4-18-2012 12:00 AM

The Role of Valence in Construct Dimensionality Debates

Chun Seng Kam, *The University of Western Ontario*

Supervisor: Dr. John P. Meyer, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Psychology

THE ROLE OF VALENCE IN CONSTRUCT DIMENSIONALITY DEBATES

(Spine title: The Role of Valence in Construct Dimensionality Debates)

(Thesis format: Monograph)

by

Chun Seng Kam

Graduate Program in Psychology

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor

Dr. John Meyer

Supervisory Committee

Dr. Sampo Paunonen

Dr. Richard Goffin

Examiners

Dr. Richard Goffin

Dr. Mitch Rothstein

Dr. Heather Laschinger

Dr. Michael Ashton

The thesis by

Chun Seng Kam

entitled:

The Role of Valence in Construct Dimensionality Debates

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Date

Chair of the Thesis Examination Board

ABSTRACT

There continues to be debate about the dimensionality of important psychological constructs (e.g., anxiety; job satisfaction). The standard procedures for determining whether a construct is unidimensional or two-dimensional have been to (a) conduct factor analyses and (b) test for differences in the nomological network of correlations with other variables. I argue, and demonstrate empirically, that these approaches can sometimes mislead researchers to draw incorrect conclusions. In Study 1, I examined how item valence (i.e., favorability of item content) can affect factor analyses and nomological network analyses with two separate samples of undergraduate students. Results consistently showed that item valence can induce regular- and reverse-keyed items to load on separate factors in factor analyses, and that it can systematically bias construct correlations in favor of the two-dimensional interpretation of a construct. Multi-trait multi-method analyses demonstrated that the item valence accounts for close to 10% of the item variance, and this effect cannot be entirely explained by social desirability response bias. In Study 2, I again tested the effect of item valence in a job satisfaction measure. In addition, I examined careless responding as an alternative explanation to the item valence effect among working adults. Results not only replicated item valence effect in Study 1, but it also showed that careless responding can amplify, but cannot explain entirely, the apparent two dimensionality results caused by valence bias in both factor analytic and nomological network analyses. I suggested several remedies for the valence problem, including the minimization of valence in item design, the use of reverse-keyed items for construct measurement, and the use of objective behavioral measures in nomological network investigations.

KEYWORDS: valence, construct dimensionality, factor analysis, nomological network.

ACKNOWLEDGMENTS

A number of people have contributed significantly to this project. First and foremost, I have to thank my advisor, Professor John P. Meyer, for his tremendous guidance, brilliant insight, and continuous encouragement. He has provided me with rigorous feedback on the theoretical component of this work. Through his years of guidance he has taught me how to have the utmost professionalism in being an academic researcher. Without exaggeration, he is a mentor with a great sense of caring and kindness; he went beyond the role of an academic supervisor and contributed to my overall development as a researcher and as a person.

I would like to extend my gratitude to my committee members –Professor Sampo V. Paunonen and Professor Richard D. Goffin. Both of these mentors have provided me with valuable knowledge regarding the methodological components of this work. I especially value the thought-provoking and valuable comments of Prof. Paunonen and the constructive and meticulous feedback from Prof. Goffin. Both of these mentors showed genuine care for my work and my career.

I am grateful to Professor Richard Neufeld, Dr. Thomas O'Neill and Dr. Etienne LeBel who have provided me with invaluable advice for my work. I also must extend my gratitude to my Master's advisor, Professor Michael Harris Bond, who has provided continuous guidance on my Ph.D. and my continuing career path. I would like to thank Professor Andreas Klein, who encouraged me to pursue my career in academia. Finally, I would like to thank my father and my family members, who provided me with love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1 - INTRODUCTION

The accurate measurement of psychological constructs is essential for research and the advancement of knowledge. A key to effective measurement is a clear conceptualization of the construct. However, we often rely on measurement research to help formulate our ideas about the true nature of a construct. Thus, theory and measurement are inextricably intertwined (Cronbach & Meehl, 1955). This link is clearly reflected in ongoing debates regarding the dimensionality of several important constructs across multiple areas of psychology, including presence of anxiety versus absence of anxiety in clinical psychology (Vagg, Spielberger, & O'Hearn, 1980; Vigneau & Cormier, 2008), optimism versus pessimism in personality and clinical psychology (e.g., Scheier, Carver, & Bridges, 1994; Rauch, Schweizer & Moosbrugger, 2007), marital satisfaction versus dissatisfaction in family psychology (Locke & Wallace, 1959), belief in the just world versus belief in the unjust world in social psychology (Rubin & Peplau, 1975), and job satisfaction versus dissatisfaction in industrial-organizational psychology (e.g., Credé, Chernyshenko, Bagraim, & Sully, 2009). To illustrate, consider the case of job satisfaction. Some scholars argue that satisfaction is a unidimensional construct, anchored at one pole by extreme dissatisfaction and at the other pole by extreme satisfaction (see Locke, 1976). In contrast, others believe that satisfaction and dissatisfaction are distinguishable, albeit related, constructs (Credé et al., 2009). These diverging beliefs can have important implications for research pertaining to the development and consequences of job satisfaction/dissatisfaction as well as for management practices. For example, should we be trying to identify common or distinct mechanisms underlying employees' satisfaction and dissatisfaction with their jobs? Is the management of these evaluations a unitary or two-step process?

To address the issue of construct dimensionality, researchers include items measuring both poles of a construct and conduct factor analyses. If these analyses yield two factors defined by items reflecting the opposing poles, it is taken as evidence that the poles actually reflect distinguishable constructs. To provide further evidence, researchers often compute separate scores and correlate them with other variables (e.g., theoretical antecedents and consequences). If the pattern of correlations differs, it adds support to the argument for distinguishable constructs (see Credé et al., 2009, for a recent example involving job satisfaction and dissatisfaction).

My objective in the present research is to offer an alternative explanation for the variance contributing to the appearance of a second factor in measures of bipolar constructs. Specifically, I argue that, in addition to variance due to content, there will be variance in a construct measure due to its *valence,* or favorability, as an attribute (e.g., marital satisfaction is more favorable than is marital dissatisfaction; emotional stability is more favorable than is neuroticism). Past methodological research has suggested that the emergence of two factors for a given measure may simply be the result of unexpected variance due to the differential keying direction of scale items (e.g., Horan, DiStefano, & Motl, 2003; Magazine, Williams & Williams, 1996; Marsh, 1996; Motl & DeStefano, 2002; Tomás & Oliver, 1999). As the first goal, this dissertation aims to advance previous research by clarifying the distinction between two concepts that are often confused: *item keying direction* vis-à-vis *item valence*. Researchers often use these two terms interchangeably but in fact there are clear distinctions between them, and this confusion can hinder communications among researchers.

The second major objective of the current research is to demonstrate that item valence can add an additional source of variance in *both* factor analytic *and* nomological network analyses. Although past research has shown that item-keying direction can cause an artificially

distinct factor in factor analyses (e.g., Horan et al., 2003; Magazine et al., 1996), investigators continued to search for evidence of the bi-dimensionality of a construct with nomological network analyses (e.g., Marshall, Wortman, Kusulas, Hervig, & Vickers, 1992). Presumably the underlying assumption among these researchers is that nomological network analysis is able to provide information *distinct* from factor analysis, and thus both types of analyses corroborate each other to reveal the true picture regarding construct dimensionality. However, with two empirical demonstrations, I will show that variance due to item valence can distort the results of both factor analyses and nomological network analyses. Specifically, when item valence causes a bias in factor analysis in favor of a two-dimensional explanation, it will likely produce the *same* bias in nomological network analysis.

In the following discussion I will first differentiate between two important concepts which researchers often confuse. This confusion may have caused researchers to neglect possible bias in the statistical outcomes they observed. Next, I will elaborate on the debates pertaining to construct dimensionality with illustrative examples. The discussion focuses on dimensionality debates in which opposite poles of a construct (e.g., presence versus absence of anxiety) are suspected to belong to separate dimensions. Further, I provide a more detailed discussion of construct valence and its implications for the dimensionality debate. Finally, I explain my analysis strategies.

Distinction between Constructs versus Measures

Researchers often confuse two distinct concepts - constructs versus measures. A construct (or, hypothetical construct) is an ideal, unobserved concept whose existence is based in a person's mind (MacCorquodale & Meehl, 1948). An example of a construct is job satisfaction. Given that job satisfaction exists in respondents' minds, there is impossible to directly observe or

objectively measure a person's job satisfaction. Similar examples in psychology are work engagement, anxiety, motivation, personality, organizational self-esteem and most of the concepts that psychologists are interested in. There are also examples outside psychology such as liberalism (in political science), gravity (in physics), romanticism (in literature), and evolution (in biology). None of these constructs are directly and objectively quantifiable. For example, it is impossible to put an objective number on the degree of romanticism in a particular piece of literature writing.

Researchers may devise an instrument to measure a person's job satisfaction indirectly. The instrument may indicate that a person's satisfaction score is five (out of a total of seven). However, a hypothetical construct at the latent level is not in the same metric as a survey measure. For example, job satisfaction at the latent level is obviously not in a Likert scale format as in an observable measure. Therefore, participants need to translate his or her own latent level of job satisfaction onto a paper-and-pencil Likert-scale survey. This translation process is subject to response bias, misunderstanding of the survey scale, misinterpretation of the survey items, inability to self-reflect, and other unintended errors. Therefore, an observed score is simply a proxy, rather than a perfect representation, of the degree of latent constructs.

Despite the differences between hypothetical (latent) constructs and observable measures, researchers often treat a measure score as if it were an accurate reflection of the latent construct. For example, they often depend on the observed correlations among the measures, and use them to infer the association among the latent variables. Unfortunately, findings based on the observed level do not necessarily reflect the state of affairs at the latent level. In most cases a researcher is interested in understanding the latent construct itself, rather than its corresponding observed measure. Perhaps the only exception is when a researcher simply aims at validating a

measurement instrument of a common construct, in which its nomological network has been already established. In the following I will focus mainly on the role of measurement methods in causing the incommensurability between a latent construct and an observed measure.

For most psychological variables, an observed score is a function of true score, methods score, and unreliability (Campbell & Fiske; 1959):

$$\text{Observed Score} = \text{True Score} + \text{Method Score} + \text{Random Error}$$

Method score here refers to systematic effects on a score due to measurement methods. When we assume a simple additive relationship between the observed score and its components, and when we further assume no multiplicative interaction among these components, the variance of the observed score is as follow:

$$\text{Var}_{\text{observed}} = \text{Var}_{\text{true}} + \text{Var}_{\text{method}} + \text{Var}_{\text{random error}}$$

Variance of an observed score can be decomposed into variance of the true score, variance of the method score, and variance of the random error. From the two formulas shown above, we can clearly see that two individuals may obtain different observed scores simply because of different method scores or differential errors.

Method effect refers to a systematic effect due to measurement methods (e.g., self-report versus peer report), which might include response bias. Response bias refers to "a systematic tendency to respond to a range of questionnaire items on some basis other than the specific item content (i.e., what the items were designed to measure)" (Paulhus, 1991, p. 17). Certain methods, such as participants' self-report of job performance, allow for response bias, while other methods, such as objective indicators of job performance (e.g., sales figures), do not. As an example, when a survey presents its items in a Likert-scale format (method), systematic response bias occurs for participants who tend to disagree rather than agree with items regardless of item content. Later in

this dissertation, I will discuss two common types of response bias – social desirability responding (Paulhus, 1991) and careless responding (Schmitt & Stuits, 1985).

Construct Dimensionality Debates

To illustrate how response bias can influence a construct dimensionality debate, I will first discuss how researchers usually develop their surveys. Hinkin (1995) has reviewed the key stages in measure development. The first stage is the creation of items. Researchers may follow a deductive approach (i.e., reviewing the relevant literature and developing items rationally based on construct definition) or an inductive approach (i.e., developing items based on respondents' descriptions of their construct-relevant behaviors, cognition, or feelings). Researchers may include or exclude reverse-keyed items at this item creation stage. Subsequently, researchers administer the measure to the population of interest and conduct reliability analyses. Afterwards, researchers are usually interested in assessing the dimensionality of their measure. For instance, if they theorize their measure is unidimensional, at this stage they will check whether the data supported that expectation. The standard procedure for assessing a measure's dimensionality is factor analysis (e.g., Carmines & Zeller, 1979; Scheier & Carver, 1985) followed, where appropriate, by nomological network analysis (see below; Cronbach & Meehl, 1955). Due to the importance of these two procedures in the assessment of a construct's dimensionality I will describe them in further detail below.

Factor Analysis and Nomological Network Analysis

As noted above, a first step in assessing the dimensionality of a construct or measure often involves the use of factor analysis. When the factor analysis of items on a measure yields two factors, it suggests that the items defining the factors share unique variance (although they may also share common variance with items defining the other factor if the factors are allowed to

correlate). In studies where the focus is on the dimensionality of one or more *measures*, the issue often has to do with the inclusion of reverse-keyed items. Reverse-keyed items are written in such a way that disagreement rather than agreement reflects a higher level of the construct. Scores on reverse-keyed items are reflected (e.g., 5 = 1; 1 = 5) before computing scale scores. When reverse-keyed items are included in a measure, factor analyses often yield two factors, one defined by regular items and one by the reverse-keyed items (e.g., Dalbert, Montada, & Schmitt, 1987). Some investigators interpret this finding as evidence that the focal construct is really two-dimensional (e.g., presence versus absence of anxiety; job satisfaction versus dissatisfaction) and that the two dimensions should be measured and studied independently (e.g., Dalbert, Lipkus, Sallay, & Goch, 2001; Fincham & Lindfield, 1997). For example, Vagg et al. (1980) performed an exploratory factor analysis and concluded that presence and absence of anxiety represent distinct factors. Very recently, Credé et al. (2009) also conducted confirmatory factor analyses of two different job satisfaction measures and discovered that a two-factor solution provided a better fit to the data than did a single-factor solution in both cases. Credé et al. interpreted their findings as evidence that job satisfaction and job dissatisfaction are distinguishable constructs.

Some investigators have argued that the two factors found in studies such as those cited above are often a function of response style related to direction of item keying, rather than a reflection of the inherent nature of the construct(s). To demonstrate that item-keying is indeed a biasing factor, investigators (e.g., Carmines & Zeller, 1979; Greenberger, Chen, Dmitrieva, & Farruggia, 2003; Vautier & Pohl, 2009) have conducted confirmatory factor analyses (CFA) and included a keying-factor(s) along with a substantive factor in their models. In such a model, the substantive factor has causal paths directed at all of the items, and the keying factor(s) has/have causal paths directed at the regular-keyed items, the reverse-keyed items, or both. This model

typically produces a better fit to the data than does a model without item keying factor(s), suggesting that the use of item keying introduces construct irrelevant variance. For example, Rauch et al. (2007) included an item-keying factor along with a substantive factor in an analysis of the revised version of Life-Orientation Test and found that the keying factor accounted for a substantial portion of the variance. Similar analyses have been conducted in studies including multi-dimensional construct measures such as employee commitment (e.g., Magazine et al., 1996), for which there are multiple content factors. In this case, the keying factor had causal paths directed at reverse-keyed items on each dimension. Again, this model fit better than one with only substantive factors, suggesting that keying introduces construct-irrelevant variance that can influence the correlations among the construct measures.

