemingly contradictory scientific statements, whereas American participants used a differentiation approach. In their study, the American and Chinese participants were randomly assigned to three conditions. In the first two conditions, the participants read only Statement A or Statement B; in the third condition, the participants read both statements. The participants rated the plausibility of the statement that they read. Both the Chinese and the American participants believed the same statement was more plausible than the other statement based on the plausibility ratings from the single-statement conditions. However, when the participants were presented with both contradictory statements, the difference between the two statements' ratings on plausibility was greater than the difference computed from the single-statement conditions for the American participants. For the Chinese participants, the difference between the two ratings on plausibility was less than the difference computed from the single-statement conditions (Peng & Nisbett, 1999, Figure 5, p. 749).

Although there are some differences in the context between Peng and Nisbett's (1999) study and this dissertation (e.g., the former did not focus on changing a subject's belief position), if one considers the more plausible statement to be a subject's own belief, the psychological discrepancy between a Chinese subject's own belief and the belief of the less plausible statement was less than the psychological discrepancy between an American subject's own belief and the belief of the less plausible statement. Based on the results of this dissertation, the less plausible statement would weigh heavier for the Chinese subjects than for the American subjects. Thus, if one considers the difference between the plausibility of the more plausible statement rated in the double-statement condition and the plausibility of the more plausible statement rated in the single-statement condition as a proxy for the measure of belief change, the mean amount of belief change in the direction toward the less plausible statement for the Chinese subjects would be greater than the mean amount of belief change for the American subjects, which was reflected in Peng and Nisbett's (1999, p. 749) Figure 5. Future studies can have a more direct test of the above hypothesis by measuring the subject's own beliefs or belief positions. Further, future studies can explore whether the dialectical thinking pattern is related to a wider perspective that then leads to a smaller psychological discrepancy between a subject's initial belief position and a contradictory message.

The methodological significance of this dissertation is that it used a computational approach to derive some of the hypotheses based on the model equations. Unlike some theories in persuasion in which propositions about the relationship between concepts are stated verbally, the goal of using mathematical models is to make assumptions explicit, to derive hypotheses symbolically, and to make precise predictions (Kaplowitz, Fink, & Bauer, 1983, p. 233). Building belief change models in mathematical form could also make measurement models possible to develop (Torgerson, 1958). In certain situations, especially when working with a complex model, deducing a hypothesis based on analytic derivation is difficult. In this dissertation, such a difficulty was found when I expressed a first partial derivative of final position with respect to a parameter in purely symbolic terms in the complex model and then tried to determine the sign of this first partial derivative. To determine the sign of this first partial derivative by using an analytic proof alone was difficult. A computational approach was a workaround to this difficulty, and it approximated some behaviors of a first partial derivative with a large number of hypothetical data points. This dissertation showed that a hypothesis can be deduced using a computational approximation even for a complex model involving seemingly too many parameters.

Regarding the practical significance, this dissertation suggests that when a persuasive message is extremely discrepant from a message receiver's initial position, the importance of the message would be discounted, whereas the substantive meaning of the message would remain the same (e.g., a $100 request for donation would be perceived as a $100 request; see the example presented in Chapter 1). Therefore, to maximize the effectiveness of the persuasive message, the persuader can pay extra attention to improve, for example, the argument quality of the message and to maintain his or her credibility. More importantly, ethically using strategies to widen the message receiver's perspective (by extending either the lower bound or the upper bound depending on the sign of message discrepancy) before presenting an extremely discrepant message would make the message seem less discrepant. In this way, the message weight would be discounted less.

NOTES

[1]Note that another way to define effectiveness is that the effectiveness of a message is the ratio of the amount of belief change to the amount of advocated change, the latter of which is the difference between the message's advocated position and a subject's initial position before message receipt. If the amount of advocated change is held constant, the difference between the above definition and the definition used in the main text does not matter. That is, either definition is appropriate for the remaining discussion in the main text.

[2]This dissertation focuses on four static belief change models rather than a dynamic model (Chung et al., 2008). In a dynamic model that examines belief change over time, after message receipt, a subject's belief position can move toward or around an equilibrium position with different patterns of oscillation (i.e., oscillating either in an overdamped, a critically damped, or an underdamped case; Kaplowitz et al., 1983). This dissertation assumes that when a subject's belief position after message receipt is measured, the subject's belief position has reached a new equilibrium position. In other words, neither time nor any oscillating patterns is examined in the models in this dissertation.

[3]Equation 3 and Equation 4 assume that when R_{i-1} is integrated with s_i, the *independent* weight of R_{i-1} is $\sum_{j=0}^{i-1} w_j$. That is, for example, if $i = 2$, the weight of R_1 is $w_0 + w_1$. A more general form (S. Chung, personal communication, December 11, 2017) of Equation 4 is: $R_i - R_{i-1} = \frac{w_i}{w_{i-1}+w_i}(s_i - R_{i-1})$, where $w_{i-1} = f(w_0, w_1, w_2, \ldots, w_{i-1})$.

[4]The traditional method of information integration theory is to use within-subject factorial designs (Anderson, 1981). The rationale is that the valuation process varies greatly across individuals. That is, two individuals may interpret the same stimulus differently and may put different values on a dimension for that stimulus. The theoretical focus of information integration theory is on generalizing integration rules across individuals. If a between-subject design is used, the idiosyncratic valuation process may make the validation of integration rules more difficult. In a within-subject design, the valuation process can be made approximately constant across responses for all cells in a factorial table. One drawback of a within-subject design, however, is that a subject's fatigue may lead to a less reliable judgment over time. Another drawback is the anchoring effect, where the extremity of a message's scale value affects judgment (Wegener, Petty, Detweiler-Bedell, & Jarvis, 2001).

Also, Anderson (1981) argued that analysis should first be done for each subject's data. If each subject's data show parallelism, the group data averaged across individuals will also show parallelism. The reverse is not true because deviations from parallelism at the individual level could cancel out, which makes the group data look like that parallelism is supported.

[5]The historical context of a 15% increase being reported as the most extremely moderate and the most moderately extreme position by the subjects was that the annual inflation rates were 11.35% and 13.50% for the year 1979 and the year 1980 in the U.S., respectively. The annual inflation rates in the U.S. were calculated based on the monthly consumer price index (CPI; using the all urban consumers, all items, and non-seasonally-

adjusted data; U.S. Bureau of Labor Statistics, 2020). The annual inflation rate for the year 1979, for example, was calculated based on this formula: $(CPI_{1979} - CPI_{1978})/CPI_{1978}$, where CPI_{1979} and CPI_{1978} are the 12-month average CPIs for the year 1979 and the year 1978, respectively.

The secondary data from the World Bank (2020) showed that the annual inflation rates were 11.25% and 13.55% for the year 1979 and the year 1980 in the U.S., respectively, which somewhat corroborates the calculated inflation rates based on the primary CPI data from the U.S. Bureau of Labor Statistics (i.e., 11.35% for 1979 and 13.50% for 1980). The small discrepancies might be due to the possibility that the World Bank calculated the inflation rate for a given year using a different version of CPI from the one specified in the last paragraph.

Note that in Fink et al. (1983, p. 423), the annual inflation rate was 11% when the Fall 1979 data were collected, and the annual inflation rate was 18% when the Spring 1980 data were collected. The source of these numbers was unclear. To verify these numbers against the primary CPI data from the U.S. Bureau of Labor Statistics (2020), I used the five-month average during a fall semester (i.e., the average over August, September, October, November, and December) and the five-month average during a spring semester (i.e., the average over January, February, March, April, and May) to calculate the annual inflation rates in Fall 1979 (compared with Fall 1978) and in Spring 1980 (compared with Spring 1979). The results were 11.03% and 14.40% for Fall 1979 and Spring 1980, respectively. For Fall 1979, the 11% inflation rate was verified. For Spring 1980, the discrepancy between the 18% inflation rate reported in Fink et al. (1983,

p. 423) and the calculated inflation rate of 14% reported in this dissertation was probably due to the fact that the reference base of the current version of the historical CPI data (U.S. Bureau of Labor Statistics, 2020) is the 36-month period from the years 1982 to 1984 (i.e., the 36-month average CPI from 1982 to 1984 was set as 100). In other words, Fink et al. (1983) must have used a set of CPI data that had a different reference base from the current version of CPI when they calculated the annual inflation rates. This discrepancy did not change the argument, given in Fink et al. (1983, p. 423), that perhaps the historical context of the increase in the annual inflation rates gave rise to the fact that the spring-study subjects recommended a higher tuition increase and found the message less psychologically discrepant than the fall-study subjects.