When evidence for two factors is obtained, a second step in the investigation of construct dimensionality often involves a nomological network analysis (Cronbach & Meehl, 1955) in which measures of the two factors are correlated with other variables presumed to be antecedents, correlates, or consequences of the underlying constructs. If this analysis demonstrates a different pattern of relations, it adds credence to the conclusion that the factors reflect distinguishable constructs. In contrast, if the two measures simply reflect polar opposites of a unidimensional construct they would correlate with the same magnitude but in opposite directions. As an illustrative example, Credé et al. (2009) proceeded with correlational analyses after finding evidence of a two-factor solution for a job satisfaction measure. These researchers discovered that job satisfaction and job dissatisfaction showed distinguishable patterns of correlations with other variables. Job satisfaction correlated more strongly with positive affect and organizational citizenship behaviors; job dissatisfaction correlated more strongly with negative affect, counterproductive work behaviors, job stress, and perceived discrimination. Credé et al. (2009)

therefore concluded that job satisfaction and job dissatisfaction were two separate constructs. Likewise, although currently recognized as a unidimensional construct (e.g., Greenberger et al., 2003; Marsh, 1996; Quilty, Oakman, & Risko, 2006), self-esteem was once argued by some researchers as being two-dimensional, with positive and negative self-esteem having separate nomological networks. Compared with positive self-esteem, negative self-esteem was found to correlate more strongly with other negative constructs such as depression and perceived discrimination (Owens, 1994; Verkuyten, 2003).

In the present study, I examined the sources of variance inherent in measures of two constructs: self-esteem (positive vs. negative) in Rosenberg (1965) Self-Esteem Scale (RSES) and extraversion (extraversion vs. introversion) in International Personality Item Pool (IPIP; Goldberg et al., 2006). Both of these existing measures included a balanced number of regular- and reverse-keyed items. In both cases, the inclusion of reverse-keyed items may introduce construct-irrelevant bias that will have implications for factor analysis and nomological network analysis. To test the magnitude of the effect of item-keying direction, I used a CFA-based analytic strategy. I provide a more detailed description of my analysis below. First, let me turn our attention to an alternative source of variance that might help to explain the factor structure of, and correlations between, (potentially) bipolar constructs.

Valence as an Additional Source of Variance

Although there may be exceptions (e.g., monochronicity versus polychronicity [Hall, 1983]; agency versus communion [Baken, 1966]; individualism versus collectivism [Hofstede, 1980]), it is often the case that the behaviors and attributes defined at one end of a bipolar construct are more favorable than the attributes at the other. That is certainly the case in the constructs under investigation in the present study. Positive self-esteem (measured by items such

as “I feel that I’m a person of worth” in RSES) is more favorable than negative self-esteem (measured by items such as “At times I think I am no good at all” in RSES). Similarly, extraversion (captured by items such as “I make friends easily” in IPIP) is generally seen as more favorable than introversion (captured by items such as “I have little to say” in IPIP). I use the term valence to refer to this characteristic of an attribute and describe favorable attributes, behaviors, or beliefs as positively valenced and less favorable attributes as negatively valenced.

When making judgments about ourselves or another individual, often we are not only giving a *description* of the target but also providing a subjective *evaluation* of the target. Researchers in the areas of personality psychology and industrial psychology are particularly interested in the composition of observers’ rating (e.g., Connelly & Ones, 2010; Funder, 1995; Oh, Wang, & Mount, 2011; Paunonen & Ashton, 2001). Because a target’s personality traits cannot be directly observed, researchers often rely on raters (e.g., a target’s friends) to report their judgments on a measurement instrument (e.g., survey, questionnaire). Although originally proposed by Dean Peabody (1967), Beauvois and Dubois (2000) recognized that observers’ ratings of a target typically are comprised of both descriptive aspects and evaluative aspects. Descriptive aspects refer to the target’s actual behaviors. Some examples are the frequency with which the target forgives other people quickly or the extent to which the target treats people with kindness. Evaluative aspects refer to the meaning or value of the target’s behaviors to the observer (i.e., what do the target’s behaviors mean to me?; Beauvois & Dubois, 2000; read also Hofstee, 1990). Examples are whether the observer likes the target’s forgiving nature, or whether the observer would like to spend time with the target because of his/her kind nature. Descriptive aspects were traditionally the focus of psychologists, and the evaluative aspects of a target did

not gain much research attention until much later (e.g., Beauvois & Dubois, 2000; Saucier, 1994; Tellegen, 1993).

This idea of descriptive and evaluative judgments in observer ratings can also be applied to self-ratings. That is, self-rating can also be comprised of both aspects of description and evaluation of oneself (e.g., Peabody, 1967). When applied to self-rating, descriptive judgments refer to the rater's actual behaviors. However, evaluative judgments refer to the meaning of these behaviors to the raters themselves. Therefore, evaluative judgment may be affected by the tendency to exaggerate one's positive qualities and to underestimate one's negative qualities.

According to Peabody (1967), the evaluative aspects of a target are naturally confounded with the descriptive aspects of a target. For example, when we report our agreement to the statements "I make others feel at ease" or "I insult others" (both items measure opposite ends of the agreeableness trait), we are not only describing our behaviors but also evaluating ourselves as a worthy or an unworthy individual in a social setting. Because both evaluative and descriptive aspects are inherent in most scale items, it is extremely difficult, if not impossible, to disentangle the two aspects. There have been attempts to minimize the evaluative aspect of personality items (e.g., Jackson, 1984; Peabody, 1967), but Bäckstrom et al. (2009) have shown empirically that evaluative content cannot be entirely eliminated, at least among items measuring Big Five personality traits. In the current article, item or construct valence refers to this evaluative aspect in a scale item or a construct.

Unconfounding construct valence with item keying direction

It is important to note that the evaluative aspect (i.e., valence) of items in a measure is distinct from their direction of keying. For example, if Emotional Stability is used to describe one of the Big Five personality traits, the reverse-keyed items measure Neuroticism – a

negatively valenced construct. However, if the scale is reflected to measure Neuroticism, then the reverse-keyed items measure the positively valenced Emotional Stability construct (see Table 1). This distinction has been blurred (or confused) in the past by authors who refer to regular-keyed items as "positively keyed" and reverse-keyed items as "negatively keyed" (e.g., Vigneau & Cormier, 2008). More often than not, the negatively-keyed items being described indeed have negative valence (i.e., we tend to put positive labels on our constructs), although it is not always the case (e.g., Credé et al., 2009). However, the term reverse-keyed is more variable and allows for the possibility that such items can actually have a positive valence when the construct under investigation is negatively valenced (e.g., neuroticism; belief in a dangerous world).

I propose that the valence of a construct can introduce an independent source of variance to a measure beyond that explained by the construct itself. There may be unique variance attributable to both the regular- and reverse-keyed items due to the fact that they have different valences. Therefore, in the current analyses I included factor(s) to reflect positive and negative valence along with each substantive construct factor.

I should note that the search for item-keying effects is not unprecedented. Some previous researchers have included what is in essence a reverse-keyed method factor in analyses involving constructs at the center of a dimensionality debate. Some example of these construct measures include State-Trait Anxiety Inventory (Vautier, Callahan, Moncany, & Sztulman, 2004), Meyer and Allen's (1984) measure of organizational commitment scale (Magazine, Williams & Williams, 1996), Life-Orientation Test Revised (measuring optimism; Rauch et al., 2007). In those studies, a model with the method effect controlled fit better than did a baseline model. Most of those studies, however, attributed the emergence of the method factor simply to the inclusion of reverse-keyed items. As I have already argued, the inclusion of reverse-keyed items

Table 1

Relations Between Item Keying Direction and Valence

	Regular-Keyed	Reverse-Keyed
Neuroticism	Negatively-valenced item Example: “I often feel blue.”	Positively-valenced item Example: “I am very pleased with myself”
Emotional Stability	Positively-valenced item Example: “I am very pleased with myself”	Negatively-valenced item Example: “I often feel blue.”

and the valence of items are naturally confounded with each other (Peabody, 1967). Therefore, the present study will examine which of these two effects most likely contributed to the discovery of the method effects in those previous studies.

Past research has often treated factor analysis and nomological network analyses as two independent investigations. Conclusions regarding a construct's dimensionality are stronger when evidence comes from both types of analysis than from any one type alone. More often than not, the conclusions from both analyses agree with each other: separate factors discovered in factor analysis also correlate differently with other variables in nomological network analyses (e.g., Fincham & Linfield, 1997), thus supporting researchers' expectations regarding a construct's bi-dimensionality. The current paper, however, questions this conventional perception with regard to the independence between the two types of analysis. My major argument is that factor analyses and nomological network analyses are *both* liable to misleading interpretations caused by the valence effect.

For nomological network investigations, when a construct is measured predominantly or exclusively by regular-keyed items, the construct score itself will be partially accounted for by a particular valence (positive or negative). The valence issue has implications for both sides of the construct dimensionality debate. For those who advocate the two-dimensional interpretation of a construct, its regular-keyed items and its reverse-keyed items will have opposite valence. If the valence is positive, it is possible that its correlations with other positively-valenced constructs will be inflated and its correlations with negatively-valenced construct will be deflated. Similarly, if the valence is negative, it is possible that its correlations with other negatively-valenced constructs will be inflated and its correlations with positively-valenced constructs will be

deflated. This will contribute to the finding that the two constructs have different nomological networks, thereby justifying their continued treatment as separate constructs.

As an illustrative example, a well-cited study by Marshall et al. (1992) demonstrated that optimism and pessimism have separate nomological networks – optimism correlates more strongly with extraversion than does pessimism; pessimism correlates more strongly with neuroticism than does optimism. Those researchers thus conclude that optimism and pessimism should be treated as distinguishable constructs. Some other investigators (Herzberg, Glaesmer, & Hoyer, 2006) concurred with this conclusion because they also found that pessimism correlates more strongly with depression than does optimism. Careful examination of those studies, however, reveals that the constructs in these nomological network studies (e.g., extraversion, neuroticism) were measured predominantly by regular-keyed items. Therefore, it is unclear whether these results were caused by item content, item valence, or both. Optimism might correlate stronger with extraversion simply because both constructs were measured by positively-valenced items. Similarly, pessimism might correlate stronger with neuroticism and depression only because these three constructs were all measured by negatively-valenced items. This methodological flaw renders any substantive interpretations of these results problematic.

Two Potential Explanations for the Valence Effect

Two potential explanations for the valence effect will be reviewed here. The first explanation, social desirability response bias, is quite obvious. Social desirability response bias reflects participants' tendency to project themselves in a positive light by showing a socially approved image of themselves. Recently, Bäckström et al. (2009) found that social desirability response bias can explain common variance among the Big Five factors that are theoretically expected to be orthogonal. Bäckström et al.'s research thus suggests that participants' responses

to the survey items are influenced by their desire to project a positive self-image and mitigate a negative self-image. The result is that respondents' answers can reflect the effect of item valence. As discussed later, Study 1 of the present research included a measure of social desirability response bias to examine its role on participants' self-ratings.

I deliberately chose the term *valence* in contradiction to past research in personality, which often uses the term *social desirability* to represent the evaluative aspect of an item (e.g., Peabody, 1967). I did this because the term *social desirability effect* traditionally has an inherently negative connotation of faking or distorting survey responses among personality and industrial psychologists. For measurement purposes, the traditional goal of many researchers is to minimize the negative effect of social desirability response bias (e.g., Bäckström et al., 2009). In contrast, the term *valence effect* has a more neutral stance that entertains a wide range of potential explanations for an item response pattern. As I will demonstrate later, item valence cannot be fully explained by participants' ego-enhancing manner (i.e., social desirability response bias) in my investigations. In addition, social desirability response bias is relevant in situations where a respondent is evaluating a characteristic about himself or herself. However, there are situations where respondents may not be influenced solely by social desirability responding. For example, when one is rating his or her own private beliefs or attitudes (e.g., attitude towards the organization), item valence may affect responses that social desirability responding does not. Therefore, I have employed a more general term *valence*.

The second explanation for the valence effect is careless responding. There are two types of careless responding (Meade & Craig, in press). The first type is random responding — respondents randomly choose a response option for each survey question (Meade & Craig, in press). Assuming that each response option has an equal probability of being selected for each

survey item and that the percentage of respondents in this type of careless responding is small (as compared to careful respondents), it should not substantially affect the underlying factor structure of the items. In contrast, the second type of careless responding is non-random responding, and it is more problematic. Schmitt and Stuits (1985) proposed one kind of non-random responding, in which respondents may read the first few items of the survey, decide on the general content of all the questions, and then give the same answer for each survey item (e.g., a rating of "5" on a 5-point Likert scale for both regular-keyed and reverse-keyed items). However, because answers to reverse-keyed items are reversed before scoring, a score of "5" for a reverse-keyed item will be converted into a score of "1". This will result in non-correspondence between the scores of regular-keyed items ("5") and the scores of reverse-keyed items ("1"). Thus, because regular-keyed items and reverse-keyed items often differ in their valence, this type of careless responding will result in positively- and negatively-valenced items differing in their scores and correlating weakly with one another. Hence, the overall result of this responding style is that a unidimensional construct will form two factors in factor analysis.

According to the results of a simulation study by Schmitt and Stuits (1985), if as little as 10% of the participants give identical response alternative across all measurement items, a factor represented by negatively-keyed items will appear. However, that study (Schmitt & Stuits, 1985) was conducted with artificial simulated data rather than with actual data. Therefore, when real respondents show a consistent pattern of endorsing both regular- and reverse-keyed items, little is known about whether it simply reflects careless responding or is the participants' genuine opinion. As discussed later, Study 2 in this research included a measure of careless responding that was used with real participants. This allowed me to distinguish the effect of careful vis-à-vis careless respondents.

Objectives of the Current Dissertation

My dissertation has four major objectives that will be pursued in two studies. Objective 1 is to test whether item valence can cause a unidimensional construct to be two-dimensional in *factor analytic* procedures. If correlations among items reflect both their content and their valence, then I would expect similarly-valenced items to have stronger correlations with each other than oppositely-valenced items. This finding will suggest that item valence is enough to cause oppositely-valenced items to load on two separate factors in factor analysis, even when they all measure similar item content. Objective 2 is to test whether valence can distort correlational results in *nomological network* studies. Particularly, I expect that when two constructs are measured by similarly-valenced items, their correlations will be inflated; when two constructs are measured by oppositely-valenced items, their correlations will be deflated. Objectives 3 and 4 are to test whether the valence effect can be fully explained by social desirability responding and careless responding, respectively (see Table 2). To examine the generalizability of my results, Study 1 will have a sample of undergraduate students completing a variety of measures commonly found in personality and social psychology, and Study 2 will have working adult samples completing measures in industrial-organizational psychology.

In Study 1 I will examine the effect of valence on using the measurement of extraversion-introversion and positive and negative self-esteem. Extraversion and self-esteem were chosen because they are two of the most common constructs in multiple areas of psychology. In Study 2 I will illustrate the effect of valence with the measurement of job satisfaction, because job satisfaction is a popular construct in industrial-organizational psychology. Some organizational behavior researchers (Credé et al., 2009) have shown that job satisfaction is a two-dimensional construct that is represented by positively-valenced items on one dimension (job satisfaction) and

Table 2

Summary of the Objectives of the Current Research

	Study 1	Study 2
Objective 1: Evaluate item valence and its effects on factor analyses	√	√
Objective 2: Evaluate construct valence and its effects on nomological networks	√	√
Objective 3: Evaluate social desirability response bias as an explanation of item valence effect	√	
Objective 4: Evaluate careless responding as an explanation of item valence effect		√

negatively-valenced items on the opposite dimension (job dissatisfaction). In addition, research by Credé et al. (2009) suggested that these two dimensions have distinct nomological networks. I will test how item valence may affect the factor analytic and nomological network results for this construct.

CHAPTER 2 — STUDY 1

The overall objective of Study 1 is to demonstrate the potential impact of item and construct valence using the measurement of extraversion and introversion and the measurement of positive and negative self-esteem as examples. My intention here is to investigate the implications of valence as it pertains to measure development and substantive research in general. My first objective is to demonstrate the impact of valence on exploratory factor analysis of extraversion-introversion and positive-negative self-esteem. This question is important because exploratory factor analysis remains a common method for organizational researchers to examine the dimensionality of a construct.

My second objective is to examine the impact of construct valence on correlations among variables (e.g., as in nomological network analysis). To do this, I created separate positive and negative self-esteem scores from a self-esteem measure and correlated them with other variables. These other variables were measured with positively- and negatively-valenced items, so I was able to create three scores: full-scale (all items included), positively-valenced, and negatively-valenced. If positive and negative self-esteem correlate similarly with the full-scale scores, but differ systematically in their correlations with the positively- and negatively-valenced scale scores of other variables, it suggests that valence is contributing to the correlations. Consequently, these correlations must be interpreted with caution when used to make decisions about whether a construct is unidimensional or two-dimensional. To examine the generalizability of my results, I created extraversion and introversion scores, and repeated the same analysis. Positive-negative self-esteem and extraversion-introversion were chosen as examples in my analyses because they are two of the most widely used scales in psychology. The purpose of my current research is to use both extraversion and self-esteem as examples to illustrate how valence

can bias both the factor analytic and nomological network results of psychological constructs in general.

As a follow-up analysis, I used the technique of multitrait-multimethod (MTMM) analyses on a wide variety of scales including Big Five traits and other psychological measures. The purpose was to estimate the magnitude of valence variance in psychological surveys. Although the current research is not the first to investigate the role of item keying direction in a construct dimensionality debate, past studies tended to focus narrowly on its role on a particular measure (e.g., Rosenberg's Self-Esteem Scale; Quilty et al., 2006; Tomás & Oliver, 1999). As such, their results cannot be generalized to a wide variety of measures in psychology. To overcome this limitation, I included several common *personality and social psychology measures* in my studies (e.g., Big Five traits, social dominance orientation, belief in zero-sum resources). Thus, my focus of examining the magnitude of valence was not only on constructs commonly used in a particular area of psychology (e.g,. personality psychology) in Study 1.

To assess the magnitude of valence, I compared a baseline model (in which all items load only on their construct factors) with a method-factor model (in which items load on both the construct factor and factor(s) reflecting positive and negative valence). If the second model fit better than the first model, it reflected the fact that item correlations reflect both their content and valence. The procedure also allowed me to calculate the variance explained by item content and item valence, thereby enabling me to estimate the magnitude of the valence effect vis-à-vis the content effect.

Some readers may question the relationship between the two valence factors, wondering whether positive valence is antipodal to negative valence. Traditionally, the evaluative content of personality ratings is assumed to be unidimensional, with positive and negative valence

representing two ends of the same pole (e.g., Peabody, 1967). However, empirical studies have shown that positive valence and negative valence load on two separate factors in MTMM analyses (e.g., Marsh, Scalas, & Nagengast, 2010; Tomás & Oliver, 1999; Vautier, Callahan, Moncany, & Sztulman, 2004; Vautier & Pohl, 2009; Vigneau & Cormier, 2008). Thus, I followed the practice of past research by having separate factors to represent positive valence and negative valence. However, I also re-investigated the dimensionality of valence in my data and imposed a one-factor solution of valence and compared its fit with the fit of a two-factor solution. Based on past empirical findings, I expect that positive and negative valence will belong to two distinct factors.

My third objective is to examine the effect of social desirability response bias. I examined how well social desirability response bias explained the variance in the valence factors in one of the samples in Study 1. Social desirability responding is comprised of two components — impression management and self-deception (Paulhus, 1991). Impression management refers to one's deliberate or intentional attempt to distort self-report in order to create a positive social image, and self-deception refers to one's non-deliberate or unintentional propensity to project an overly positive self-image (e.g., Li & Bagger, 2007). If the two components of social desirability responding do not correlate strongly with the valence factors (> .80), it means that valence factors are due to more than simply self- enhancement responding.

I expect that positively-valenced items and negatively-valenced items will load on two separate factors in exploratory factor analysis (for Objective 1). In addition, I expect that the valence effect can distort the results of a nomological network analysis in favour of a two-dimensional interpretation (for Objective 2). Furthermore, multitrait-multimethod (MTMM) analysis should show that the variance in scale items can be attributed to both constructs and

valence. However, I have no a priori expectation regarding how well social desirability responding will explain the valence effect (i.e., Objective 3).

Method

Participants

I collected data from two samples of introductory psychology students at a large Canadian university. These students were participating in online mass testing surveys conducted in consecutive years. Sample 1A consisted of 1094 students (332 male, 760 female and two unidentified; mean age = 18.45) and Sample 1B consisted of 1254 students (380 male, 873 female and one unidentified; mean age = 18.38). One participant from Sample 1A and two participants from Sample 1B did not fill out any of the measures, and thus were excluded from the analyses. The final sample size for Sample 1A and Sample 1B were 1093 and 1252 respectively.

Measures

For purposes of this investigation, I used data from measures obtained in online mass testing sessions (see below and APPENDIX A). I selected these measures because all except one (Belief in Zero-Sum Resources) included equal numbers of positively- and negatively-valenced items. This allowed me to compute full-scale scores as well as positively- and negatively-valenced scale scores for each construct. I also selected a measure of social desirability response bias (available for Sample 1B only) to be used in analyses to compare the effects of valence and social desirability response bias. The reliability information for each of the measures is shown in Table 3.

Table 3

Scale Reliabilities in Study 1

	Total number of items	Full scale	Positively-valenced scale	Negatively-valenced scale
Sample 1A				
Conscientiousness	10	.80	.67	.73
Extraversion	10	.87	.81	.77
Agreeableness	10	.76	.67	.70
Openness	10	.76	.52	.64
Neuroticism	10	.82	.66	.73
Self-Esteem	10	.88	.80	.82
SDO	16	.91	.87	.86
BDW	12	.81	.71	.72
BZSR	6	.91	-	.93
Average reliability		.84	.71	.77
Sample 1B				
Conscientiousness	10	.77	.65	.70
Extraversion	10	.87	.81	.78
Agreeableness	10	.76	.64	.72
Openness	10	.75	.53	.59
Neuroticism	10	.84	.67	.78
Self-Esteem	10	.90	.84	.84
SDO	16	.94	.87	.87
Average reliability		.83	.72	.75
BIDR Social Desirability Total	38	.77	-	-
BIDR Impression Management	19	.76	-	-
BIDR Self-Deception	19	.66	-	-

Note. No reliability (internal consistency) was estimated for the single positively-valenced item of BZSR. SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World; BZSR = Belief in Zero-Sum Resources; BIDR = Balanced Inventory of Desirable Responding. All measurement instrument include a balanced number of positively-valenced and negatively-valenced items, except BZSR (only one positively-valenced item), BIDR Social Desirability total (18 regular-keyed items), BIDR Impression Management (nine regular-keyed items), and BIDR Self-Deception (nine regular-keyed items).

Personality. The Big Five personality factors (Conscientiousness, Extraversion, Agreeableness, Openness to Experience, and Neuroticism) were measured with scales (NEO domain) taken from the International Personality Item Pool (IPIP; Goldberg et al., 2006). Each factor consisted of 10 items with a 5-point Likert-type scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*).

Self-Esteem. The Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965) consisted of 10 items, and measured respondents' global evaluation of self-worth. Each item was measured with a 4-point scale from 1 (*Strongly Disagree*) to 4 (*Strongly Agree*). A sample item is "On the whole, I am satisfied with myself."

Social Dominance Orientation (SDO). The SDO scale measures respondents' preference for inequality and hierarchical differentiation in a social context. SDO was measured by 16 items, developed by Sidanius and Pratto (2001), with a 7-point scale from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). A sample item is "To get ahead in life, it is sometimes necessary to step on other groups."

Belief in a Dangerous World (BDW). The BDW Scale (Altemeyer, 1988) was composed of 12 items, each of which was measured with a 9-point scale from -4 (*Very Strongly Disagree*) to +4 (*Very Strongly Agree*). BDW measured respondents' belief that the world is a dangerous and threatening place. A sample item is "It seems that every year there are fewer and fewer truly respectable people, and more and more persons with no morals at all who threaten everyone else." This scale was only available for Sample 1A.

Belief in Zero-Sum Resources (BZSR). This 6-item BZSR Scale is a revised and shortened version of the original BZSR measure by Esses, Jackson, and Armstrong (1998). BZSR measures one's beliefs that immigrants are competing with Canadians for valuable resources in society. A sample item is "Money spent on social services for immigrants means less money for services for Canadians already living here". The current measure contains one

reverse-keyed item. The ZSBI Scale was only included in Sample 1A. Each item was measured with a 7-point scale from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*).

Balanced Inventory of Desirable Responding (BIDR). The BIDR scale (Paulhus, 1991) was included in Sample 1B to measure social desirability responding, or one's tendency to project a positive social image in a survey. It consisted of 38 items[1] measured with a 5-point Likert-type scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*) in my study. Half of the items are reverse-keyed. The original scale consists of two subscales, namely impression management and self-deception. Researchers (e.g., Li & Bagger, 2007) generally conceptualized impression management as *intentional* distortion of a self-image whereas self-deception as an *unintentional* propensity to exaggerate positive attributes. The author of the scale (Paulhus, 2002) has also stated that impression management and self-deception differ on the level of awareness. A sample item for impression management is "My first impression of people usually turns out to be right." A sample item for self-deception is "I never cover up my mistakes." I used two methods to obtain scores for impression management and self-deception. The first method involved averaging participants' ratings on relevant items (see Stöber, Dette, & Musch, 2002). I refer to this method as averaged scores method. The second method using dichotomous item scores was suggested by the original author (Paulhus, 1991). Using this method, I first reversed the scores for reverse-keyed items. For the self-deception scale, item scores of 5 were converted to 1 and scores of 4 or less were converted to 0. For the impression management scales, scores of 4 or 5 were converted to 1 and scores of 3 or less were converted to 0. The item scores were then summed with high

[1] Two items relevant to sex-relevant behaviors (an impression management item) and love-related cognition (a self-deception item) were removed from the data collection.

scores reflecting greater self-deception and impression management, respectively. I refer to this method as Paulhus scoring method.

Results

Demonstrating the Role of Item Valence in Exploratory Factor Analysis

I first evaluated the potential impact of item valence on exploratory factor analyses (EFA). Despite the availability of more advanced factor analytic techniques, EFA continues to be the most common method in examining the factor structure of a construct (Furr, 2011), and it is usually the first step in data analysis before a more advanced procedure (such as confirmatory factor analysis). Therefore, I subjected both self-esteem items and extraversion items to principal component analyses (PCA) as the extraction method and direct oblimin transformation as the rotation method. I also conducted principal axis factor analysis (PAF) with direct oblimin rotation. However, because this method yielded a similar conclusion to the PCA method, the PAF results are not further elaborated here.[2] To determine the number of factors to extract, I used the scree test because most researchers are more familiar with this method compared to others (Furr, 2011). The scree test retains factors before a natural bend in a plot of eigenvalues. The results of these scree test analyses are shown in Figure 1 and Figure 2.

For self-esteem in both Sample 1A and 1B, the scree tests suggested the retention of two factors (and the possibility of three factors in Sample 1A). In Table 4 I examined factor loadings of the self-esteem items in both a one-factor solution and a two-factor solution. In

---

[2] PAF results were consistent with the conclusion that item valence exists in both extraversion and self-esteem measures in Sample 1A, and the extraversion measure in Sample 1B.

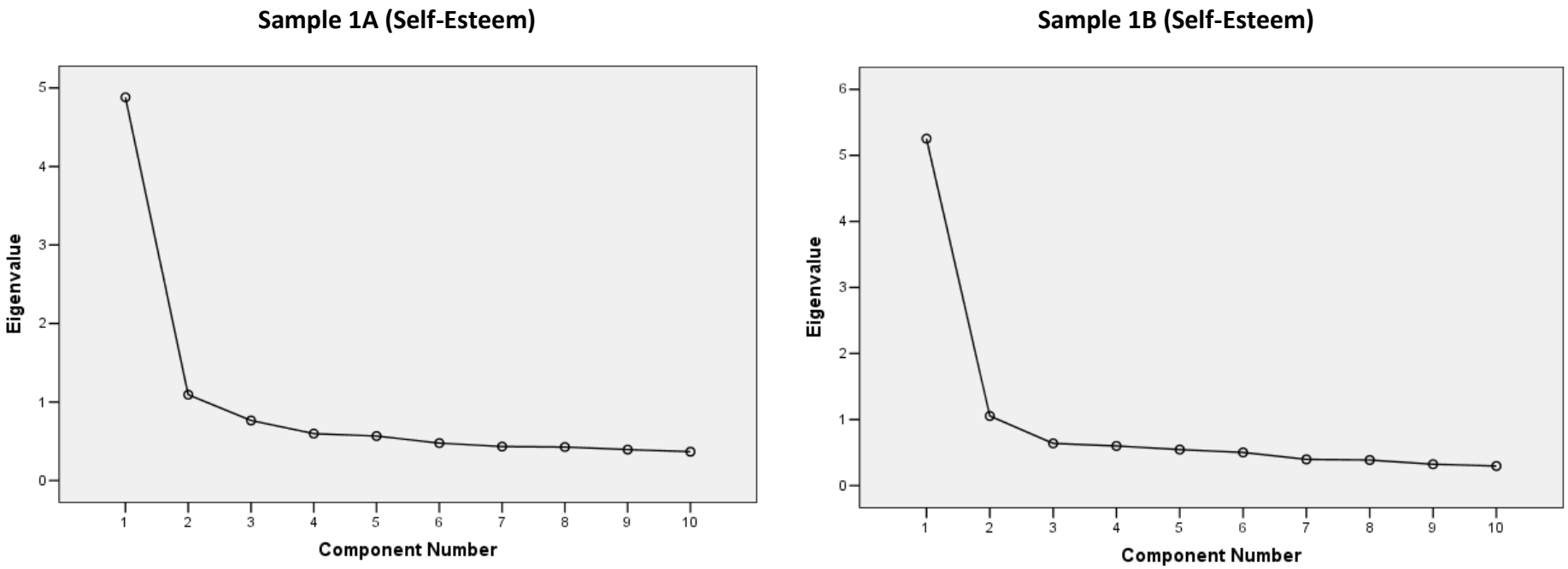

Figure 1. Plots of eigenvalues for Rosenberg Self-Esteem Scale.

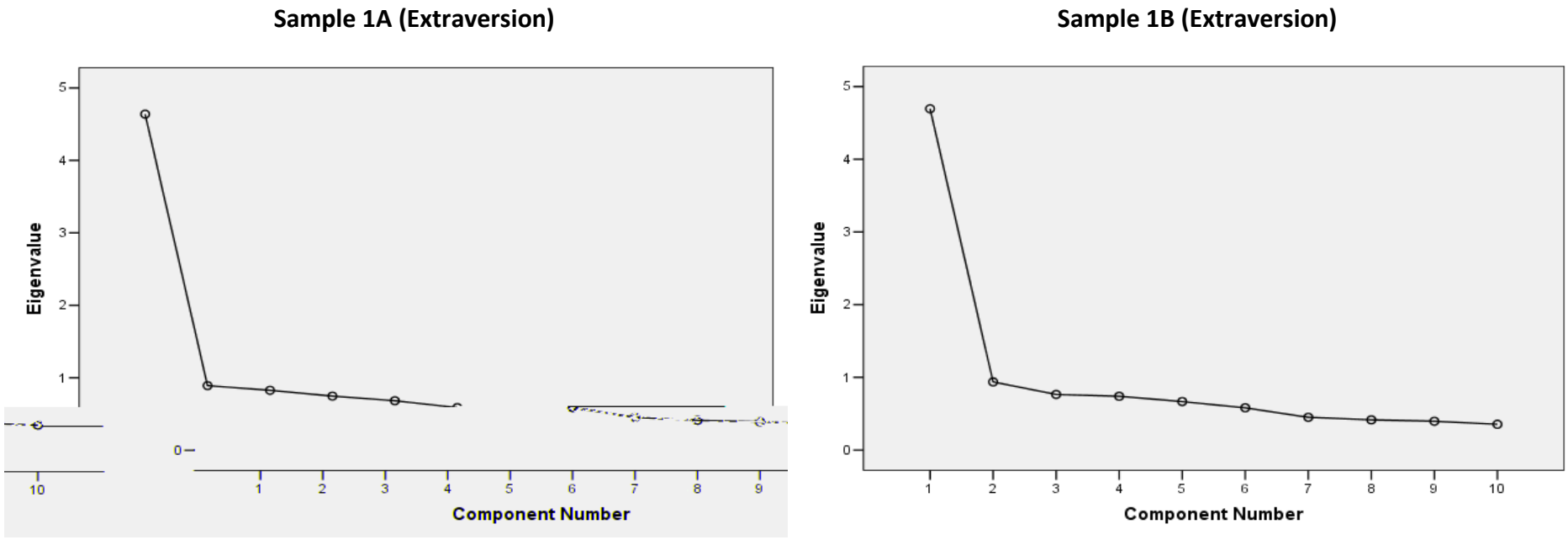

Figure 2. Plots of eigenvalues for Extraversion scale from IPIP.