[6]The goal of translating a partial η^2 to f was to get a sense of the magnitude of an effect size according to Cohen's (1988) proposed convention (see the following paragraph in the main text). "To get a sense" means that caution had been taken to not equate the *sample* effect size with the *population* effect size (Faul et al., 2007, p. 176). This is also why getting a 95% confidence interval of the effect size can improve one's understanding of the population effect size based on the sample effect size.

The effect size index f is σ_m/σ where $\sigma_m = \{[\sum_{i=1}^{k}(m_i - m)^2]/k\}^{1/2}$, k is the number of means to be compared, m_i is the ith mean, m is the grand mean, and σ is the standard deviation of the total population (Cohen, 1988, pp. 274-275). Cohen (1988) described f as "the standard deviation of the standardized means" (p. 275). The conversion formula from f to η^2 is: $\eta^2 = f^2/(1 + f^2)$ (Cohen, 1988, p. 281).

[7]Cohen (1988, p. 56) proposed a power of .80 to be a minimally acceptable level, given an alpha level of .05. With the same alpha level, Holbert et al. (2018) considered a .95 power level as “more than adequate” (p. 83).

[8]Using AIC in model selection assumes a normal distribution of the regression residuals (Burnham & Anderson, 2004, p. 268). An examination on the distributions of the regression residuals from the models in Table 32 revealed a significant Shapiro-Wilk test of normality in all four models. Thus, the estimated residuals did not follow a normal distribution across the models. The skewness statistics reported in Table 33 also showed that the residuals were significantly positively skewed across the models, although for the sake of model selection, the most plausible weight-discounting model also had the least skewed residuals.

The nonnormal distribution of the residuals indicated some other alternative models to the ones in Table 32 could be more plausible. An alternative model was specified by adding a linear term for religiosity. This was done for all four models in Table 32 in order to give these models a fair comparison in model selection. Choosing religiosity was based on the observation that religiosity significantly correlated with the subject’s final position, $r = .223$, $p < .001$, $n = 402$. The magnitude of correlation with final position was the highest among several control variables. Religiosity was measured with a four-item magnitude scale based on the one used in Curry’s (1996) study. After each item was winsorized at the 95th percentile and square-root transformed, the four-item measure was reliable, Cronbach’s $\alpha = .98$, $n = 402$. The average of the four items was used in the following analysis, $M = 4.59$, $SD = 5.31$, $Sk = 0.71$ ($SE = 0.12$),

Ku = -0.93 (SE = 0.24), n = 402.

For the weight-discounting model, with the inclusion of the linear term for religiosity, the skewness, although still significantly different from zero, reduced to 0.73 (SE = 0.12). This alternative weight-discounting model (AIC = 721.92; AIC_c = 722.14; BIC = 745.90) fit the data better than the weight-discounting model in Table 32 (see also Table 33). The estimated coefficient for the religiosity term was 0.03, SE = 0.01, 95% CI = [0.02, 0.04]. For the scale-value-pullback model, with the inclusion of the linear term for religiosity, the skewness reduced to 0.73 (SE = 0.12). This alternative scale-value-pullback model (AIC = 722.95; AIC_c = 723.16; BIC = 746.93) fit the data significantly better than the scale-value-pullback model in Table 32 (see also Table 33). The estimated coefficient for the religiosity term was 0.03, SE = 0.01, 95% CI = [0.02, 0.04]. For the restricted complex model and the full complex model, after adding the linear term for religiosity, an optimal solution could not be found after 999 iterations (the maximum allowed in SPSS).

In summary, this sensitivity test specified alternative models to the ones in Table 32. The alternative weight-discounting model and the alternative scale-value-pullback model had a milder violation of the normality assumption for the regression residuals than the models in Table 32. The results based on AIC, AIC_c, and BIC revealed that the weight-discounting model was still the most plausible model with the scale-value-pullback model as a close second.

REFERENCES CITED

Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression formation. *Journal of Experimental Psychology*, *70*, 394–400. doi:10.1037/h0022280

Anderson, N. H. (1968). Application of a linear-serial model to a personality-impression task using serial presentation. *Journal of Personality and Social Psychology*, *10*, 354–362. doi:10.1037/h0026816

Anderson, N. H. (1971a). Integration theory and attitude change. *Psychological Review, 78*, 171–206. doi:10.1037/h0030834

Anderson, N. H. (1971b). Two more tests against change of meaning in adjective combinations. *Journal of Verbal Learning and Verbal Behavior*, *10*, 75–85. doi:10.1016/s0022-5371(71)80097-x

Anderson, N. H. (1973). Serial position curves in impression formation. *Journal of Experimental Psychology*, *97*, 8–12. doi:10.1037/h0033774

Anderson, N. H. (1981). *Foundations of information integration theory*. San Diego, CA: Academic Press.

Anderson, N. H. (2008). *Unified social cognition*. New York, NY: Psychology Press.

Anderson, N. H., & Alexander, G. R. (1971). Choice test of the averaging hypothesis for information integration. *Cognitive Psychology*, *2*, 313–324. doi:10.1016/0010-0285(71)90017-x

Anderson, N. H., & Hubert, S. (1963). Effects of concomitant verbal recall on order effects in personality impression formation. *Journal of Verbal Learning and Verbal Behavior*, *2*, 379–391. doi:10.1016/s0022-5371(63)80039-0

Anderson, N. H., & Jacobson, A. (1965). Effect of stimulus inconsistency and discounting instructions in personality impression formation. *Journal of Personality and Social Psychology*, *2*, 531–539. doi:10.1037/h0022484

Aronson, E., Turner, J. A., & Carlsmith, J. M. (1963). Communicator credibility and communication discrepancy as determinants of opinion change. *The Journal of Abnormal and Social Psychology*, *67*, 31–36. doi:10.1037/h0045513

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258–290. doi:10.1037/h0055756

Bentler, P. M., & Yuan, K.-H. (1999). Structural equation modeling with small samples: Test statistics. *Multivariate Behavioral Research*, *34*, 181–197. doi:10.1207/s15327906mb340203

Blanton, H., & Jaccard, J. (2006). Tests of multiplicative models in psychology: A case study using the unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *113*, 155–166. doi:10.1037/0033-295x.113.1.155

Bochner, S., & Insko, C. A. (1966). Communicator discrepancy, source credibility, and opinion change. *Journal of Personality and Social Psychology*, *4*, 614–621. doi:10.1037/h0021192

Brock, T. C. (1967). Communication discrepancy and intent to persuade as determinants of counterargument production. *Journal of Experimental Social Psychology*, *3*, 296–309. doi:10.1016/0022-1031(67)90031-5

Bullock, J. G., Green, D. P., & Ha, S. E. (2008). *Experimental approaches to mediation: A new guide for assessing causal pathways*. Unpublished manuscript, Yale University, New Haven, CT.

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*, 550–558. doi:10.1037/a0018933

Bullock, J. G., & Ha, S. E. (2011). Mediation analysis is harder than it looks. In J. N. Druckman, D. P. Green, J. H. Kuklinski, & A. Lupia (Eds.), *Cambridge handbook of experimental political science* (pp. 508–521). New York, NY: Cambridge University Press.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*, 261–304. doi:10.1177/0049124104268644

Cacioppo, J. T., Petty, R. E., & Morris, K. J. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology*, *45*, 805–818. doi:10.1037/0022-3514.45.4.805

Chung, S., & Fink, E. L. (2017). *Mathematical models of the effect of message discrepancy on belief change: Previous models and a modified psychological discounting model*. Manuscript submitted for publication.

Chung, S., Fink, E. L., & Kaplowitz, S. A. (2008). The comparative statics and dynamics of beliefs: The effect of message discrepancy and source credibility. *Communication Monographs*, *75*, 158–189. doi:10.1080/03637750802082060

Chung, S., Fink, E. L., Waks, L., Meffert, M. F., & Xie, X. (2012). Sequential information integration and belief trajectories: An experimental study using candidate evaluations. *Communication Monographs*, *79*, 160–180. doi:10.1080/03637751.2012.673001

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Curry, T. R. (1996). Conservative Protestantism and the perceived wrongfulness of crimes: A research note. *Criminology*, *34*, 453–464. doi:10.1111/j.1745-9125.1996.tb01215.x

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, *81*, 95–106. doi:10.1037/h0037613

Dillard, J. P., & Shen, L. (Eds.). (2012). *The SAGE handbook of persuasion: Developments in theory and practice* (2nd ed.). Thousand Oaks, CA: Sage.