Table 4

Factor Loadings From Exploratory Factor Analyses

	Sample 1A			Sample 1B		
	One-Factor	Two-Factor		One-Factor	Two-Factor	
		First Factor	Second Factor		First Factor	Second Factor
Self-Esteem						
Item 1	**.65**	.20	-.53	**.78**	.51	.38
Item 2	**.69**	**.82**	-.04	**.69**	**.86**	-.06
Item 3	**.61**	**.86**	.08	**.65**	**.89**	-.13
Item 4	**.71**	**.70**	-.17	**.70**	**.63**	.19
Item 5	**.79**	.29	**-.60**	**.81**	.40	.53
Item 6*	**-.68**	.12	**.84**	**-.71**	.10	**-.90**
Item 7*	**-.73**	-.30	**.54**	**-.74**	-.45	-.39
Item 8*	**-.68**	.11	**.83**	**-.70**	.08	**-.86**
Item 9*	**-.65**	.07	**.77**	**-.70**	-.04	**-.74**
Item 10*	**-.77**	-.27	**.61**	**-.76**	-.32	-.54
Factor Correlation			-.50			-.53
Extraversion						
Item 1	**.69**	**.84**	-.10	**.70**	**.81**	.06
Item 2	**.73**	**.88**	-.10	**.75**	**.81**	.00
Item 3	**.76**	**.76**	.08	**.74**	**.86**	.07
Item 4	**.68**	.42	.37	**.70**	.49	-.30
Item 5	**.64**	.53	.18	**.62**	**.64**	-.03
Item 6*	**-.66**	.21	.58	**-.67**	-.19	**.70**
Item 7*	-.56	.36	.28	**-.77**	-.12	**.68**
Item 8*	**-.77**	.39	.51	-.56	-.28	**.61**
Item 9*	-.53	-.19	**.86**	-.54	-.49	.12
Item 10*	**-.74**	.22	**.67**	**-.76**	.16	**.83**
Factor Correlation			-.52			-.56

Note. * = reverse-keyed item. Principal Components Analysis with Direct Oblimin Rotation. Factor loadings at or above .30 are underlined. Factor loadings at or above .60 are underlined and bolded. Factor correlation refers to the correlation between the first factor and the second factor.

the two-factor solution, interestingly, positively-valenced and negatively-valenced items with strong loadings (defined as loading ≥ .60) tended to load on separate factors (see Table 4). For extraversion in both Sample 1A and 1B, the scree tests supported a one-factor solution in Sample 1A but suggested the possibility of a second factor in Sample 1B, as evidenced by a slight bend in the eigenvalue plot after the second factor. Table 4 shows the factor loadings for both a one-factor and a two-factor solution for extraversion. When I forced a solution based on two factors, regularly-keyed and reverse-keyed items with strong loadings (defined as loading ≥ .60) tended to cluster in separate factors. Even though the scree test suggested a one-factor solution for the construct extraversion, the two-factor solution appeared to reveal additional information (Paunonen & Jackson, 1979).

Multitrait-Multimethod Analysis

In the current study I used the correlated trait-uncorrelated method (CTUM) MTMM model to investigate the potential impact of item valence. My decision to use the CTUM model was based on past studies that demonstrated its estimation accuracy (Marsh et al., 2010). In addition, CTUM does not overestimate the method effect, which is a problem that is found in the correlated trait-correlated method (CTCM) model (Marsh, 1989; see also Marsh & Bailey, 1991). Similar to many other MTMM methods, CTUM requires a large sample size for successful model convergence. The sample size issue is not a problem of the current study.

As explained earlier, I compared a baseline model to a method-factor model. In the baseline model ($M_{baseline}$), item indicators load on the intended construct factors only (see Figure 3). For example, all of the indicators for extraversion only loaded on the intended construct factor for extraversion. In the method-factor model ($M_{2valence}$), I included two valence (method) factors and

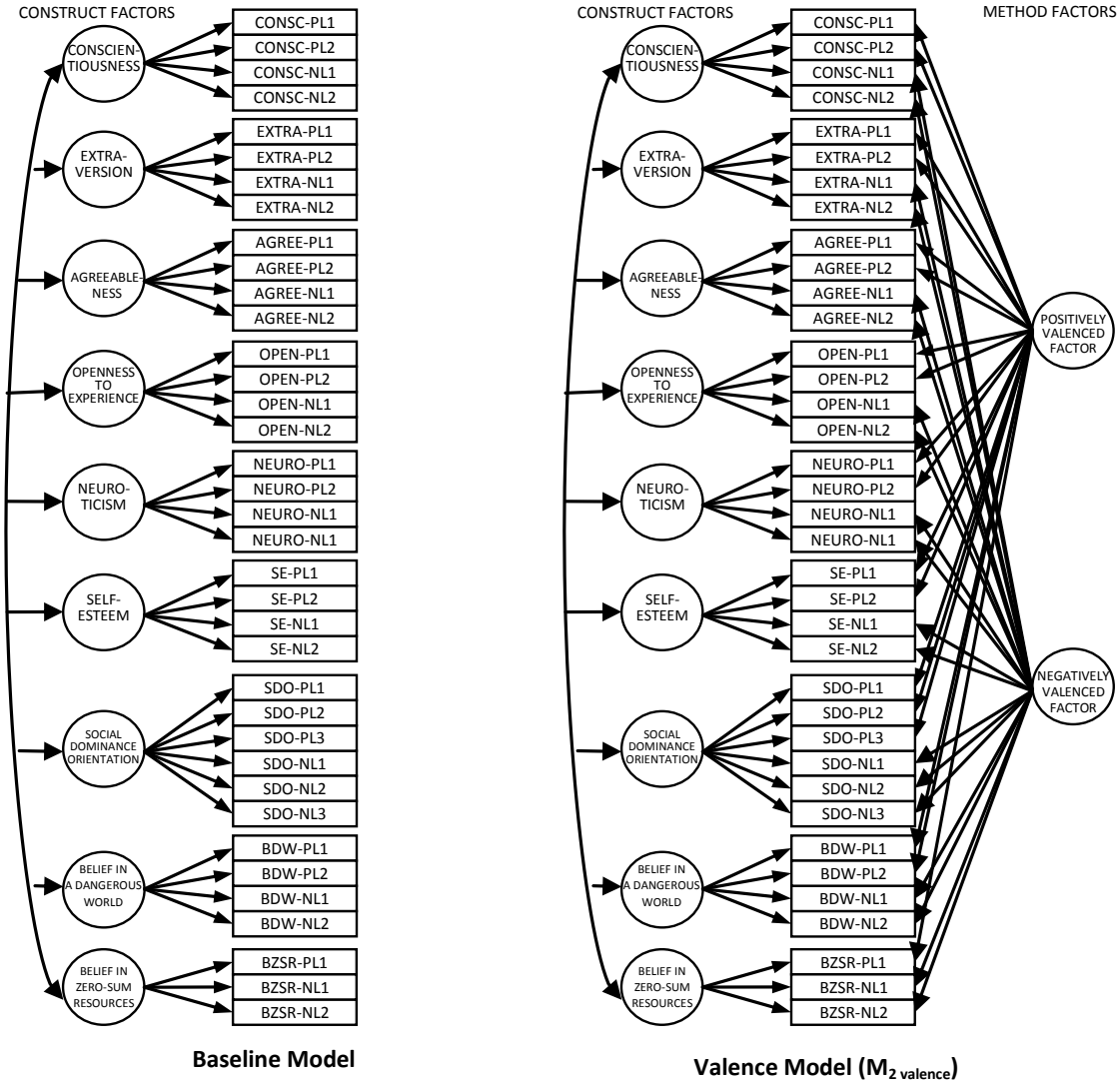

Figure 3. Models comparisons in the current study. PL = parcel for positively-valenced items; NL = parcel for negatively-valenced items.

allowed item indicators to load on *both* the intended construct factor *and* the corresponding valence factor (positive or negative). If $M_{2valence}$ fits the data better than does $M_{baseline}$, it supports the notion that shared valence contributes to the correlations among the items.

I included all variables in the MTMM CFA, because I am interested in examining the effect of item valence in psychological scales in general rather than in a particular measurement instrument (i.e., a self-esteem measure or an extraversion measure). Before the analyses, all of the scale items were parceled because the MTMM CFA solutions often do not converge satisfactorily when too many item indicators are included in the model (Bentler & Chou, 1987; Yuan, Bentler, & Kano, 1997). In my empirical study, I aggregated individual items together within a measure to form parcel indicators because parcel indicators are more psychometrically reliable and have better distributional characteristics than individual items (Nasser & Wisenbaker, 2003). I adopted the most recent recommendation by Alhija and Wisenbaker (2006) to have four to six-item parcels for each construct factor in my model. To create a parcel indicator I followed the same practice as Kwan, Bond, and Singelis (1997; see also Yuan, Bentler, & Kano, 1997). For items that belonged to the same construct and had the same valence, I first conducted exploratory factor analyses. I then averaged the items with the highest and lowest factor loadings to form the first parcel, the items with the second highest and second lowest factor loadings to form the second parcel, and so on. Because ZSBI only has one positively-valenced item, this item was not parceled.

The MTMM CFA analyses were conducted with the program Mplus 6.0 (Muthén & Muthén, 1998-2010) with the maximum likelihood robust estimator (MLR). Missing data were estimated with the default option of full-information maximum likelihood (FIMR) in the Mplus program. All construct factors and method factors were set to have a variance of unity and all factor loadings were freely estimated. $M_{baseline}$ was nested within $M_{2valence}$ and thus the

two models were compared directly with the chi-square difference test (Satorra & Bentler, 2001).

For both Samples 1A and 1B, the fit of $M_{2valence}$ was a significant improvement over that for $M_{baseline}$, indicating the existence of the two valence factors (see Table 5 for fit indices; $M_{2keying}$ and M_{common} in Table 5 will be explained later in this results section; see Table 6 [Comparison 1] for chi-square difference test statistics). Therefore, across two independent samples these results are consistent with the hypothesis that correlations among the scale items reflect both their content and their valence. This finding shows that the valence effect has a strong potential to cause a construct to form two factors in factor analysis, regardless of the dimensionality of the construct itself. The factor loadings of the MTMM model in two samples are shown in Table 7. Readers may notice that the loadings for SDO items on the negative valence factor are much higher than those of the other items. To ensure that my results are not simply dominated by SDO items, I conducted another MTMM analysis excluding SDO items in the model. The analysis continued to suggest the existence of two valence factors.

Nomological network analyses

To determine how construct valence affects the correlations of measures with external variables, I compared how positive and negative self-esteem, and extraversion and introversion, correlated with the other variables measured in Samples 1A and 1B.

Self-Esteem. As mentioned previously, I created positive self-esteem scores (from the positively-valenced self-esteem items) and negative self-esteem scores (from the negatively-valenced self-esteem items in the same scale) and compared their correlations with other variables in each sample. For these other variables, I also created full, positively-valenced,

Table 5

Model Comparisons in Both Sample 1A and 1B

	Absolute Fit Indices							Predictive Fit Indices	
	χ^2	*df*	*p*	TLI	CFI	RMSEA	SRMR	AIC	SABIC
Sample 1A									
Baseline model	3251.72	593	<.001	.82	.84	.06	.06	88148.38	88415.99
$M_{2valence}$	1766.46	556	<.001	.91	.93	.05	.05	86552.46	86887.42
$M_{2keying}$	1987.11	556	<.001	.90	.92	.05	.06	86798.44	87133.40
M_{common}	2201.62	556	<.001	.88	.90	.05	.05	86998.88	87323.84
Sample 1B									
Baseline model	3140.73	384	<.001	.82	.84	.08	.07	76883.34	77100.46
$M_{2valence}$	1577.67	354	<.001	.91	.93	.05	.06	75241.83	75517.63
$M_{2keying}$	1732.07	354	<.001	.90	.92	.06	.06	75404.57	75680.37
M_{common}	2154.50	354	<.001	.87	.90	.06	.05	75793.88	76069.68

Note. TLI = Tucker Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; AIC = Akaike Information Criterion; SABIC = Sample-size Adjusted Bayesian Information Criterion. Maximum likelihood estimator with robust standard errors is used in these test results.

Table 6

Model Comparisons

	Are they nested models?	Sample #	Comparison between nested models			Comparison between non-nested models		Preferred Model
			$\Delta\chi^2$	Δdf	p	AIC	SABIC	
1st set of Comparison								
(1) Baseline vs $M_{2valence}$	Y	1	1215.91	37	<.001			$M_{2valence}$
		2	1489.21	30	<.001			$M_{2valence}$
2nd set of Comparison								
(2a) Baseline vs $M_{2keying}$	Y	1	1083.23	37	<.001			$M_{2keying}$
		2	1537.70	30	<.001			$M_{2keying}$
(2b) $M_{2valence}$ vs. $M_{2keying}$	N	1				86552 vs. 86798	86887 vs. 87133	$M_{2valence}$
		2				75242 vs. 75405	75518 vs. 75680	$M_{2valence}$
3rd set of Comparison								
(3a) Baseline vs M_{common}	Y	1	1233.12	37	<.001			M_{common}
		2	750.92	30	<.001			M_{common}
(3b) $M_{2valence}$ vs. M_{common}	N	1				86552 vs. 86999	86887 vs. 87324	$M_{2valence}$
		2				75242 vs. 75794	75518 vs. 76070	$M_{2valence}$

Note. AIC = Akaike Information Criterion; SABIC = Sample-size Adjusted Bayesian Information Criterion. Satorra-Bentler (2001) scaled difference chi-square tests are used to compare between nested models. Absolute fit indices (AIC and SABIC) are used to compare between non-nested models. A model with lower AIC and SABIC values is preferred.

Table 7

Factor Loadings (Standardized) for the Model with Both Latent Constructs and Two Keying Method Factors (M_{method} Model)

		Sample 1A Factor loading on			Sample 1B Factor loading on		
construct	parcel #	construct factor	positively-valenced factor	negatively-valenced factor	construct factor	positively-valenced factor	negatively-valenced factor
CONSC	PL-1	.55***	.37***		.51***	.30***	
CONSC	PL-2	.60***	.16***		.56***	.14***	
CONSC	NL-1	.79***		.15***	.75***		.12***
CONSC	NL-2	.63***		.26***	.68***		.18***
EXTRA	PL-1	.72***	.45***		.70***	.45***	
EXTRA	PL-2	.74***	.28***		.72***	.29***	
EXTRA	NL-1	.78***		.10***	.86***		.10***
EXTRA	NL-2	.79***		.11***	.70***		.11***
AGREE	PL-1	.54***	.55***		.50***	.40***	
AGREE	PL-2	.46***	.40***		.42***	.51***	
AGREE	NL-1	.78***		.20***	.82***		.18***
AGREE	NL-2	.58***		.19***	.68***		.15***
OPEN	PL-1	.77***	.18***		.86***	.19***	
OPEN	PL-2	.34***	.45***		.30***	.42***	
OPEN	NL-1	.86***		.14***	.80***		.04
OPEN	NL-2	.60***		.14***	.42***		.15***
NEURO	PL-1	.73***	.30***		.73***	.31***	
NEURO	PL-2	.61***	.12***		.66***	.10***	
NEURO	NL-1	.80***		.10***	.82***		.03
NEURO	NL-2	.70***		.10***	.84***		.07**
SE	PL-1	.74***	.31***		.77***	.29***	
SE	PL-2	.73***	.30***		.79***	.30***	
SE	NL-1	.84***		.09***	.81***		.08***
SE	NL-2	.78***		.08***	.82***		.07**
SDO	PL-1	.85***	.04		.83***	.10***	
SDO	PL-2	.86***	.12***		.86***	.10***	
SDO	PL-3	.76***	.13***		.83***	.10***	
SDO	NL-1	.66***		.47***	.62***		.59***
SDO	NL-2	.63***		.60***	.68***		.51***
SDO	NL-3	.51***		.63***	.52***		.63***

(continue to the next page)

(continue from the previous page)

		Sample 1A Factor loading on			Sample 1B Factor loading on		
construct	parcel #	construct factor	positively-valenced factor	negatively-valenced factor	construct factor	positively-valenced factor	negatively-valenced factor
BDW	PL-1	.68***	.17***				
BDW	PL-2	.63***	.11***				
BDW	NL-1	.70***		.25***			
BDW	NL-2	.65***		.19***			
BZSR	PL-1	.49***	.16***				
BZSR	NL-1	.95***		.12***			
BZSR	NL-2	.85***		.10***			
Variance Explained		.50	.09	.07	.51	.09	.08

Note. CONSC = Conscientiousness; AGREE = Agreeableness; EXTRA = Extraversion; OPEN = Openness to Experience; NEURO = Neuroticism; SE = Self-Esteem; SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World; BZSR = Belief in Zero-Sum Resources; PL = parcels for positive-keyed constructs; NL = parcels for negative-keyed constructs

Example: "AGREE PL-2" stands for the second parcel of the agreeableness score that were measured with positive-valenced items.

The positive or negative signs for the factor loadings are not shown for simplicity. Within a construct, positively- and negatively-valenced item parcels have opposite signs for their factor loadings. All factor loadings for positively-valenced factor are significant and share the same sign; all factor loadings for negatively-valenced factor are also significant and share the same sign *$p < .05$; ***$p < .001$.

and negatively-valenced scale scores[3]. All correlation coefficient comparisons were based on the statistical formula provided by Meng, Rosenthal, and Rubin (1992). The results of the analyses are shown in Table 8.

For Sample 1A, neither positive nor negative self-esteem had much advantage over each other in terms of their correlations with other full-scale variables. In cases where the correlations differed significantly, two favored positive self-esteem and one favored negative self-esteem (see the top left panel). However, a different picture emerged when correlations with positively- and negatively-valenced scale scores were compared. For positively-valenced scale scores, four out of eight cases favored positive self-esteem (see the middle left panel). In contrast, for negatively-valenced scale scores, four comparisons favored negative self-esteem and only one comparison favored positive self-esteem (see the bottom left panel).

Note that even for the one case that favored positive self-esteem, the significance level was rather low ($p < .03$). Therefore, Sample 1A showed that positive self-esteem scores tend to correlate more strongly with positively-valenced scores than do negative self-esteem scores, and conversely, negative self-esteem scores tend to correlate more strongly with negatively-valenced scores than do positive self-esteem scores. The results for Sample 1B were similar to the results for Sample 1A, and I will not further elaborate on these results here.

I further added in full-scale scores of self-esteem from the Sample 1A and Sample 1B data and compared how half-scale positive self-esteem, full-scale self-esteem, and negative self-esteem scores correlate with other constructs. By comparing the pattern of correlations between

[3] Belief in Zero Sum Resources (BZSR) only contains one positively-valenced item. To avoid its full-scale score being over-represented by its five negatively-valenced items, I first calculated an overall mean score of these negatively-valenced items. Afterwards, I took the average between its positive-valanced item score and the overall negatively-valenced item score.

Table 8

Comparing the Nomological Networks between Positive and Negative Self-Esteem in Sample 1A and Sample 1B

	Sample 1A				Sample 1B			
	Positive self-esteem (PSE)	Negative self-esteem (NSE)	*Z*		Positive self-esteem (PSE)	Negative self-esteem (NSE)	*Z*	
Full-scale scores								
Conscientiousness	.350	.366	-0.73		.326	.382	-2.85**	(NSE stronger)
Extraversion	.453	.442	0.53		.404	.377	1.40	
Agreeableness	.207	.194	0.56		.179	.191	-0.58	
Openness	.152	.056	4.06***	(PSE stronger)	.068	.012	2.64**	(PSE stronger)
Neuroticism	-.621	-.679	-3.41***	(NSE stronger)	-.646	-.688	-2.82**	(NSE stronger)
SDO	-.090	-.085	0.21		-.022	-.044	-1.04	
BDW	-.154	-.167	-0.55					
BZSR	-.074	-.019	2.31*	(PSE stronger)				
Positive-valenced scores								
Conscientiousness	.318	.279	1.73		.283	.285	-0.10	
Extraversion	.440	.392	2.25*	(PSE stronger)	.422	.364	3.01**	(PSE stronger)
Agreeableness	.303	.194	4.75***	(PSE stronger)	.248	.168	3.86***	(PSE stronger)
Openness	.214	.081	5.66***	(PSE stronger)	.129	.045	3.97***	(PSE stronger)
Neuroticism	-.571	-.584	-0.70		-.611	-.604	0.43	
SDO	-.092	-.054	1.60		.049	.033	0.75	
BDW	-.156	-.116	1.70					
BZSR	-.111	-.029	3.45***	(PSE stronger)				
Negative-valenced scores								
Conscientiousness	.300	.358	-2.61**	(NSE stronger)	.285	.375	-4.53***	(NSE stronger)
Extraversion	.395	.420	-1.17		.324	.329	-0.25	
Agreeableness	.083	.145	-2.62**	(NSE stronger)	.079	.157	-3.70***	(NSE stronger)
Openness	.082	.029	2.23*	(PSE stronger)	.010	-.016	-0.28	
Neuroticism	-.569	-.656	-4.87***	(NSE stronger)	-.595	-.673	-5.00***	(NSE stronger)
SDO	-.071	-.094	-0.97		-.005	.046	-1.93	
BDW	-.116	-.177	-2.60***	(NSE stronger)				
BZSR	-.012	-.003	-0.38					

Note. SE = Self-Esteem; SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World; BZSR = Belief in Zero-Sum Resources. SDO, BDW and BZSR were not measured in Sample 1B. *$p < .05$, **$p < .01$, ***$p < .001$.

full-scale scores and half-scale scores, I can examine how correlations can be inflated or deflated by construct valence.[4] In addition, I tested this predicted pattern of results with all correlation coefficients corrected for unreliability, because the reliability of a full-scale variable is generally higher than the reliability of its half-scale counterparts, and unreliability tends to attenuate the strength of correlation coefficients (Meeker & Escobar, 1998).

If scores sharing the same valence direction have an inflated correlation and scores with opposing valence directions have a deflated correlation, I would expect that in a majority of cases the positively-valenced variables would correlate most highly with the positive self-esteem score, followed by the full-scale self-esteem score and finally by the negative self-esteem score. The results are shown in the middle panel of Table 9.[5] Indeed, for Sample 1A (see the left side of the analyses), when other variables in the investigations were positively valenced, the expected pattern of correlations was discovered the majority of the time (six out of seven times; see the middle left panel). Conversely, I predicted that negatively-valenced variables would correlate most highly with the negative self-esteem score, followed by the full-scale self-esteem score and finally by the positive self-esteem score. The result is shown in the bottom panel of Table 9. Again, the expected correlational pattern was found five out of seven times (see the bottom left panel). A similar pattern of results was found for Sample 1B (see the right panel in Table 9).

---

[4] I assume a full-scale score as the standard of comparison because it gives a more balanced representation of a bipolar construct than any half-scale scores. For example, a half-scale score of introversion may be unable to fully represent its opposite pole of extraversion.

[5] BZSR is not included in the analysis because the unreliability of its single positively-valenced item cannot be estimated. Therefore, the correlation cannot be corrected.

Table 9

Comparing the Nomological Networks among Positive Self-Esteem, Negative Self-Esteem, and Full-scale Self-Esteem After Unreliability Correction in Sample 1A and 1B

	Sample 1A				Sample 1B			
	Positive self-esteem (PSE)	Full self-esteem (SE)	Negative self-esteem (NSE)	Expected Pattern of Correlational Strength	Positive self-esteem (PSE)	Full self-esteem (SE)	Negative self-esteem (NSE)	Expected Pattern of Correlational Strength
Full-scale scores								
Conscientiousness	.437	.467	.454		.405	.463	.475	
Extraversion	.543	.556	.525		.472	.474	.442	
Agreeableness	.265	.267	.246		.224	.242	.239	
Openness	.195	.133	.072		.086	.048	.015	
Neuroticism	-.766	-.837	-.830		-.767	-.830	-.820	
SDO	-.106	-.108	-.098		-.025	-.041	-.050	
BDW	-.191	-.208	-.205					
Positive-valenced scores				PSE>SE>NSE?				PSE>SE>NSE?
Conscientiousness	.432	.419	.376	Yes	.384	.403	.388	No
Agreeableness	.545	.532	.482	Yes	.511	.492	.442	Yes
Openness	.412	.344	.261	Yes	.339	.290	.231	Yes
Neuroticism	.331	.226	.124	Yes	.193	.127	.068	Yes
SE	-.783	-.823	-.793	No	-.814	-.844	-.808	No
SDO	-.110	-.089	-.064	Yes	-.057	-.049	-.039	Yes
BDW	-.207	-.184	-.152	Yes				
Negative-valenced scores				NSE>SE>PSE?				NSE>SE>PSE?
Conscientiousness	.392	.452	.464	Yes	.373	.457	.492	Yes
Agreeableness	.503	.541	.530	No	.401	.423	.409	No
Openness	.110	.163	.191	Yes	.102	.164	.202	Yes
Neuroticism	.114	.077	.040	No	.014	-.007	-.023	No
SE	-.742	-.838	-.849	Yes	-.736	-.824	-.835	Yes
SDO	-.085	-.105	-.111	Yes	.006	-.028	-.054	Yes
BDW	-.153	-.206	-.230	Yes				

Note. SE = Self-Esteem; SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World. BZSR was excluded from the analysis because it only had one positively-valenced item. The corrected correlations of this item with other variables thus could not be examined.

Extraversion. My previous analyses clearly demonstrated that item valence can affect a researcher's substantive conclusion regarding the dimensionality of self-esteem. Here, I further examined the replicability and the generalizability of this effect with a common personality dimension – introversion-extraversion. I was particularly interested in how extraversion scores (measured by positively-valenced extraversion items) and introversion scores (measured by negatively-valenced introversion items) correlated with other variables. The results of the analyses are shown in Table 10. Extraversion and introversion did not have much advantage over each other in terms of their correlations with other full-scale variables; two favored extraversion and one favored introversion (see the top left panel). For correlations with positively-valenced scores, however, extraversion was significantly stronger than introversion for four out of eight comparisons (see the middle left panel). For correlation with negatively-valenced scores, introversion was significantly stronger than extraversion for four out of eight comparisons (see the bottom left panel). Extraversion only showed advantage in one of these eight comparisons. Similar results were found in Sample 1B, with the exception that extraversion seemed to have somewhat more advantage over introversion in correlating with other full-scale variables (see the top right panel of Table 10). Nevertheless, the noteworthy results here were that the correlations favored extraversion even more when the other variables were positively valenced. Specifically, four out of six correlation comparisons favored extraversion (see the middle right panel). In contrast, the advantage of extraversion diminished when the other variables were negatively valenced (see the bottom right panel). Correlation comparisons favored introversion in two out of six comparisons and favored extraversion in one comparison.

Summarizing the results for extraversion here, its positively-valenced scores and its negatively-valenced scores correlate somewhat differently with other variables based on the

Table 10

Comparing the Nomological Networks between Extraversion and Introversion in Sample 1A and Sample 1B

	Sample 1A				Sample 1B			
	Half-Scale Extraversion (EXT)	Half-Scale Introversion (INT)	*Z*		Half-Scale Extraversion (EXT)	Half-Scale Introversion (INT)	*Z*	
Full-scale scores								
Conscientiousness	.167	.209	-1.80		.164	.154	0.46	
Agreeableness	.169	.095	3.15**	(EXT stronger)	.126	.032	4.23***	(EXT stronger)
Openness	.172	.221	-2.11*	(INT stronger)	.148	.176	-1.28	
Neuroticism	-.429	-.380	2.29*	(EXT stronger)	-.404	-.300	5.05***	(EXT stronger)
SE	.449	.445	0.19		.420	.353	3.31***	(EXT stronger)
SDO	-.027	-.057	-1.26		.019	-.031	-0.54	
BDW	-.040	-.086	-1.94					
BZSR	.024	-.029	-0.21					
Positive-valenced scores								
Conscientiousness	.193	.130	2.69**	(EXT stronger)	.180	.096	3.81***	(EXT stronger)
Agreeableness	.364	.192	7.59***	(EXT stronger)	.308	.147	7.45***	(EXT stronger)
Openness	.247	.242	0.22		.267	.244	1.07	
Neuroticism	-.428	-.306	5.59***	(EXT stronger)	-.397	-.246	7.23***	(EXT stronger)
SE	.440	.395	2.12*	(EXT stronger)	.422	.324	4.81***	(EXT stronger)
SDO	-.056	-.038	0.76		-.015	-.010	0.22	
BDW	-.056	-.054	0.08					
BZSR	-.026	-.066	-1.69					
Negative-valenced scores								
Conscientiousness	.108	.227	-5.09***	(INT stronger)	.111	.166	-2.50**	(INT stronger)
Agreeableness	-.019	-.003	0.67		-.041	-.064	-1.03	
Openness	.087	.171	-3.57***	(INT stronger)	.032	.097	-2.92**	(INT stronger)
Neuroticism	-.364	-.385	-0.97		-.360	-.308	2.50*	(EXT stronger)
SE	.392	.420	-1.31		.364	.329	1.69	
SDO	.001	-.064	-2.65**	(INT stronger)	.045	-.043	0.09	
BDW	-.016	-.091	-3.16**	(INT stronger)				
BZSR	.070	.019	2.15*	(EXT stronger)				

Note. SE = Self-Esteem; SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World; BZSR = Belief in Zero-Sum Resources. SDO, BDW and BZSR were not measured in Sample 1B. *$p < .05$, **$p < .01$, ***$p < .001$.

valence of these other variables. The slight difference between the Sample 1A results and the Sample 1B results should not be alarming because it can simply the outcome of sampling errors (see Schmidt & Hunter, 1996, 1999). Random sampling errors can cause the observed statistical results to fluctuate across samples (Schmidt, 1992). The overall conclusion drawn from the extraversion measure, however, is the same as that drawn from the self-esteem measure. It should also be noted that these results for both extraversion and self-esteem cannot simply be due to the impact of differential reliabilities on correlation coefficients, because both extraversion and introversion possess a similar level of reliability (α = .81 vs. .77 in Sample 1A; .81 vs. .78 in Sample 1B), as do positive and negative self-esteem (α = .80 vs. .82 in Sample 1A; both αs = .84 in Sample 1B).