Dillehay, R. C., & Berger, P. K. (1969). Permissive introduction and anchoring in altering perceptual-judgmental processes and the impact of a persuasive communication. *Journal of Experimental Social Psychology*, *5*, 417–428. doi:10.1016/0022-1031(69)90034-1

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace.

Eagly, A. H., & Telaak, K. (1972). Width of the latitude of acceptance as a determinant of attitude change. *Journal of Personality and Social Psychology*, *23*, 388–397. doi:10.1037/h0033161

Eisinga, R., te Grotenhuis, M., & Pelzer, B. (2012). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, *58*, 637–642. doi:10.1007/s00038-012-0416-3

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. doi:10.3758/bf03193146

Fink, E. L. (1996). Dynamic social impact theory and the study of human communication. *Journal of Communication, 46*(4), 4–12. doi:10.1111/j.1460-2466.1996.tb01500.x

Fink, E. L. (2009). The FAQs on data transformation. *Communication Monographs, 76*, 379–397. doi:10.1080/03637750903310352

Fink, E. L., & Cai, D. (2012). Discrepancy models of belief change. In J. P. Dillard & L. Shen (Eds.), *The SAGE handbook of persuasion: Developments in theory and practice* (2nd ed., pp. 84–103). Thousand Oaks, CA: Sage.

Fink, E. L., Kaplowitz, S. A., & Bauer, C. L. (1983). Positional discrepancy, psychological discrepancy, and attitude change: Experimental tests of some mathematical models. *Communication Monographs*, *50*, 413–430. doi:10.1080/03637758309390178

Fink, E. L., Monahan, J. L., & Kaplowitz, S. A. (1989). A spatial model of the mere exposure effect. *Communication Research*, *16*, 746–769. doi:10.1177/009365089016006002

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*, 889–906. doi:10.1037/0022-3514.38.6.889

Goodboy, A. K., & Kline, R. B. (2017). Statistical and practical concerns with published communication research featuring structural equation modeling. *Communication Research Reports*, *34*, 68–77. doi:10.1080/08824096.2016.1214121

Granberg, D., & Steele, L. (1974). Procedural considerations in measuring latitudes of acceptance, rejection, and noncommitment. *Social Forces*, *52*, 538–542. doi:10.2307/2576997

Hamilton, M. A., Hunter, J. E., & Boster, F. J. (1993). The elaboration likelihood model as a theory of attitude formation: A mathematical analysis. *Communication Theory*, *3*, 50–65. doi:10.1111/j.1468-2885.1993.tb00056_3_1.x

Hanushek, E. A., & Jackson, J. E. (1977). *Statistical methods for social scientists*. Orlando, FL: Academic Press.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (1st ed.). New York, NY: Guilford.

Himmelfarb, S. (1975). On scale value and weight in the weighted averaging model of integration theory. *Personality and Social Psychology Bulletin, 1*, 580–583. doi:10.1177/014616727500100406

Himmelfarb, S., & Anderson, N. H. (1975). Integration theory applied to opinion attribution. *Journal of Personality and Social Psychology, 31*, 1064–1072. doi:10.1037/h0076955

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, *24*, 1–55. doi:10.1016/0010-0285(92)90002-j

Holbert, R. L., Hardy, B. W., Park, E., Robinson, N. W., Jung, H., Zeng, C., . . . Sweeney, K. (2018). Addressing a statistical power-alpha level blind spot in political- and health-related media research: Discontinuous criterion power analyses. *Annals of the International Communication Association*, *42*, 75–92. doi:10.1080/23808985.2018.1459198

Hunter, J. E., Danes, J. E., & Cohen, S. H. (Eds.). (1984). *Mathematical Models of Attitude Change: Vol. 1*. *Change in single attitudes and cognitive structure*. New York, NY: Academic Press.

IBM. (2011). *IBM SPSS Statistics 20 command syntax reference*. Armonk, NY: IBM Corp.

Johnson, B. T., Lin, H. Y., Symons, C. S., Campbell, L. A., & Ekstein, G. (1995). Initial beliefs and attitudinal latitudes as factors in persuasion. *Personality and Social Psychology Bulletin*, *21*, 502–511. doi:10.1177/0146167295215008

Johnson, B. T., Maio, G. R., & Smith-McLallen, A. (2005). Communication and attitude change: Causes, processes, and effects. In D. Albarracín, B. T. Johnson, & M. P.

Zanna (Eds.), *The handbook of attitudes* (pp. 617–669). Mahwah, NJ: Lawrence Erlbaum Associates.

Judd, C. M., & DePaulo, B. M. (1979). The effect of perspective differences on the measurement of involving attitudes. *Social Psychology Quarterly*, *42*, 185–189. doi:10.2307/3033700

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kaplowitz, S. A., & Fink, E. L. (with Mulcrone, J., Atkin, D., & Dabil, S.). (1991). Disentangling the effects of discrepant and disconfirming information. *Social Psychology Quarterly*, *54*, 191–207. doi:10.2307/2786650

Kaplowitz, S. A., & Fink, E. L. (1997). Message discrepancy and persuasion. In G. A. Barnett & F. J. Boster (Eds.), *Progress in communication sciences* (Vol. 13, pp. 75-106). Greenwich, CT: Ablex.

Kaplowitz, S. A., Fink, E. L., Armstrong, G. B., & Bauer, C. L. (1986). Message discrepancy and the persistence of attitude change: Implications of an information integration model. *Journal of Experimental Social Psychology*, *22*, 507–530. doi:10.1016/0022-1031(86)90048-x

Kaplowitz, S. A., Fink, E. L., & Bauer, C. L. (1983). A dynamic model of the effect of discrepant information on unidimensional attitude change. *Behavioral Science*, *28*, 233–250. doi:10.1002/bs.3830280306

Kline, R. B. (2012). Assumptions in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 111–125). New York, NY: Guilford.

Laroche, M. (1977). A model of attitude change in groups following a persuasive communication: An attempt at formalizing research findings. *Behavioral Science*, *22*, 246–257. doi:10.1002/bs.3830220403

Latané, B. (1996). Dynamic social impact: The creation of culture by communication. *Journal of Communication, 46*(4), 13–25. doi:10.1111/j.1460-2466.1996.tb01501.x

Levine, T. R., Weber, R., Park, H. S., & Hullett, C. R. (2008). A communication researchers' guide to null hypothesis significance testing and alternatives. *Human Communication Research*, *34*, 188–209. doi:10.1111/j.1468-2958.2008.00318.x

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198–1202. doi:10.1080/01621459.1988.10478722

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*, 433–442. doi:10.3758/s13428-016-0727-z

Meffert, M. F., Chung, S., Joiner, A. J., Waks, L., & Garst, J. (2006). The effects of negativity and motivated information processing during a political campaign. *Journal of Communication*, *56*, 27–51. doi:10.1111/j.1460-2466.2006.00003.x

Meurer, A., Smith, C. P., Paprocki, M., Čertík, O., Kirpichev, S. B., Rocklin, M., . . . Scopatz, A. (2016). SymPy: Symbolic computing in Python. *PeerJ Computer Science*, *3*. doi:10.7717/peerj-cs.103

Miller, G. R. (2012). On being persuaded: Some basic distinctions. In J. P. Dillard & L. Shen (Eds.), *The SAGE handbook of persuasion: Developments in theory and practice* (2nd ed., pp. 84–103). Thousand Oaks, CA: Sage.

O'Keefe, D. J. (1990). *Persuasion: Theory and research*. Newbury Park, CA: Sage.

Ostrom, T. M. (1970). Perspective as a determinant of attitude change. *Journal of Experimental Social Psychology*, *6*, 280–292. doi:10.1016/0022-1031(70)90063-6

Ostrom, T. M., & Davis, D. (1979). Idiosyncratic weighting of trait information in impression formation. *Journal of Personality and Social Psychology*, *37*, 2025–2043. doi:10.1037/0022-3514.37.11.2025

Ostrom, T. M., & Upshaw, H. S. (1968). Psychological perspective and attitude change. In A. G. Greenwald, T. C. Brock, & T. M. Ostrom (Eds.), *Psychological foundations of attitudes* (pp. 217–242). New York, NY: Academic Press.