Next I compared the corrected correlations among the extraversion, full-scale, and introversion scores with other constructs. To recall, for positively-valenced variables I expected them to correlate most strongly with extraversion scores in a majority of cases, followed by full-scale scores and introversion scores. For negatively-valenced variables I expected a reversed pattern of correlation strength. The results are shown in Table 11. For Sample 1A, when the other variables in the investigation were positively valenced (see the middle left panel), the expected pattern of correlations was found five out of seven times. Similarly, when the other variables in the investigations were negatively valenced (see the bottom left panel), the expected pattern of correlations was found six out of seven times. The same trend was also found for the Sample 1B results (see the right panel in Table 11), where the expected pattern was found five out of six times for positively-valenced variables and three out of six times for negatively-valenced variables. The results for the negatively-valenced constructs in Sample 1B were somewhat weaker, but were nonetheless consistent with my hypothesis. Thus, my previous

Table 11

Comparing the Nomological Networks among Half-scale Extraversion, Half-scale Introversion, and Full-scale After Unreliability Correction in Sample 1A and 1B

	Sample 1A				Sample 1B			
	Half-Scale Extraversion (EXT)	Full-Scale (FS)	Half-Scale Introversion (INT)	Expected Pattern of Correlational Strength	Half-Scale Extraversion (EXT)	Full-Scale (FS)	Half-Scale Introversion (INT)	Expected Pattern of Correlational Strength
Full-scale scores								
Conscientiousness	.208	.247	.267		.206	.210	.199	
Agreeableness	.215	.176	.124		.160	.104	.041	
Openness	.220	.264	.290		.190	.219	.232	
Neuroticism	-.527	-.521	-.479		-.488	-.445	-.371	
SE	.532	.556	.541		.493	.475	.423	
SDO	-.031	-.052	-.068		.021	-.009	-.037	
BDW	-.049	-.082	-.109					
Positive-valenced scores				EXT>FS>INT?				EXT>FS>INT?
Conscientiousness	.262	.229	.181	Yes	.248	.199	.136	Yes
Agreeableness	.493	.393	.267	Yes	.428	.330	.209	Yes
Openness	.380	.395	.382	No	.405	.407	.379	No
Neuroticism	-.585	-.524	-.429	Yes	-.538	-.455	-.341	Yes
SE	.546	.542	.502	Yes	.511	.473	.401	Yes
SDO	-.067	-.059	-.046	Yes	-.018	-.016	-.012	Yes
BDW	-.074	.077	-.073	No				
Negative-valenced scores				INT>FS>EXT?				INT>FS>EXT?
Conscientiousness	.141	.231	.303	Yes	.148	.194	.226	Yes
Agreeableness	-.025	-.015	-.004	No	-.054	-.072	-.085	Yes
Openness	.121	.189	.244	Yes	.046	.099	.143	Yes
Neuroticism	-.474	-.511	-.514	Yes	-.452	-.440	-.396	No
SE	.482	.525	.530	Yes	.442	.442	.409	No
SDO	.001	-.040	-.079	Yes	.054	.000	-.052	No
BDW	-.021	-.075	-.122	Yes				

Note. SE = Self-Esteem; SDO = Social Dominance Orientation; BDW = Belief in a Dangerous World. BZSR was excluded from the analysis because it only had one positively-valenced item. The corrected correlations of this item with other variables thus could not be examined.

conclusion regarding how valence can inflate or mitigate correlation coefficients is again supported.

Testing some alternative explanations with MTMM

Before concluding that item valence was the underlying cause for my results, I considered other alternative explanations for my findings. The first alternative explanation is that item keying directions (regular-keyed versus reverse-keyed items), rather than item valence (positive versus negative), cause the method effect. It is extremely important to clarify which of these explanations drives my results because researchers constantly confuse these two concepts. The second alternative explanation is that participants have an identical response style among all items, regardless of whether the items are positively valenced or negatively valenced. I tested both of these explanations with the MTMM technique.

I examined the first explanation — item keying directions causes the method effect. I constructed an item-keying method factor model ($M_{2keying}$) in which observed indicators were loaded on two method factors based on their keying direction (i.e., regular-keyed and reverse-keyed) instead of on their item valence (positive and negative). I first investigated whether this new model ($M_{2keying}$) fit the data better than did the baseline model ($M_{baseline}$), and I then investigated whether this new model fit better than did my valence model ($M_{2valence}$). My baseline model ($M_{baseline}$) was nested within $M_{2keying}$ but $M_{2keying}$ and $M_{2valence}$ were not nested within each other. Therefore, the chi-square difference test was used to compare between $M_{baseline}$ and $M_{2keying}$, and the predictive fit indices were used to compare between $M_{2keying}$ and $M_{2valence}$ (information concerning whether two models were nested is also shown in Table 6 and will not be discussed further in the text). I used two common predictive fit indices, namely Akaike Information Criterion (AIC) and Sample Size Adjusted Bayesian Information Criterion (SABIC). A model with lower AIC and SABIC indices is more preferable because it is more likely to replicate in a separate sample (Kline, 2005). The results showed that

although $M_{2keying}$ fit better than $M_{baseline}$, it fit worse than $M_{2valence}$ (see the second set of comparison in Table 6). In addition, in $M_{2keying}$ most loadings for the reverse-keyed method factor (13 out of 18 indicators in Sample 1A; 9 out of 15 indicators in Sample 1B) were not significant, further indicating an unacceptable model fit. Therefore, item valence is a preferred explanation compared to item keying direction.

I then examined the second explanation — participants have the same response style among all measurement items. I constructed a model (M_{common}) in which all items are loaded on a common method factor regardless of whether these items are valenced positively or negatively. M_{common} fit better than $M_{baseline}$ but worse than $M_{2valence}$ in both samples (see the third set of comparison in Table 6). In addition, in M_{common}, 14 out of 37 method factor loadings in Sample 1A, and 9 out of 20 method factor loadings in Sample 1B, were not significant. Therefore, not only do participants' response styles exist for both positively- and negatively-valenced items, but my result also suggests that the response patterns *differ* for the two types of items.

Percentage of Variance Explained by Item Valence

The MTMM analyses allowed me to calculate the magnitude of the overall construct effect and the overall valence effect of $M_{2valence}$ in two samples. The percentage of variance explained by the constructs is calculated by averaging the squared standardized factor loadings of all the constructs. Similarly, the percentage of variance explained by the valence is calculated by averaging the squared standardized factor loadings of the two valence factors. The two valence factors accounted for 7.87% of the variance in the observed scores in Sample 1A and 8.33% of the variance in the observed scores in Sample 1B. The valence effect, although not negligible in its magnitude, accounts for slightly less than 10% of the total variance in participants' responses in my samples. In contrast, the construct effect accounts for a substantial percentage of variance (49.71% in Sample 1A; 50.62% in Sample

1B). The rest of the variance is unexplained. It appears that both factor analysis and nomological network analysis can reveal the influence of valence even though it accounts for slightly less than 10% of the variance in a dataset.

Effect of Social Desirability Response Bias

To examine the nature of the method effects estimated in the MTMM procedures described above, I extracted the factor scores for positive valence and negative valence from $M_{2valence}$ and correlated these scores with the two facets of social desirability responding in Sample 1B. Recall that I computed social desirability scores on the BIDR using two different approaches (a) the averaged scores method and (b) Paulhus original scoring method. The results are shown in Table 12. Interestingly, with the averaged scores method, I found that the positive valence factor correlated both positively and significantly with self-deception but not with impression management. In contrast, the negative-valence factor correlated both negatively and significantly with impression management but not with self-deception. With Paulhus original scoring method, the positive valence factor correlated positively and significantly with both self-deception and impression management, whereas the negative valence method factor correlated negatively and significantly with impression management but not self-deception. In summary, social desirability responding correlated positively with the positive valence factor and negatively with the negative valenced factor, supporting the validity of my MTMM model in extracting method effect. However, I controlled for the effect of the two facets of social desirability response bias with partial correlation analysis when comparing the correlations of positive and negative self-esteem and of extraversion and introversion with other constructs in Sample 1B. The overall pattern of results did not disappear. This further suggests that social desirability response bias cannot fully explain my results.

Table 12

Correlation between Valence Factors and Social Desirability Response Bias

	Averaged scores		Paulhus Scoring Method	
	Impression Management	Self-Deception	Impression Management	Self-Deception
Positive valence factor scores	.01	.22***	.20***	.24***
Negative valence factor scores	-.21***	-.03	-.09**	-.02

** $p < .01$; *** $p < .001$.

Summary

Study 1 showed that item valence can affect both factor analysis and nomological network analysis. My first hypothesis stated that positively-valenced items and negatively-valenced items will load on two separate factors in factor analysis, and this hypothesis was supported. My second hypothesis stated that valence can distort the results of a nomological network analysis. Construct correlations reflected both content and valence, and this hypothesis was also supported. My third research question addressed whether social desirability response bias can explain my findings, and my results suggest that it can only partially explain the item valence effect. Although the results for Objectives 1, 2, and 3 show the importance of item valence, there are still two important research questions that remain unexplored. The first question is whether my results might be explained by careless responding (Objective 4). The second question is whether the findings have implications for research pertaining to a central construct in I/O psychology, namely job satisfaction. Therefore, I conducted a second study to explore these two issues.

CHAPTER 3 — STUDY 2

The overarching goal of Study 2 is to examine how valence can affect the dimensionality decision of a job satisfaction measure. Recently, Credé et al. (2009) argued that job satisfaction and job dissatisfaction are two separate constructs. As already mentioned in the introduction, the researchers used the CFA technique and found that a two-factor solution fit better than a one-factor solution for job satisfaction. In addition, they also found that job satisfaction scores and job dissatisfaction scores had different nomological networks — job satisfaction scores correlated more strongly with positive affect, OCB and extraversion; job dissatisfaction scores correlated more strongly with negative affect, CWB and neuroticism. I suspect, however, that their results could be partially explained by the valence effect.

My first objective in Study 2 is to examine the factor structure of a job satisfaction measure with a series of factor analytic model comparisons, similar to the comparisons I conducted in Studies 1A and 1B. In their analyses, Credé et al. (2009) did not include a one-factor solution with the item valence effect being controlled. Therefore, I compared three models — a one-factor model, a two-factor model, and a one-factor model with the item valence effect being modelled. Following factor analytic model comparisons, nomological network analyses were conducted to examine whether item valence can influence the correlation of job satisfaction scores and job dissatisfaction scores with other constructs.

My second objective in this study is to examine the research question of how careless responding may affect the factor analytic and nomological network results for a construct. To achieve this purpose I compared the factor analytic results and nomological network results between careful respondents and careless respondents, and I examined the effect of item valence for these two types of respondents. If the valence effect existed among careless but not careful

respondents, it implied that the two-dimensional result of a construct is simply caused by respondents' inattention in completion of survey items. In contrast, if I still found the valence effect among careful respondents, it supported my argument that respondents attend to and answer partly due to the valence of an item.

Study 1 showed that item valence may cause a unidimensional construct to be two-dimensional in factor analyses. Based on these results, my first hypothesis is that item valence may induce job satisfaction items and job dissatisfaction items to load on separate factors in factor analysis. In addition, I expect that this result is perhaps stronger among careless respondents than among careful respondents, because the former group probably does not read and consider the actual item content as carefully as do the latter group. If some careless respondents systematically give the same rating (e.g., strongly agree) to both regular-keyed and reverse-keyed items, the scores of regular-keyed items and reverse-keyed items will not correspond to each other and give rise to a two-factor solution in factor analysis (Schmitt & Stuits, 1985).

Study 1 has demonstrated that construct correlation is inflated when two constructs are of similar valence. In contrast, construct correlation is deflated when two constructs are of opposite valence. Thus, my second hypothesis is as follows: when job satisfaction and dissatisfaction are treated as two separate constructs, job satisfaction will correlate higher with other positive constructs (such as positive affect, positive occupational health, organizational citizenship behaviors). Job dissatisfaction will correlate higher with other negative constructs (such as negative affect, negative occupational health, counterproductive work behaviors). Again, I assume that careful respondents may notice the direction of item keying and thus item valence

more attentively than will careless respondents. Therefore, I expect that careful respondents are less likely to show this differential pattern of correlations compared to careless respondents.

Method

Participants

Employed workers were recruited through StudyResponse, an online respondent recruitment service operated by Syracuse University (2005). StudyResponse sampled participants from a wide variety of industries and backgrounds in the United States, resulting in an enhanced generalizability of my findings. StudyResponse randomly invited 1227 participants into the study through recruitment e-mails. Each participant completed a 20-minute survey in exchange for a remuneration of US$5. A total of 666 workers participated in my survey (response rate = 54.28%), and the mean age of the entire sample was 41.42 (*SD* = 10.78). The ethnic composition is as follow: 78.95% Caucasians, 6.77% Asians/Pacific Islanders, 6.01% Hispanic, 4.81% African Americans, 2.41% Native Americans, 1.20% unidentified.

StudyResponse sent the recruitment e-mail to potential respondents who originally reported themselves in the database as working full-time. I specifically wanted respondents who worked full time because their employment takes a larger part of their time per week, as compared to part-time workers, and the former may experience their jobs more extensively compared to the latter group. Thus, job satisfaction is arguably more important to full-time workers than to part-time workers. I confirmed the respondents' work status at the end of the survey. Indeed, most of them (96.29%) reported themselves as working full-time. Due to the small percentage of part-time workers, however, I included all respondents in the reported analyses.

StudyResponse provided me the demographic characteristics between respondents and non-respondents. There were only slight, albeit statistically significant differences, between the two groups. Compared to non-respondents, respondents were generally 1.51 years older (t[1224] = 2.48, p = .01), more likely to have an undergraduate degree or above (59.31% vs. 53.69%, χ^2[1] = 3.91, p = .05), and more likely to be male (53.88% vs. 40.40%, χ^2[1] = 22.05, p < .001). There was, however, no significant difference in terms of ethnic composition between the respondents and non-respondents (χ^2[5] = 4.62, p = .46).

Measures

All participants completed the following measures. Cronbach's alphas of the scales are shown in Table 13.

Job Satisfaction. Two different job satisfaction measures were used. The first measure was the work satisfaction subscale of the Illinois Job Satisfaction Index (JSI; Chernyshenko, Stark, Crede, Wadlington, & Lee, 2003). The JSI was selected because Credé et al. (2009) discovered evidence of separate dimensions defined by job satisfaction (positively-valenced items) and job dissatisfaction (negatively-valenced items) with each facet of this measure. The JSI originally measured four facets of job satisfaction, namely satisfaction with work, satisfaction with immediate supervisor, satisfaction with coworkers, and satisfaction with pay and benefits. I used only the work satisfaction subscale (15 items) because the inclusion of the entire measure would involve too many items (60 items) and because the work satisfaction facet is more closely related to respondents' general opinions of their jobs. Out of 15 items in the work satisfaction facet, eight are positively valenced (satisfaction items) and seven are negatively valenced (dissatisfaction items). Only three of these negatively-valenced items involve the use of

Table 13

Scale Reliabilities in Study 2

	Total number of items	Full scale	Positively-valenced scale	Negatively-valenced scale
JSI Job Satisfaction	15	.91	.91 (8 items)	.87 (7 items)
BIJS Job Satisfaction total	16	.89	.85 (8 items)	.88 (8 items)
BIJS Work Satisfaction	4	.79	.78 (2 items)	.76 (2 items)
BIJS Supervisor Satisfaction	4	.85	.85 (2 items)	.85 (2 items)
BIJS Coworker Satisfaction	4	.75	.81 (2 items)	.79 (2 items)
BIJS Pay and Benefits Satisfaction	4	.77	.72 (2 items)	.65 (2 items)
Intent to Leave	2	.70	- (1 item)	- (1 item)
Exhaustion	8	.73	.69 (4 items)	.76 (4 items)
OCB	10	.88	.88 (10 items)	- (0 item)
CWB	11	.96	- (0 item)	.96 (11 items)
Balanced Affect	12	.91	.93 (6 items)	.92 (6 items)
Extraversion	14	.85	.81 (7 items)	.88 (7 items)
Neuroticism	14	.82	.76 (7 items)	.89 (7 items)
Average reliability		.83	.82	.84

Note. Cronbach's alphas were estimated in this table. Cronbach's alphas can be estimated when the number of two items is two or more. JSI = Job Satisfaction Index; BIJS = Balanced Inventory of Job Satisfaction; OCB = Organizational Citizenship Behaviors; CWB = Counterproductive Work Behaviors.

the word "not'. Sample items are "I am proud of my work" and "This job is frustrating" (see APPENDIX B).