Peer, E., Vosgerau, J., & Acquisti, A. (2013). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, *46*, 1023–1031. doi:10.3758/s13428-013-0434-y

Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist*, *54*, 741–754. doi:10.1037/0003-066x.54.9.741

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879–891. doi:10.3758/brm.40.3.879

Python Software Foundation. (2017). Python documentation (Version 2.7.14). Retrieved from: https://docs.python.org/2/

Qualtrics [Computer software]. (2018). Provo, UT: Qualtrics.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*, 1–36. doi:10.18637/jss.v048.i02

Saltiel, J., & Woelfel, J. (1975). Inertia in cognitive processes: The role of accumulated information in attitude change. *Human Communication Research*, *1*, 333–344. doi:10.1111/j.1468-2958.1975.tb00282.x

Sherif, C. W., Kelly, M., Rodgers, H. L., Sarup, G., Jr., & Tittler, B. I. (1973). Personal involvement, social judgment, and action. *Journal of Personality and Social Psychology*, *27*, 311–328. doi:10.1037/h0034948

Sherif, C. W., Sherif, M., & Nebergall, R. E. (1965). *Attitude and attitude change*. Philadelphia, PA: Saunders.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven, CT: Yale University Press.

Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. *Sociological Methodology*, *13*, 290–312. doi:10.2307/270723

Spiess, A.-N., & Neumeyer, N. (2010). An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach. *BMC Pharmacology*, *10*. doi:10.1186/1471-2210-10-6

SPSS (Version 20.0.0) [Computer software]. (2011). Armonk, NY: IBM Corp.

Steiger, J. H. (2004). Beyond the *F* test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychological Methods*, *9*, 164–182. doi:10.1037/1082-989x.9.2.164

Sulfaro, V. A., & Crislip, M. N. (1997). How Americans perceive foreign policy threats: A magnitude scaling analysis. *Political Psychology*, *18*, 103–126. doi:10.1111/0162-895x.00047

Tesser, A. (1968). Differential weighting and directed meaning as explanations of primacy in impression formation. *Psychonomic Science*, *11*, 299–300. doi:10.3758/bf03328201

Torgerson, W. S. (1958). *Theory and methods of scaling*. New York, NY: Wiley.

Ullman, J. B. (2006). Structural equation modeling: Reviewing the basics and moving forward. *Journal of Personality Assessment*, *87*, 35–50. doi:10.1207/s15327752jpa8701_03

U.S. Bureau of Labor Statistics. (2020). *Historical consumer price index for all urban consumers (CPI-U): U.S. city average, all items, by month* [Data file]. Retrieved from https://www.bls.gov/cpi/tables/supplemental-files/historical-cpi-u-202001.pdf

U.S. Department of Labor. (2009). *Minimum wage*. Retrieved from https://www.dol.gov/agencies/whd/minimum-wage

Weber, R., & Popova, L. (2012). Testing equivalence in communication research: Theory and application. *Communication Methods and Measures*, *6*, 190–213. doi:10.1080/19312458.2012.703834

Wegener, D. T., Petty, R. E., Detweiler-Bedell, B. T., & Jarvis, W. B. G. (2001). Implications of attitude change theories for numerical anchoring: Anchor plausibility and the limits of anchor effectiveness. *Journal of Experimental Social Psychology*, *37*, 62–69. doi:10.1006/jesp.2000.1431

World Bank. (2020). *Inflation, consumer prices (annual %)* [Data file]. Retrieved from http://api.worldbank.org/v2/en/indicator/FP.CPI.TOTL.ZG?downloadformat=excel

Zalinski, J., & Anderson, N. H. (1989). Measurement of importance in multiattribute models. In J. B. Sidowski (Ed.), *Conditioning, cognition, and methodology: Contemporary issues in experimental psychology* (pp. 177–205). Lanham, MD: University Press of America.

Zhao, X., & Fink, E. L. (2018, May). *Two routes to the boomerang effect: Proattitudinal versus counterattitudinal message*s. Paper presented at the annual meeting of the International Communication Association, Prague, Czech Republic.

APPENDIX A

PROOF FOR THE RELATIONSHIP BETWEEN *s* AND *R* IN THE WEIGHT-DISCOUNTING MODEL

The goal of the following procedure is to provide proof for the nonmonotonic relationship between s and R (with R first increases and then decreases, as s increases) in the weight-discounting model (Equation 8) for $s > s_0$, where all the parameters are assumed to be positive and $P = U - L > 0$. The first partial derivative of R with respect to s is as follows: $\frac{\partial R}{\partial s} = \frac{\exp(\gamma k|s_0\text{-}s|/P)ww_0[P|s_0\text{-}s|\text{-}\gamma k(s_0\text{-}s)^2] + P|s_0\text{-}s|w^2}{Pw_0^2\exp(2\gamma k|s_0\text{-}s|/P)|s_0\text{-}s| + 2Pww_0\exp(\gamma k|s_0\text{-}s|/P)|s_0\text{-}s| + Pw^2|s_0\text{-}s|}$.

To determine the sign of $\frac{\partial R}{\partial s}$, because the denominator is positive, the task is to evaluate the sign of the numerator. Let the numerator be greater than zero. Then, the algebraic manipulation can lead to: $Pw/w_0 > \exp(\gamma k|s_0\text{-}s|/P)(\gamma k|s_0\text{-}s|\text{-}P)$. When $s \to s_0$, $|s_0 - s| \to 0$, and $\gamma k|s_0\text{-}s|\text{-}P$ is negative. Because Pw/w_0 is positive and $\exp(\gamma k|s_0\text{-}s|/P) \to 1$, the inequality $Pw/w_0 > \exp(\gamma k|s_0\text{-}s|/P)(\gamma k|s_0\text{-}s|\text{-}P)$ holds. Therefore, when $s \to s_0$, the numerator of $\frac{\partial R}{\partial s}$ is positive. As s approaches positive infinity, ceteris paribus, there exists one and only one value of s at which the right side of the above inequality becomes greater than the inequality's left side. That is, there exists one and only value of s at which the numerator becomes negative. In other words, there exists one and only value of s at which R starts to decrease.

Q.E.D.

APPENDIX B

PROOF FOR THE RELATIONSHIP BETWEEN *s* AND *R* IN THE SCALE-VALUE-PULLBACK MODEL

The goal of the following procedure is to provide proof for the relationship between s and R depending on the value of P in the scale-value-pullback model (Equation 10) for $s > s_0$, where all the parameters are assumed to be positive and $P = U - L > 0$. The first partial derivative of R with respect to s is as follows: $\frac{\partial R}{\partial s} = \frac{[P|s_0\text{-}s|+\gamma ks(s_0\text{-}s)]w}{P\exp(\gamma k|s_0\text{-}s|/P)|s_0\text{-}s|(w_0+w)}$.

To determine the sign of $\frac{\partial R}{\partial s}$, because the denominator is positive, the task is to evaluate the sign of the numerator. Note that because $s > s_0$, the numerator of $\frac{\partial R}{\partial s}$ can be rewritten as: $[-P(s_0\text{-}s) + \gamma ks(s_0\text{-}s)]w$. Let $[-P(s_0\text{-}s) + \gamma ks(s_0\text{-}s)]w > 0$. After algebraic manipulation, the following inequality can be derived: $P > \gamma ks$. That is, when $s \to s_0$, if $P > \gamma ks$ holds, R is an increasing function of s. As s approaches positive infinity, ceteris paribus, there exists one and only one value of s at which γks becomes greater than P. That is, there exists one and only value of s at which the numerator becomes negative. In other words, there exists one and only value of s at which R starts to decrease.

When does $P > \gamma ks$ hold? Given a value for each of the four parameters, there are three possible situations: $P > \gamma ks$, $P = \gamma ks$, and $P < \gamma ks$. If γ, k, and s are held constant, the value of P determines whether $P > \gamma ks$, $P = \gamma ks$, or $P < \gamma ks$ holds.

If $P = \gamma ks$, when $s \rightarrow s_0$, $\frac{\partial R}{\partial s} \rightarrow 0$. As s approaches positive infinity, ceteris paribus, $P < \gamma ks$ always holds. That is, as s approaches positive infinity, ceteris paribus, R is a decreasing function of s.

If $P < \gamma ks$, when $s \rightarrow s_0$, R is a decreasing function of s. As s approaches positive infinity, ceteris paribus, $P < \gamma ks$ always holds. That is, as s approaches positive infinity, ceteris paribus, R remains to be a decreasing function of s.

Q.E.D.

APPENDIX C

PROOF FOR THE BOOMERANG EFFECT IN THE WEIGHT-DISCOUNTING MODEL, THE SCALE-VALUE-PULLBACK MODEL, AND THE COMPLEX MODEL

The goal of the following procedure is to provide proof for the nonexistence of the boomerang effect in the weight-discounting model and in the complex model, and the existence of the boomerang effect in the scale-value-pullback model. A boomerang effect is defined as the situation when $R < s_0$.

For the weight-discounting model (Equation 8),

$R - s_0 = \frac{w_0 s_0 + w\Delta(\psi)s}{w_0 + w\Delta(\psi)} - s_0 = \frac{w\exp(-\gamma k|s_0 - s|/P)(s - s_0)}{w_0 + w\exp(-\gamma k|s_0 - s|/P)}$. Because all parameters are assumed to be positive and $s > s_0$, $R - s_0 > 0$ always holds.

For the scale-value-pullback model (Equation 10 when $s > s_0$),

$R - s_0 = \frac{w_0 s_0 + w\Delta(\psi)s}{w_0 + w} - s_0 = \frac{w(s\exp(-\gamma k|s_0 - s|/P) - s_0)}{w_0 + w}$. As $s \rightarrow s_0$, $\exp(-\gamma k|s_0\text{-}s|/P) \rightarrow 1$, because $s > s_0$, $R > s_0$. As $s \rightarrow +\infty$, $s\exp(-\gamma k|s_0\text{-}s|/P) \rightarrow 0$, $s\exp(-\gamma k|s_0\text{-}s|/P) - s_0 < 0$. Therefore, as s increases, there exists one and only one point of s when $R - s_0$ become negative.

For the complex model (Equation 11), $R - s_0 = \frac{w_0 s_0 + w\Delta(\psi)_w \cdot [s\Delta(\psi)_s + s_0]}{w_0 + w\Delta(\psi)_w} - s_0 =$

$\frac{w\exp(-\gamma k|s_0 - s|/P)sk'[1 - \exp(-\gamma k|s_0 - s|/P)]}{w_0 + w\exp(-\gamma k|s_0 - s|/P)}$. Because all parameters are assumed to be positive and $s > s_0$, $R - s_0 > 0$ always holds.

Q.E.D.

APPENDIX D

PROOF FOR THE RELATIONSHIP BETWEEN *s* AND *R* IN THE COMPLEX MODEL

The goal of the following procedure is to provide proof for the nonmonotonic relationship between s and R (with R first increases and then decreases, as s increases) in the complex model (Equation 11) for $s > s_0$, where all the parameters are assumed to be positive and $P > 0$. A partial analytic proof shows that when s is close to s_0, R is an increasing function of s. A computational approximation further shows that, after s reaches certain value, R is a decreasing function of s. The relationship between s and R demonstrated by the partial derivative (see below) assumes that all the parameters except s are constants.

The first partial derivative of R with respect to s is as follows:

$$\frac{\partial R}{\partial s} = \frac{[(2\gamma kk's+(L\text{-}U)k')\exp((2\gamma ks_0+\gamma ks)/(U\text{-}L))+(\text{-}\gamma kk's+(U\text{-}L)k')\exp((\gamma ks_0+2\gamma ks)/(U\text{-}L))]ww_0}{(U\text{-}L)w_0^2\exp(3\gamma ks_0/(U\text{-}L))+2(U\text{-}L)ww_0\exp((\gamma ks_0+2\gamma ks)/(U\text{-}L))+(U\text{-}L)w^2\exp((2\gamma ks_0+\gamma ks)/(U\text{-}L))}$$

$$+\frac{[(U\text{-}L)k'\exp((2\gamma ks_0+\gamma ks)/(U\text{-}L))+(\gamma kk's+(L\text{-}U)k')\exp(3\gamma ks_0/(U\text{-}L))]w^2}{(U\text{-}L)w_0^2\exp(3\gamma ks_0/(U\text{-}L))+2(U\text{-}L)ww_0\exp((\gamma ks_0+2\gamma ks)/(U\text{-}L))+(U\text{-}L)w^2\exp((2\gamma ks_0+\gamma ks)/(U\text{-}L))}.$$

With the assumptions that all parameters are positive, and $U - L > 0$, the denominator of $\frac{\partial R}{\partial s}$ is positive. Now examine the numerator of $\frac{\partial R}{\partial s}$. Let $s \to s_0$, then

$\exp[(2\gamma ks_0+\gamma ks)/(U\text{-}L)] \to \exp[3\gamma ks_0/(U\text{-}L)]$, and

$\exp[(\gamma ks_0+2\gamma ks)/(U\text{-}L)] \to \exp[3\gamma ks_0/(U\text{-}L)]$. The limit of the numerator of $\frac{\partial R}{\partial s}$ with respect to s, as s approaches s_0, is $\gamma kk's(ww_0+w^2)\exp[3\gamma ks_0/(U\text{-}L)]$.

With the assumption that all parameters are positive, the value of the limit above, $\gamma kk's(ww_0+w^2)\exp[3\gamma ks_0/(U\text{-}L)]$, is positive. Therefore, the partial derivative, $\frac{\partial R}{\partial s}$, is positive. This means that when s approaches s_0 from the right, R increases, as s increases.

It is difficult to prove analytically that after s reaches a certain value, the partial derivative becomes negative. This dissertation provides a computational approximation for demonstrating the nonmonotonic relationship between s and R.

The strategy of the computational demonstration is to calculate the value of $\frac{\partial R}{\partial s}$ given the values of the parameters in $\frac{\partial R}{\partial s}$. For each of the five parameters, w_0, w, γ, k, and k', a number of five values were randomly drawn from normal distributions (see Table 46). For the normal distributions of w_0, w, and γ, the means and the standard deviations were set to 0.50 and 0.12 respectively, where $0 < w_0 < 1$, $0 < w < 1$, and $0 < \gamma < 1$. For the normal distribution of k, the mean and the standard deviation were set to 1.5 and 0.2 respectively. For the normal distribution of k', the mean and the standard deviation were set to 1.5 and 0.12 respectively. For the four parameters, w_0, w, k, and k', the specified ranges were arbitrary. However, without more information, there was either no reason for rejecting the selections of those ranges. For γ, the range was specified based on the estimated value in Fink et al.'s (1983) study.

The four selected values of $U - L$ were 7.0, 22.0, 52.0, and 112.0. The value of s_0 was fixed to 10.0. The values of $U - L$ and s_0 were selected based on the topic of an experimental study (Kaplowitz & Fink, 1991). The unidimensional scale of belief

Table 46

Normal Distributions of the Five Parameters (w_0*,* w*,* γ*,* k*, and* k'*) and the Randomly Drawn Values Used in the Computational Approximation That Demonstrates the Nonmonotonic Relationship Between* s *and* R *in the Complex Model*

	Normal distribution			
Parameter	Mean	Standard deviation	Constraint	5 drawn values
w_0	0.50	0.12	$0 < w_0 < 1$	.188, .447, .635, .298, .295
w	0.50	0.12	$0 < w < 1$	.764, .506, .413, .552, .431
γ	0.50	0.12	$0 < \gamma < 1$	.569, .728, .531, .594, .430
k	1.50	0.20	$k > 0$	1.207, 1.509, 1.487, 1.508, 1.694
k'	1.50	0.12	$k' > 0$	1.481, 1.474, 1.604, 1.408, 1.524

position was the number of years of sentencing for armed robbery. In the pilot of that study, participants believed that a reasonable number of years of sentencing for armed robbery was 10-years. Then, 10-years was used in a bogus sentencing guideline presented to the participants at the beginning of the main study, which served the purpose of stabilizing the scale values of initial belief across individuals as 10-years. Therefore, the value of s_0 was fixed to 10.0 in the computational approximation.