In addition to the work satisfaction scale by Credé et al. (2009), I created an ad-hoc four-facet job satisfaction scale, which was named Balanced Inventory of Job Satisfaction (BIJS), to use in this study (16 items). The four facets in this new scale were the same four facets originally measured by the JSI, and some of the scale items in the new scale were closely based on the original items in Credé et al. (2009). One noticeable characteristic of BIJS was that it controlled for the content of its positively-valenced items and negatively-valenced items so that they would be extremely similar, and as a result what differed between these items is only their valence. Another important characteristic of this new scale was the avoidance of negation items (i.e., the use of 'not') in the measure. I chose to do this because previous researchers (e.g., Holden & Fekken, 1990) suggested that the use of "negation" words in item wording might cause construct-irrelevant error responding. They argued that negation words would increase the cognitive burden of the respondents because they needed to mentally 'reverse' the meaning of the items before responding. Each facet of BIJS was measured by four items (see APPENDIX B).

The items of both the JSI and BIJS were responded to in one of the most common types of Likert-scale formats (5-point Likert scale, with responses ranging from 1 [*Strongly Disagree*] to 5 [*Strongly Agree*]). The use of a common scale format facilitates the generalizability of my results in situations in which other scales are used.

Turnover Intention. Participants answered two questions related to turnover intention: "I intend to *leave* my organization in the near future" and "I intend to *remain* in my organization for at least the next three years." Participants responded to the items in a 5-point Likert-scale format (1 = *Strongly Disagree*; 5 = *Strongly Agree*).

Exhaustion. The experience of exhaustion is one characteristic of work burnout and was measured in the current study using an 8-item subscale of the Oldenberg Burnout Inventory (Demerouti, Mostert, & Bakker, 2010). Half of the items used were positively valenced and the remaining half were negatively valenced (see Appendix B). Participants responded to the items in a 5-point Likert scale (1 = *Strongly Disagree*; 5 = *Strongly Agree*).

Organizational Citizenship Behaviors (OCB) and Counterproductive Work Behaviors (CWB). I used a short form of the organizational citizenship behavior (OCB) scale and a short form of the counterproductive work behavior (CWB) scale in this study. The long-versions of these scales were originally developed by Spector and colleagues (Fox et al., 2007, for OCB; Spector et al., 2006, for CWB). Later, Spector, Bauer and Fox (2010) developed a 10-item version of OCB and an 10-item version of CWB, and these scales are what I used here. One additional item was added ("Took supplies or tools home without permission") to measure CWB. Similar to most other OCB and CWB measures, all OCB items were positively valenced and all CWB items were negatively valenced. Participants responded to the items in a 5-point Likert scale (1 = *Never*; 5 = *Everyday*).

Positive Affect and Negative Affect. The Scale of Positive and Negative Experience (SPANE) by Diener et al. (2010) was used to measure positive affect (6 items) and negative affect (6 items) in this study. According to Diener et al. (2010), the SPANE has an advantage over the commonly used Positive Affect and Negative Affect Schedule (PANAS; Watson et al., 1988), because the PANAS does not have items that measure low arousal feelings, and because some PANAS items (e.g., 'strong') do not measure emotions. In contrast, the SPANE includes items measuring the whole spectrum of emotion arousal. Participants in this study reported their frequency of experiencing each emotion in a 5-point Likert scale (1 = *Very rarely or Never*; 5 =

Very often or Always). Sample items for positive affect are "joyful" and "happy", and sample items for negative affect are "sad" and "angry". All positive affect items were positively valenced and all negative affect items were negatively valenced. According to Deiner et al. (2010), the positive affect items (SPANE-P) and negative affect items (SPANE-N) can be combined to form a balanced measure of affective well-being (SPANE-B).

Extraversion and Neuroticism. Extraversion and neuroticism were each measured using a short version of the adjective mini-markers for the by Goldberg (1992). This scale can be administered quickly with good reliability and validity (see Goldberg, 1992). Participants answered the extent to which each adjective accurately describes himself or herself in a 5-point Likert scale (1 = *Very Inaccurate*; 5 = *Very Accurate*). Each adjective item used in this scale has another adjective item that is directly antithetical in meaning. Examples of two extraversion items are "silent" and "talkative", and examples of two neuroticism items are "tense" and "relaxed". The use of antithetical items can control for the fact that they assess the same item content, and that only their valence differs. Extraversion and neuroticism were each measured by 14 items. Half of the items were positively-valenced and the remaining half were negatively-valenced.

Check of Careless Responding. Before they began the survey, participants were forewarned that, as an accuracy check, some survey items would require that they select a designated response. In these cases they were asked to follow the instructions and select the identified response alternative. In total there were four accuracy check items embedded throughout the survey. Examples of two of these check items are "Please select strongly agree for this item" and "Choose 'never' for this item". These items were included because respondents who did not pay adequate attention to the content of these items were unable to

choose the correct response alternative. Out of 614 participants (92.19%) who completed my check items, 447 (72.80%) followed the instructions correctly for all of these items and 167 (27.20%) did not follow at least one instruction. The former group was identified as being careful respondents and the latter group was regarded as being careless respondents.

Analysis Strategies

I first conducted a series of confirmatory factor analyses on each of the two job satisfaction measures (i.e., JSI work satisfaction and BIJS). I compared among three models — a one-factor model, a two-factor model, and a one-factor model with valence factors being modelled. After running this analysis with the entire sample, I repeated this analysis for two groups of respondents — careful respondents and careless respondents. It allowed me to examine the magnitude of the valence effect for each group of respondents in the factor analyses. Afterwards, I studied how item valence may affect the nomological networks of job satisfaction scores and job dissatisfaction scores for all respondents. I then repeated the same analyses for careful respondents and careless respondents separately. This allowed me to examine how item valence influenced nomological network analyses for each group of respondents. Finally, I conducted exploratory factor analyses (EFAs) and investigated how item valence can potentially affect the EFAs for all respondents, careful respondents and careless respondents.

Results

I first conducted a confirmatory factor analytic model of job satisfaction measures with and without the valence method factors. The results of these analyses are presented below.

MTMM Factor Analyses on Job Satisfaction

Entire Sample for JSI Work Satisfaction. I first examined the item valence effect in the JSI work satisfaction subscale with the entire respondent sample. In the one-factor model

($M_{baseline}$), all work satisfaction and work dissatisfaction items loaded on one latent construct factor. In the two-factor model ($M_{2factors}$), work satisfaction items loaded on one latent construct factor and work dissatisfaction items loaded on another latent construct factor. The two latent construct factors were allowed to covary with one another. In the valence factor model ($M_{valence}$), all items loaded on the same construct factor and a valence factor (either positive or negative where appropriate). The two valence factors were not allowed to covary with each other. The results are shown in Table 14. The model comparison of $M_{baseline}$ with $M_{2factors}$ or with $M_{valence}$ is based on a chi-square difference test, because $M_{baseline}$ is nested within both $M_{2factors}$ and $M_{valence}$. The model comparison between $M_{2factors}$ and $M_{valence}$ is based on predictive fit indices, namely AIC and SABIC, because these two models are not nested within each other.

For the JSI Work Satisfaction subscale, the one-factor model ($M_{baseline}$) did not fit the data well, as indicated by its poor fit indices. The two-factor model ($M_{2factors}$) was a significant improvement over the one-factor model, as indicated by the chi-square difference test. The job satisfaction factor and the job dissatisfaction factor were moderately correlated (-.56). The best fitting model, however, was the method factor model ($M_{valence}$), suggesting the existence of two method factors based on item valence. The chi-square difference test indicated the superior fit of $M_{valence}$ over $M_{baseline}$, and the predictive fit indices indicate the better fit of $M_{valence}$ compared to $M_{2factors}$. Construct factors accounted for 34.89% of the variance and valence method factors account for 21.27% of the variance. The rest of the percentage is unexplained variance.

Entire Sample for BIJS. I also conducted model comparisons for BIJS. The small number of items (four items only) within each facet of job satisfaction meant that I could not set up a complicated, second-order CFA model in which each item loads on its corresponding facet,

Table 14

Fit Indices for Models Among All Respondents in Study 2

	Model Fit Indices									Model Comparison with $M_{baseline}$		
	χ^2	*df*	*p*	TLI	CFI	RMSEA	SRMR	AIC	SABIC	χ^2 improvement	Δ*df*	*p*
Work Satisfaction Subscale in Job Satisfaction Inventory												
$M_{baseline}$	1584.13	77	<.001	.64	.70	.17	.12	23894.67	23950.24			
$M_{2factors}$	379.37	76	<.001	.93	.93	.08	.05	22691.91	22748.81	1204.76	1	<.001
$M_{valence}$	**273.57**	**63**	**<.001**	**.94**	**.96**	**.07**	**.04**	**22612.10**	**22686.20**	**1310.56**	**14**	**<.001**
Balanced Inventory of Job Satisfaction												
$M_{baseline}$	2753.30	98	<.001	.52	.61	.20	.13	26755.09	26826.54			
$M_{valence}$	**650.12**	**90**	**<.001**	**.89**	**.92**	**.10**	**.09**	**24667.90**	**24749.94**	**2103.18**	**8**	**<.001**

Note. $M_{baseline}$ = Baseline model with construct indicators loaded on a construct; $M_{2factors}$ = Model with a job satisfaction factor and a job dissatisfaction factor; $M_{valence}$ = Model with valence method factor(s). $M_{baseline}$ and $M_{2factors}$ were nested within each other, so were $M_{baseline}$ and $M_{valence}$. This property allows for direct model comparisons with chi-square difference statistics. However, $M_{2factors}$ and $M_{valence}$ can also be compared with predictive fit indices (AIC and SABIC) because they are not nested with each other. The best fit model is boldfaced.

which in turn loads on a higher order factor of job satisfaction, because the model did not converge properly. Instead, I had the items load directly on their corresponding facets, and the facets were allowed to covary with each other.[6] The model comparison process here was similar to the case of the JSI Work Satisfaction subscale. The only exception was that one latent method factor was included in $M_{valence}$ because the model could not converge properly with two valence factors in the current case. The model used here is called the correlated-trait methods minus one model (CT[M-1] model; Eid, Lischetzke, Nussbeck, & Trierweiler, 2003). "Methods minus one" implies that only one method factor is modeled when there are two measurement methods (i.e., the use of positively-valenced items and negatively-valenced items). The CT(M-1) model has an important advantage over my previous model with two method factors because it is computationally less demanding and because it is easier to converge to an interpretable solution. It is particularly useful for estimation of a complicated model, such as in the current situation that involves multiple construct factors[7]. Because only the negatively-valenced method factor was being modeled in this case, the construct factors represented the construct scores that were measured with positively-valenced items. The method factor represented the *unique* variance captured by the negatively-valenced items only.

---

[6] I also considered having the items loaded directly onto a higher-order latent factor ignoring the facet level. However, later in the analyses, the CFA model showed poor fit to the data because the facet level information was discounted. Therefore, the results are not presented.

[7] Failure to extract the two valence factors can be also due to model misspecification, such as the non-existence of positive and negative valence in the data. However, I believe that this explanation is unlikely because other MTMM analyses in my dissertation consistently demonstrated the existence of positive and negative valence factors in the data.

The facet baseline model ($M_{baseline}$) and the valence factor model ($M_{valence}$) were estimated and the results of the model comparison are shown in the bottom section of Table 14. The model comparison was based on the chi-square difference test because $M_{baseline}$ is nested within $M_{valence}$. As expected, the valence factor model ($M_{valence}$) fit significantly better than the baseline model ($M_{baseline}$). The construct effect is estimated to explain 32.52% of the variance and the valence effect is estimated to explain 37.49% of the variance.

One may be surprised by the amount of valence effect in my data. The valence effect explained over one-fifth of the total variance in the JSI Work Satisfaction subscale and accounts for even more variance than the construct effect in the BIJS scale. As such, the data may convince researchers that positively-valenced items substantially differ from negatively-valenced items in the case of job satisfaction. To further explore this issue, I separated the data from careful versus careless respondents and re-conducted the analyses.

Careless versus Careful Respondents for JSI Work Satisfaction. To recall, I classified the carefulness of the respondents into two groups based on their accuracy in answering four check items. I then conducted model comparisons separately for the careful and careless respondents using the JSI Work Satisfaction data. The results are shown in Table 15. For careless respondents, the one-factor model ($M_{baseline}$) fit poorly to the data. The two-factor model ($M_{2factors}$) was a dramatic improvement over the one-factor model ($M_{baseline}$), as indicated by the chi-square difference test. The correlation between the job satisfaction and the job dissatisfaction factors were negligible, $r = -.09$, $p = .25$. However, the best fitting model for the careless respondents was the valence factor model ($M_{valence}$). The chi-square difference test shows its superior fit to the baseline model ($M_{baseline}$) and the predictive fit indices showed its better fit compared to the

Table 15

Fit Indices for Models Among Careful versus Careless Respondents in Study 2

	Model Fit Indices									Model Comparison with $M_{baseline}$		
	χ^2	*df*	*P*	TLI	CFI	RMSEA	SRMR	AIC	SABIC	χ^2 improvement	Δ*df*	*p*
Work Satisfaction Subscale in JSI												
Careless Respondents												
$M_{baseline}$	708.21	77	<.001	.48	.56	.20	.21	7757.42	7763.31			
$M_{2factors}$	216.81	76	<.001	.88	.90	.10	.08	7268.03	7274.05	491.40	1	<.001
$M_{valence}$	**145.20**	**63**	**<.001**	**.92**	**.94**	**.08**	**.05**	**7222.41**	**7230.26**	**563.01**	**14**	**<.001**
Careful Respondents												
$M_{baseline}$	423.49	77	<.001	.89	.91	.10	.05	14999.59	15039.99			
$M_{2factors}$	268.34	76	<.001	.94	.95	.07	.04	14846.43	15024.26	155.25	1	<.001
$M_{valence}$	**231.98**	**63**	**<.001**	**.94**	**.96**	**.08**	**.04**	**14836.08**	**14889.94**	**191.51**	**14**	**<.001**
Balanced Inventory of Job Satisfaction												
Careless Respondents												
$M_{baseline}$	1303.11	98	<.001	.25	.39	.25	.27	9129.07	9136.63			
$M_{valence}$	**259.20**	**90**	**<.001**	**.89**	**.91**	**.10**	**.19**	**8101.16**	**8306.27**	**1043.91**	**8**	**<.001**
Careful Respondents												
$M_{baseline}$	907.14	98	<.001	.80	.84	.13	.05	16026.87	16078.81			
$M_{valence}$	**654.79**	**90**	**<.001**	**.85**	**.89**	**.12**	**.05**	**15790.52**	**15850.15**	**252.35**	**8**	**<.001**

Note. $M_{baseline}$ = Baseline model with construct indicators loaded on a construct; $M_{2factors}$ = Model with a job satisfaction factor and a job dissatisfaction factor; $M_{valence}$ = Model with two valence method factors. $M_{baseline}$ and $M_{2factors}$ were nested within each other, so did $M_{baseline}$ and $M_{valence}$. This property allows for model comparisons with chi-square difference statistics. However, $M_{2factors}$ and $M_{valence}$ can also be compared with predictive fit indices (AIC and SABIC) because they are not nested with each other. The best fit model is boldfaced.

two-factor model ($M_{2factors}$). In the method factor model, the construct factor accounts for merely 11.78% of the variance. In contrast, the valence method factors explain 45.39% of the variance, close to half of the total variance in the data. Therefore, for careless respondents, much of the item variance apparently came from the valence effect.