For the values of $P = U - L$, suppose that a manipulation of psychological discrepancy is done by varying the difference between U and L (see Equation 11, where $\psi = kD/(U - L)$), where U represents the harshest number of years of sentencing from all the previous trials in the U.S., and L represents the most lenient number of years of sentencing. If L is fixed as 8-years, and U has four values (15-years, 30-years, 60-years, and 120-years), then $U - L$ has four values: 7.0, 22.0, 52.0, and 112.0. These values were used in the computational approximation.

Based on the above description, there were 5 × 5 × 5 × 5 × 5 × 4 × 1 = 12,500 combinations for the seven parameters (i.e., w_0, w, γ, k, k', $U - L$, and s_0). Then, for each of the 12,500 combinations, $\frac{\partial R}{\partial s}$ was calculated by varying the value of s. For the initial run of the computational approximation, the range from which s varied was set to all integers between and including 10 and 200. Again, considering the topic of Kaplowitz and Fink's (1991) study, the participants in the main study received one message that contained the suggested number of years of sentencing from a bogus judge's decision. The number of years in the judge's decision was treated as the scale value of the message, s, in Equation 11. A lower bound of 10-year for s was chosen, because s_0 was fixed as 10-year, and the constraint of $s > s_0$ was imposed. In Equation 11, when $s = s_0$, there is no belief change. Thus, the domain of s in Equation 11 was set as $s > s_0$. The reason for including 10-years for s in this computational approximation was that I wanted to compute the instant rate of change in R with respect to s, as $s \rightarrow s_0$ (see also the partial analytic proof above). An upper bound of 200-year for s was chosen for the initial run. Two considerations determined what value to be used for the upper bound of s. First, the upper bound should be large enough to detect a critical point (see below). Second, the upper bound should be reasonable to represent a positive extreme value on the scale of belief position (i.e., the number of years of sentencing for armed robbery in this case).

Thus, for each of the 12,500 combinations, a number of $200 - 10 + 1 = 191$ derivatives were calculated. In total, a number of $12{,}500 \times 191 = 2{,}387{,}500$ derivatives were calculated in the initial run.

Then, for each of the 12,500 combinations, the pattern of how the 191 derivatives varied, as *s* increases, was examined. To demonstrate the nonmonotonic relationship between *s* and *R* (with *R* first increases and then decreases, as *s* increases), there are three necessary and sufficient conditions given the boundary of *s* as discussed above. First, when $s = 10$, the derivative is positive. Second, when *s* approaches positive infinity in theory, the derivative is negative. Practically, *s* needs to be large enough to detect a critical point (see the third condition below). Here, the value of 200 for *s* was considered a sufficiently large number and a reasonable positive extreme value on the scale of belief position. Note that this computational approximation cannot provide information about how the derivatives behave beyond the upper bound of *s*. That is, the findings of the computational approximation only apply to the specified domain of *s*, as well as the selected values of the other parameters. Third, there is one and only one critical point at which *R* begins to decrease, as *s* increases. This third condition means that for the series of 191 derivatives, there exists one and only one point where a positive derivative becomes a negative derivative. To determine whether a critical point exists, each pair of adjacent derivatives was evaluated with regard to whether the first derivative in a pair was positive, and the second derivative in that pair was negative. If the above situation is evaluated to be true for one pair of adjacent derivatives, one is added to the count of critical point for one combination. In total, $191 - 1 = 190$ pairs of adjacent derivatives were evaluated for each of the 12,500 combinations.

The above procedures were implemented in the Python programming language, Version 2.7.14 (Python Software Foundation, 2017). The results showed that, among the 12,500 combinations, all had a positive value of $\frac{\partial R}{\partial s}$ when $s = 10$; a number of 9,930 (79.44%) combinations had exactly one critical point, and a number of 2,570 (20.56%) had zero critical point; for those combinations that had exactly one critical point, all had a negative value of $\frac{\partial R}{\partial s}$ when $s = 200$; for those combinations that had zero critical point, all had a positive value of $\frac{\partial R}{\partial s}$ when $s = 200$. In other words, 79.44% of the combinations displayed the predicted nonmonotonic relationship between s and R. The remaining combinations showed an increasing monotonic relationship between s and R. The above results implicated that, if the upper bound of s was extended, the percentage of the combinations that displayed the predicted nonmonotonic relationship would increase.

The second run of the computational approximation focused only on those 2,570 combinations that had zero critical point in the initial run. The upper bound of s was extended to 400. The lower bound of s was still 10. The results showed that for these 2,570 combinations, a number of 2,535 (98.64%) combinations had exactly one critical point. A number of 35 (1.36%) combinations had zero critical point. For those 2,535 combinations that had exactly one critical point, all had a negative value of $\frac{\partial R}{\partial s}$ when $s = 400$. For those 35 combinations that had zero critical point, all had a positive value of $\frac{\partial R}{\partial s}$ when $s = 400$. These results supported the expectation that, when the upper bound of s

was extended, more combinations would display the predicted nonmonotonic relationship between s and R.

What if extending the upper bound of s for the 9,930 combinations that had exactly one critical point in the initial run? Would more critical points be detected beyond the initial upper bound of 200 for s? To answer these questions, a third run was done for all 12,500 combinations, where the upper bound of s for all combinations was extended to 500. The results showed that for the 12,500 combinations, all had a positive value of $\frac{\partial R}{\partial s}$ when $s = 10$; all had a negative value of $\frac{\partial R}{\partial s}$ when $s = 500$; all had one and only one critical point. Therefore, given the selected parameter values in the above procedure, this computational approximation demonstrated that the relationship between s and R is nonmonotonic (with R first increases and then decreases, as s increases). The nonmonotonic relationship between s and R can be seen in the function graph in Figure 1.

One point needs to be reemphasized regarding the computational approximation above. The approximation provided no information about the relationship between s and R beyond the upper bound of s (i.e., 500), within the selected increment interval of s, or outside the range of the selected values of the other parameters. However, given the large number of combinations of parameter values (12,500), and the large number of derivatives calculated in the third run ($12{,}500 \times 491 = 6{,}137{,}500$), the computational approximation provided a convincing case to support the predicted relationship between s and R in a complex mathematical system.

APPENDIX E

PROOF FOR THE RELATIONSHIP BETWEEN *P* AND *R* IN THE COMPLEX MODEL

The goal of the following procedure is to provide proof for the nonmonotonic relationship between P and R (with R first increases and then decreases, as P increases) in the complex model (Equation 11) for $s > s_0$, where all the parameters are assumed to be positive. A partial analytic proof shows that when P approaches positive infinity, R is a decreasing function of P. A computational approximation further shows that, when P approaches zero, R is an increasing function of P. The relationship between P and R demonstrated by the partial derivative (see below) assumes that all the parameters except P are constants.

The first partial derivative of R with respect to P is:

$$\frac{\partial R}{\partial P} = \frac{(\gamma kk's\exp[2\gamma k|s_0\text{-}s|/P]\text{-}2\gamma kk's\exp[\gamma k|s_0\text{-}s|/P])|s_0\text{-}s|ww_0\text{-}\gamma kk's|s_0\text{-}s|w^2}{P^2\exp[3\gamma k|s_0\text{-}s|/P]w_0^2 + 2P^2ww_0\exp[2\gamma k|s_0\text{-}s|/P] + P^2\exp[\gamma k|s_0\text{-}s|/P]w^2}.$$

Partial analytic proof:

With the assumptions that all parameters are positive, the denominator of $\frac{\partial R}{\partial P}$ is positive. Now examine the numerator of $\frac{\partial R}{\partial P}$. Let $P \to +\infty$, then $\exp[2\gamma k|s_0\text{-}s|/P] \to 1$ and $\exp[\gamma k|s_0\text{-}s|/P] \to 1$. Then, the limit of the numerator of $\frac{\partial R}{\partial P}$, as P approaches positive infinity, is $\text{-}\gamma kk's|s_0 - s|(ww_0 + w^2)$.

With the assumption that all parameters are positive, the value of the limit above, $-\gamma kk's|s_0 - s|(ww_0 + w^2)$, is negative. Therefore, the partial derivative $\frac{\partial R}{\partial P}$ is negative. This means that when *P* approaches positive infinity, *R* decreases, as *P* increases.

It is difficult to prove analytically that when *P* approaches zero, *R* is an increasing function of *P*, and there is one and only one critical point at which $\frac{\partial R}{\partial P}$ becomes negative from a positive value, as *P* increases. This dissertation uses a computational approximation to provide proof for the nonmonotonic relationship between *P* and *R*.

A computational approximation:

The strategy of the computational approximation is to calculate the value of $\frac{\partial R}{\partial P}$ given the values of the parameters in $\frac{\partial R}{\partial P}$. For each of the five parameters, w_0, w, γ, k, and k', a number of five values were randomly drawn from normal distributions (see Table 47). The four selected values of s were 11.0, 15.0, 30.0, and 50.0. Thus, there were $5 \times 5 \times 5 \times 5 \times 5 \times 4 = 12{,}500$ combinations for the 6 parameters (i.e., w_0, w, γ, s, k, and k').

Then, for each of the 12,500 combinations, $\frac{\partial R}{\partial P}$ was calculated by varying the value of *P*. For the initial run, the range from which *P* varied was set to all integers between and including 1 and 100. Thus, for each of the 12,500 combinations, a number of 100 derivatives were calculated. In total, a number of $12{,}500 \times 100 = 1{,}250{,}000$ derivatives were calculated.

Table 47

Normal Distributions of the Five Parameters (w_0*,* w*,* γ*,* k*, and* k'*) and the Randomly Drawn Values Used in the Computational Approximation That Demonstrates the Nonmonotonic Relationship Between* P *and* R *in the Complex Model*

Parameter	Normal distribution: Mean	Normal distribution: Standard deviation	Constraint	5 drawn values
w_0	0.50	0.12	$0 < w_0 < 1$	.578, .706, .633, .460, .366
w	0.50	0.12	$0 < w < 1$	.267, .465, .209, .471, .497
γ	0.50	0.12	$0 < \gamma < 1$	.524, .429, .571, .434, .449
k	1.50	0.20	$k > 0$	1.306, 1.948, 1.589, 1.149, 1.342
k'	1.50	0.20	$k' > 0$	1.593, 1.347, 1.484, 1.592, 0.924

Then, for each of the 12,500 combinations, the pattern of how the 100 derivatives varied as P increases was examined. To provide proof for the nonmonotonic relationship between P and R (with R first increases and then decreases, as P increases), there are three necessary and sufficient conditions. First, when $P = 1$, the derivative is positive. Second, when P approaches positive infinity, the derivative is negative. Here, the value of 100 for P was considered a sufficiently large number to detect a critical point (see below). Third, there is one and only one critical point at which R begins to decrease, as P increases. This third condition means that for the series of 100 derivatives, there exists one and only one point where a positive derivative becomes a negative derivative. To determine whether a critical point exists, each pair of adjacent derivatives was evaluated with regard to whether the first derivative in a pair was positive and the second derivative in that pair was negative.

With the above procedure, the results showed that, for the 12,500 combinations, 10,025 (80.20%) had a positive value of $\frac{\partial R}{\partial P}$ when $P = 1$ and had a negative value of $\frac{\partial R}{\partial P}$ when $P = 100$; for these 10,025 combinations, all had one and only one critical point. For the remaining 2,475 (19.80%) combinations, all had a negative value of $\frac{\partial R}{\partial P}$ when $P = 1$ and had a negative value of $\frac{\partial R}{\partial P}$ when $P = 100$; for these 2,475 combinations, all had zero critical point.

A second run let P vary from 0.3 to 80.2 with an increment step of 0.3. The results showed that, for the 2,500 combinations, all had a positive value of $\frac{\partial R}{\partial P}$ when $P =$ 0.3; all had a negative value of $\frac{\partial R}{\partial P}$ when $P = 80.2$; all had one and only one critical point. Therefore, given the selected parameter values in the above procedure, the computational approximation demonstrated that the relationship between P and R is nonmonotonic (with R first increases and then decreases, as P increases). The nonmonotonic relationship between P and R can be seen in the graph in Figure 2. The limitations of the computational demonstration are the same as the limitations discussed in Appendix D.

APPENDIX F

QUESTIONNAIRE FOR THE MAIN STUDY

We invite you to take part in research that studies public beliefs towards the criminal justice system. Please click and review the Consent Form.

Summary of key information in the consent form: Approximate completion time is 15 to 18 minutes; payment: $2.20; you will not receive the payment if you leave 3 questions unanswered in a row.

When you finish reviewing, please select one of the choices below. Your clicking on the "I consent" option indicates your permission to take part in this research.

Important survey instructions: There will be no back buttons throughout the survey. Please read all questions carefully and give each the carefully thought out answer it deserves. This is not a speed test. Please answer all the questions. Please also note that for some pages, the next button (represented by an arrow) will become visible after a few seconds. When you are done, please copy the 10-digit numerical code found at the end of this survey and paste the code to your MTurk task submission page.

There is great concern over whether the criminal justice system has the support of the community. Some people may believe that judges are very lenient towards criminals but ignore the suffering of victims of crimes. Others, however, may believe that overly harsh punishments make criminals more likely to commit crimes in the future and encourage cruelty in the society. If the law is to be supported, people must believe that the criminal justice system serves the needs of society. Although not everyone needs to agree with every judicial decision, it is necessary that the public understand the way in which the criminal justice system functions and the factors that can enter into sentencing decisions.

In this study, you will first be asked about your own experience with the criminal justice system and your acquaintance with people who work in this system. You will then be asked about your beliefs about the sentencing of a specific federal crime—armed bank robbery. The answers you provide will help make judges more aware of your concerns because your answers will be reported to the United States Sentencing Commission. This body is currently studying sentencing policies and practices in various states.

Below you will be asked some questions about your experience with the criminal justice system. Please answer them honestly. Be assured that your response is confidential.

Have you, or anyone you know, ever been a judge in criminal trials? Yes or no.

Have you, or anyone you know, ever been an attorney in criminal trials? Yes or no.

Have you, or anyone you know, ever been a police officer? Yes or no.

Have you, or anyone you know, ever been a member of assistant personnel in court (e.g., law clerk, security, staff attorney, etc.)? Yes or no.

Have you, or anyone you know, ever been a witness in a criminal trial? Yes or no.

Have you or anyone you know, ever been a defendant in a criminal trial? Yes or no.

How old are you?

The sentence you will be examining is for the crime of armed bank robbery. The United States Sentencing Commission has issued a Sentencing Guidelines Manual. These Guidelines, which are based on a consensus of legal experts, are to assist judges and provide some degree of consistency in sentencing. The Guidelines have also been found to be supported by the public.

Please review the following description of a recent case. In June 2017, Convict X robbed a bank by brandishing a lethal weapon in a suburban area of Pennsylvania. Because there was no resistance, the convict did not fire the weapon. Convict X took over $50,000. Also, the record shows that this was not the first time that Convict X violated the laws.

According to the Federal Sentencing Guidelines, the sentence for Convict X is 10 years in prison.

To make sure that you have absorbed all of the information above, please answer the following questions.

On the views of which group of people are the Sentencing Guidelines based?

Does the public support the Sentencing Guidelines? Yes or no.

Please think of some reasons that might justify the suggested sentence for Convict X.

What do you believe is the most appropriate sentence (in number of years) for Convict X?

The Sentencing Guidelines, however, are recommendations, not laws. Because many people believe that a judge must be able to take into account the special features of each case, the law permits a judge to pass a sentence that is greater or less than the Guidelines.

In the United States, for those armed bank robbery cases similar to Convict X's, the documented harshest sentence has been either 15, 30, or 50 years (subject's upper bound manipulated).

What do you believe is a maximally harsh sentence (in number of years) for Convict X?

What do you believe is a maximally lenient sentence (in number of years) for Convict X?

We will now present the statement of a judge made for Convict X's case. You will be asked to evaluate the judge's decision later.

When Judge F. Walters, a judge in a United States district court, passed the sentence on Convict X, he made the following statement.

> By threat of force and violence, you gained access to a great deal of money which was not rightfully yours. You brandished a lethal weapon and made it quite clear that you would not hesitate to use it if your crime were in any way resisted. Because there was no resistance, you did not fire your weapon, but the terror you instilled in all of those present will be with them for a very long time. Clearly, you played a major role in the planning and execution of this crime. Finally, your record shows that this is not the first time that you have flouted the laws in a civilized society. Therefore, I conclude that society must be protected from you for many years and I sentence you to either 15, 30, or 50 years in the penitentiary (message scale value manipulated).

We will next be asking you some questions using a kind of scale with which you may not be familiar. Although some scales have 100 as a maximum value, the scale that follows treats 100 as a moderate value and has no upper limit.

Here is a practice task for you to become more familiar with the scale you will use. Imagine that there is a scale to measure the psychological difference between any two colors. A value of 100 represents a moderate level of difference. Zero represents "Not at all different." There is no upper limit. You may use any number between 0 and 100 for two colors that are less than moderately different, and you may use any number greater than 100 for two colors that are more than moderately different. If you think a difference is twice as much as a moderate difference, rate it 200. If you think a difference is half as much as a moderate difference, rate it 50. If you think two colors are not at all different, rate it 0 (zero). Please answer the following questions.

What is the psychological difference between red and pink?

What is the psychological difference between red and blue?

What is the psychological difference between red and black?

Let's return to the armed bank robbery case. Please use the same scale you just practiced for to answer the following question.

How different is Judge Walters' decision about Convict X from your own view? Imagine that 100 is moderately different from your own view. If you think the difference between your view and Judge Walters' sentence is twice as much as a moderate difference, rate it 200. If you think the difference between your view and Judge Walters' view is half as much as a moderate difference, rate it 50. If Judge Walters' sentence is not at all different from your view, rate it 0 (zero). You may use any number between 0 and 100 for views that are less than moderately different from your own, and you may use any number greater than 100 for views that are more than moderately different from your own. You may use any number you wish, from zero on up. Please enter a number that represents the difference between Judge Walters' decision and yours.

When a jury hears a case, the members typically have their initial views. Then, after they have heard additional information, they give their final views. In this study, you expressed an initial opinion, but since then, you have received additional information. Therefore, please feel free to change your views. To how many years in prison do you think Convict X should have been sentenced?

We would like to know more about how you evaluate Judge Walters' decision. Before you answer more questions, here is another practice task for you: Imagine that there is a scale to measure the hostility between two countries. A value of 100 represents a moderate level of hostility. Zero represents “No hostility at all.” There is no upper limit. Please answer the following questions.

What is the hostility level between the U.S. and Canada?

What is the hostility level between the U.S. and India?

What is the hostility level between the U.S. and Russia?

Let's return again to the armed-robbery case, please answer the following questions using the scale you just practiced for.

How surprising did you find the sentence Judge Walters imposed on Convict X? If you found the sentence not at all surprising, write 0 (zero). If you found the sentence moderately surprising, write 100. If you found the sentence twice as surprising as moderately surprising, rate it 200. If you found the sentence half as surprising as moderately surprising, rate it 50. You may use any number between 0 and 100 to indicate a level of surprise that is less than moderate and any number greater than 100 to indicate a level of surprise that is greater than moderate. You may use any number you wish, from zero on up. Please enter a number in the box below.

Please use the same scale to answer the following questions.

How unbiased was Judge Walters? (Imagine that 100 is moderately unbiased.)

How credible was Judge Walters? (Imagine that 100 is moderately credible.)

How important was Judge Walters' sentence when you decided what an appropriate sentence should have been for Convict X? (Imagine that 100 is moderately important.)

How effectively did Judge Walters' statement make its point? (Imagine that 100 is moderately effective.)

How well written was Judge Walters' statement? (Imagine that 100 is moderately well written.)

How compelling was Judge Walters' statement? (Imagine that 100 is moderately compelling.)

How convincing was Judge Walters' statement? (Imagine that 100 is moderately convincing.)

How fair was Judge Walters' decision? (Imagine that 100 is moderately fair.)

How unexpected did you find the sentence Judge Walters imposed on Convict X? (Imagine that 100 is moderately unexpected.)

How expert was Judge Walters? (Imagine that 100 is moderately expert.)

How trustworthy was Judge Walters? (Imagine that 100 is moderately trustworthy.)

How harsh was the sentence Judge Walters imposed on Convict X? (Imagine that 100 is moderately harsh).

How punitive was the sentence Judge Walters imposed on Convict X? (Imagine that 100 is moderately punitive).

Think about your own view about the appropriate years of sentence for armed robbery after you read the Sentencing Guideline but before reading Judge Walters' decision. How different in years was Judge Walters' sentence of Convict X from your own view? You have answered a similar question a moment ago with the scale where 100 represented a moderate difference. For the current question, please enter a positive number that represents a difference in years:

You are almost done. Thank you for the patience and effort. In this last section of the survey, we would like to know more about you.

What is your gender? Male, female, or other.

Are you Spanish, Hispanic, or Latinx? Yes or no.

Here is a list of race categories. Please choose one or more races that you consider yourself to be. White, Black or African-American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, or other.

What is the highest level of school you have completed or the highest degree you have received? Twelfth grade (no diploma) or less, high school graduate with high school diploma or equivalent (for example: GED), some college but no degree, associate degree, bachelor's degree (for example: BA, AB, BS), master's degree (for example: MA, MS, MEng, MEd, MSW, MBA, MLS), professional school degree (for example: MD, DDS, DVM, LLB, JD), doctorate degree (for example: PhD, EdD), or other (please specify).

Which of the following best describes your political party identification? Strong Republican, Republican, Independent and leaning toward Republican, Independent, Independent and leaning toward Democrat, Democrat, strong Democrat, no preference, or other (please specify).

In general, where do you place yourself on a liberal (0) to conservative (100) scale? Please drag the slider below to indicate your political ideology. You need to click on or move the slider bar at least slightly for the question to count as answered.

Please use a number to indicate how much you agree with each of the following statements. You should use a scale where 100 indicates that you moderately agree with a statement, zero indicates that you do not agree with a statement at all, and the scale has no upper bound. You may use any number you wish, from zero on up.

Religion is a very important part of my life.

I would describe myself as a religious person.

My religious beliefs influence how I make my decisions.

Prayer is a very important part of my life.

We would like to know more about how you react to new information in general. Please use a number to indicate how much you agree with each of the following statements. You should use a scale where 100 indicates that you moderately agree with a statement, zero indicates that you do not agree with a statement at all, and the scale has no upper bound. You may use any number you wish, from zero on up.

My ideas are very stable and remain the same over time.

It is hard for me to change my ideas.

I have never changed the way I see most things.

After forming an impression of something, it's often hard for me to modify that impression.

I have often changed my opinions.

If it is necessary I can easily alter my beliefs.

My opinions fluctuate a lot.

I often vary or alter my views when I discover new information.

It could be said that I am likely to shift my attitudes.

I find my opinions to be changeable.

Have you previously heard about this study? Yes or no.

Have you previously participated in this study? Yes or no.

In your own words, please describe what you think the purpose of this survey is.

You have reached the end of the survey. By selecting "Submit" and hitting the "Next" button, you will submit your responses. Thank you for your time and effort. Now that you have finished the survey, the research team is responsible for explaining the purpose of this study. Please click and review the debriefing form. When you finish reviewing the debriefing form, please select one of the options below: I have been debriefed by the research team, and I understand the intent of and the purpose of my participation in the study. I agree that my data collected during the study can be included for the purpose of the study; I have been debriefed by the research team, and I understand the intent of and the purpose of my participation in the study. I do not give permission for my data collected during the study to be included for the purpose of the study.

APPENDIX G

A DEMONSTRATION OF SUBJECT EQUIVALENCE BETWEEN PILOT STUDY 5 AND PILOT STUDY 6

This appendix shows that the two sets of data (one from the fifth pilot study with $n = 75$ and the other one from the sixth pilot study with $n = 75$) did not differ significantly on the variables that were not supposed to be affected by the message-scale-value manipulation. For all the relevant variables below, data winsorization (at the 95 percentile) and data transformation were performed when necessary as stated in Chapter 6.

The square root of the winsorized survey completion time did not differ significantly between the two datasets, $t(148) = -1.31$, $p = .19$. The square root of the winsorized subject's age did not differ significantly between the two datasets, $t(148) = -1.67$, $p = .10$. The square root of the winsorized initial position did not differ significantly between the two datasets, $t(148) = -0.47$, $p = .64$. The winsorized subject's upper bound did not differ significantly between the two datasets, $t(148) = -0.52$, $p = .60$. Lastly, the square root of the winsorized subject's lower bound did not differ significantly between the two datasets, $t(148) = 0.74$, $p = .46$.