Interestingly, the picture is dramatically different when looking at the data for the careful respondents. For this group, the one-factor model fit the data acceptably well (TLI was close to .90 and CFI exceeded .90 in fit), and the two-factor model ($M_{2factors}$) fit the data even better than does the one-factor model ($M_{baseline}$). However, for this group the negative correlation between the job satisfaction factor and the job dissatisfaction factor was extremely high (r = -.88, Z = -56.23, p < .001), suggesting that job dissatisfaction is likely the antipode of job satisfaction rather than being a separate dimension. The valence factor model ($M_{valence}$) also showed improvement over both the one-factor model ($M_{baseline}$), as indicated by the significant chi-square difference test) and the two-factor model ($M_{2factors}$, as indicated by the lower predictive fit indices). In addition, the construct factor accounted for 45.49% of the variance and the valence method factors explained only 8.99% of the variance. Therefore, in contrast to the results for careless respondents wherein valence explained nearly half of the variance, for careful respondents valence only accounted for one-fifth of the total variance.

Careless versus Careful Respondents for BIJS. I also conducted separate model comparisons for careless and careful respondents using the BIJS scale data. The results are shown in the bottom part of Table 15. Nested models ($M_{baseline}$ and $M_{valence}$) were compared in this case, and I relied on chi-square difference tests to make the comparisons. For both careless and careful respondents, $M_{valence}$ fit better than $M_{baseline}$. However, the variance decomposition differed dramatically between these two groups. For careless respondents, the construct factors

accounted for much less variance than did the valence factor (15.58% vs. 54.24%). For careful respondents, the situation was reversed — construct factors explained much more variance than did the valence factor (51.58% vs. 11.07%). Therefore, the two dimensional explanation of job satisfaction is much more apparent for careless respondents than for careful respondents, because the method effect was stronger for the former group than for the latter group.

Nomological Network of Job Satisfaction and Job Dissatisfaction

Entire Sample. To examine the effect of item valence on the nomological network investigations in the present study, I first created separate work satisfaction and work dissatisfaction scores (in addition to total scores) from the JSI, and correlated these two scores with other variables. Each of these other variables was measured by both positively- and negatively-valenced items. Therefore, I was able to calculate three different sets of scores: balanced scale scores, positively-valenced scores, and negatively-valenced scores. I expected that the effect of item valence would be strongest when using positively-valenced or negatively-valenced scores, as opposed to using full-scale scores. The results of the analyses are shown in Table 16.

When balanced scale scores were used, work dissatisfaction appeared to have an advantage in correlating with the scores as compared to work satisfaction (see the top panel in Table 16). These results were completely reversed when positively-valenced scores were used. In this case, work satisfaction correlated much stronger with positively-valenced scores than did work dissatisfaction (see the middle panel in Table 16). Interestingly, the pattern was completely reversed when negatively-valenced scores were used. In this case, work dissatisfaction correlated much stronger with negatively-valenced scores than did work satisfaction (see the bottom panel in Table 16). Thus, work dissatisfaction had an advantage in that it correlated better with both

Table 16

Comparing the Correlation Between JSI Work Satisfaction Scores and JSI Work Dissatisfaction Scores with Other Constructs

	Work Satisfaction (S)	Work Dissatisfaction (D)	Absolute Difference	*Z*	
Balanced scale scores					
Extraversion	.295	-.376	.081	-2.17*	(D stronger)
Neuroticism	-.364	.554	.190	-5.54***	(D stronger)
Exhaustion	-.503	.659	.156	-5.16***	(D stronger)
Intent to Leave	-.567	.602	.035	-1.16	
Balanced Affect	.664	-.707	.103	-1.68	
Positively-valenced scores					
Extraversion	.431	-.047	.384	9.93***	(S stronger)
Neuroticism	-.461	.175	.286	7.64***	(S stronger)
Exhaustion	.707	-.306	.401	12.50***	(S stronger)
Intent to Leave	.644	-.349	.295	8.97***	(S stronger)
Positive Affect	.792	-.458	.334	12.28***	(S stronger)
OCB	.347	.201	.146	3.82***	(S stronger)
Negatively-valenced scores					
Extraversion	.086	-.505	.419	-11.11***	(D stronger)
Neuroticism	-.151	.615	.464	-12.99***	(D stronger)
Exhaustion	-.165	.691	.526	-15.39***	(D stronger)
Intent to Leave	-.373	.687	.314	-9.93***	(D stronger)
Negative Affect	-.338	.718	.380	-27.57***	(D stronger)
CWB	.070	.569	.499	-13.44***	(D stronger)

Note. $N = 666$. The parenthesis after a Z-value indicates which score, work satisfaction (S) or work dissatisfaction (D) showed a significantly stronger correlation. JSI = Illinois Job Satisfaction Index; OCB = Organizational Citizenship Behaviors; CWB = Counterproductive Work Behaviors. *$p < .05$; **$p < .01$; ***$p < .001$.

balanced scale scores and negatively-valenced scores. However, the *Z*-values for the negatively-valenced scores were substantially higher than the *Z*-values for the balanced scale scores, suggesting that the valence effect is intensified when an unbalanced, negatively-valenced scale score is used.

Next, I created job satisfaction and job dissatisfaction scores from BIJS and correlated them with balanced scale scores, positively-valenced scores, and negatively-valenced scores (see Table 17). The pattern of results was essentially the same as that described in the aforementioned analysis of the JSI, and thus I will not repeat the description of these results here. Once again, to further examine the role of careless responding on my results, I conducted two sets of analysis comparing the result patterns between careful and careless respondents.

Careless versus Careful Respondents. I examined the role of item valence in nomological network analyses using data from careful and careless respondents. Specifically, I correlated the JSI work satisfaction and work dissatisfaction scores with other variables for the two groups (see Table 18). Similar to the results of all respondents (Table 16), work dissatisfaction had a slight advantage in correlating with balanced scale scores for both careful respondents and careless respondents. This advantage of work dissatisfaction intensified when negatively-valenced scores were used (as indicated by the higher *Z*-values for the negatively-valenced scores) and it was reversed when positively-valenced scores were used. This pattern held true for both careless and careful respondents. When examining the absolute difference in the correlation coefficients between work satisfaction and dissatisfaction, I found that the differences were generally much larger for careless respondents as compared to careful respondents. Therefore, although in this case the valence effect was discovered among both careful and careless respondents, it was much higher for the latter group.

Table 17

Comparing the Correlation Between Job Satisfaction Scores and Job Dissatisfaction Scores in the Balanced Inventory of Job Satisfaction with Other Constructs

	Job Satisfaction (S)	Job Dissatisfaction (D)	Absolute Difference	*Z*	
Balanced scale scores					
Extraversion	.247	-.414	.167	-4.35***	(D stronger)
Neuroticism	-.344	.558	.214	-6.04***	(D stronger)
Exhaustion	-.499	.532	.033	-0.98	
Intent to Leave	-.521	.593	.072	-2.24*	(D stronger)
Balanced Affect	.615	-.670	.055	1.94	
Positively-valenced scores					
Extraversion	.399	-.041	.358	8.93***	(S stronger)
Neuroticism	-.458	.118	.340	8.71***	(S stronger)
Vigor	.702	-.213	.489	14.27***	(S stronger)
Intent to Stay	.595	-.324	.271	7.73***	(S stronger)
Positive Affect	.729	-.381	.348	11.15***	(S stronger)
OCB	.310	.241	.069	1.75	
Negatively-valenced scores					
Extraversion	.042	-.566	.524	-13.62***	(D stronger)
Neuroticism	-.128	.661	.533	-14.75***	(D stronger)
Exhaustion	-.162	.583	.421	-11.34***	(D stronger)
Intent to Leave	-.339	.694	.352	-10.83***	(D stronger)
Negative Affect	-.315	.730	.415	-12.90***	(D stronger)
CWB	.070	.657	.587	-15.94***	(D stronger)

Note. $N = 666$. The parenthesis after a Z-value indicates which score, work satisfaction (S) or work dissatisfaction (D) showed a significantly stronger correlation. *$p < .05$; **$p < .01$; ***$p < .001$.

Table 18

Comparing the Nomological Network of JSI Work Satisfaction and Work Dissatisfaction Among Careless and Careful Respondents

	Careless Respondents				Careful Respondents			
	Work Satisfaction (S)	Work Dissatisfaction (D)	Abs. Diff.	*Z*	Work Satisfaction (S)	Work Dissatisfaction (D)	Abs. Diff.	*Z*
Balanced scale scores								
Extraversion	.048	-.325	.277	-2.71** (D)	.393	-.353	.040	0.90
Neuroticism	-.093	.424	.331	-3.33*** (D)	-.493	.560	.067	-1.69
Exhaustion	-.448	.513	.065	-0.76	-.565	.719	.154	-4.58*** (D)
Intent to Leave	-.458	.538	.080	-0.95	-.641	.600	.041	1.17
Balanced Affect	.516	-.555	.039	0.49	.779	-.761	.018	0.74
Positively-valenced scores								
Extraversion	.421	.268	.153	1.60	.420	-.262	.158	3.48*** (S)
Neuroticism	-.449	-.194	.255	2.65** (S)	-.446	.417	.029	0.67
Vigor	.771	.150	.621	7.72*** (S)	.680	-.632	.048	1.45
Intent to Stay	.674	-.076	.598	6.66*** (S)	.629	-.506	.123	3.26*** (S)
Positive Affect	.776	-.156	.620	7.77*** (S)	.799	-.668	.131	4.81*** (S)
OCB	.570	.314	.256	2.91** (S)	.212	-.019	.193	3.96*** (S)
Negatively-valenced scores								
Extraversion	-.254	-.531	.277	-3.02** (D)	.317	-.386	.069	-1.53
Neuroticism	.228	.571	.343	-3.78*** (D)	-.431	.561	.130	-3.20*** (D)
Exhaustion	.167	.744	.577	-7.02*** (D)	-.356	.629	.273	-6.75*** (D)
Intent to Leave	-.083	.710	.627	-7.18*** (D)	-.572	.613	.041	-1.12
Negative Affect	-.056	.618	.562	-6.03*** (D)	-.616	.731	.115	-3.62*** (D)
CWB	.375	.550	.175	-2.02* (D)	-.191	.441	.250	-5.45*** (D)

Note. N = 167 for careless respondents and 447 for careful respondents. The parenthesis after a Z-value indicates which score, work satisfaction (S) or work dissatisfaction (D) showed a significantly stronger correlation. Abs. Diff. = Difference in the Absolute Value of the Correlation Coefficients

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

A similar pattern of results was found when the job satisfaction-dissatisfaction scores were created with the BIJS scale, as shown in Table 19. Job satisfaction tended to correlate better with other scores when these scores were also positively valenced; job dissatisfaction tended to correlate with other scores when these scores were also negatively valenced. These results showed that participants responded to item valence, even after excluding careless respondents from the analysis (although the effect was larger in the careless respondent group).

Past studies on the effect of careless responding on factor analyses have relied heavily on simulated data rather than using data from actual respondents (Schmitt and Stuits, 1985; Wood, 2006), and those studies found that a construct is likely to show a two-factor solution as long as at least 10% of the respondents are responding carelessly. The current research advanced upon this previous research in several important ways. First, I used data from real respondents from a worker sample rather than hypothetical respondents from a computer simulation, and this enhances the generalizability of my results. In doing so, I discovered that an alarming number of my respondents (close to 30%) did not answer my survey carefully. This result occurred even after respondents received forewarning regarding the existence of items checking their attention to my survey. Second, and even more important, this impact of careless responding on item valence is even larger than the estimated effect of common method variance on self-report surveys (around 25%; Williams, Cote, & Buckley, 1989). Careless respondents showed over 36% more variance that is due to item valence than did careful respondents (45.49% vs. 8.99% for JSI Work Satisfaction data; 54.24% vs. 11.07% for BIJS data). Finally, these results show that careless responding affects results in not only factor analysis but also in nomological network analysis. Although item valence affects both careless and careful respondents in nomological network analyses, the effect is more pronounced for the careless respondents.

Table 19

Comparing the Nomological Network of Job Satisfaction and Job Dissatisfaction Scores in the Balanced Inventory of Job Satisfaction Among Careless and Careful Respondents

	Careless Respondents				Careful Respondents			
	Job Satisfaction (S)	Job Dissatisfaction (D)	Abs. Diff.	*Z*	Job Satisfaction (S)	Job Dissatisfaction (D)	Abs. Diff.	*Z*
Balanced scale scores								
Extraversion	.012	-.421	.409	-8.34*** (D)	.341	-.388	.047	-1.67
Neuroticism	-.095	.505	.410	-8.61*** (D)	-.459	.548	.089	-3.44*** (D)
Exhaustion	-.416	.420	.004	-0.09	-.566	.582	.016	-0.66
Intent to Leave	-.412	.544	.132	-3.04** (D)	-.599	.603	.004	-0.17
Balanced Affect	.477	-.615	.138	1.75	.719	-.690	.029	1.41
Positively-valenced scores								
Extraversion	.436	.230	.206	2.15* (S)	.362	-.287	.075	2.61** (S)
Neuroticism	-.474	-.210	.264	2.77** (S)	-.413	.382	.031	1.12
Vigor	.761	.123	.638	7.76*** (S)	.665	-.533	.132	5.59*** (S)
Intent to Stay	.654	-.091	.563	6.22*** (S)	.567	-.498	.069	2.73** (S)
Positive Affect	.750	-.149	.581	7.30*** (S)	.717	-.599	.118	5.38*** (S)
OCB	.565	.306	.259	2.92** (S)	.139	.007	.132	4.33*** (S)
Negatively-valenced scores								
Extraversion	-.302	-.603	.301	-3.47*** (D)	.276	-.425	.149	-5.25*** (D)
Neuroticism	.244	.663	.419	-4.91*** (D)	-.402	.569	.167	-6.38*** (D)
Exhaustion	.196	.609	.413	-4.59*** (D)	-.369	.493	.124	-4.56*** (D)
Intent to Leave	-.034	.705	.671	-7.54*** (D)	-.553	.625	.072	-3.01** (D)
Negative Affect	-.025	.705	.680	-7.62*** (D)	-.593	.670	.077	-3.38*** (D)
CWB	.381	.627	.246	-2.99** (D)	-.171	.493	.322	-11.27*** (D)

Note. $N = 167$ for careless respondents and 447 for careful respondents. The parenthesis after a Z-value indicates which score, work satisfaction (S) or work dissatisfaction (D) showed a significantly stronger correlation. Abs. Diff. = Difference in the Absolute Value of the Correlation Coefficients
*$p < .05$; **$p < .01$; ***$p < .001$.

Effect of Careful versus Careless Responding on Exploratory Factor Analyses

As a final analysis, I examined the effect of careful versus careless responding on exploratory factor analyses, with principal axis factoring as the extraction method and direct oblimin as the rotation method. I used scree tests to decide the number of factors to be extracted in each case. For the JSI Work Satisfaction Scale, careless respondents showed a two-factor solution and careful respondents showed a one-factor solution (Figure 4). Table 20 shows the factor loadings of the items. For careless respondents, a one-factor solution was apparently inferior to a two-factor solution in the factor loading table because some job dissatisfaction items did not load on the one-factor solution (Items 11-13; |loadings| < .30). In contrast, for careful respondents, a one-factor solution appeared to fit the data apparently as well as a two-factor solution. In addition, both job satisfaction and dissatisfaction items appeared to load on the one-factor solution better among careful respondents than among careless respondents.

For BIJS, careful respondents showed a six-factor solution and careless respondents showed anywhere from three to perhaps a six-factor solution (Figure 5). I thus chose to compare the 3-factor, 4-factor, and 6-factor solutions in each case. The results are shown in Table 21 for careless respondents and Table 22 for careful respondents. For the 3-factor solution for careless respondents, job dissatisfaction items tended to load on the first factor and job satisfaction items tended to load on the second factor. The third factor was mainly occupied by pay and benefit satisfaction items. Overall, positive valence and negative valence occupied two of the three factors. A similar situation occurred for the four-factor solution for the careless respondents. In this case the first factor corresponded to the job satisfaction items, the second factor corresponded to the job dissatisfaction items, the third factor corresponded to pay and benefits satisfaction items, and the fourth factor was loaded by two supervisor satisfaction items. Again,

Table 20

Factor Loadings for JSI Work Satisfaction Scale

	Careless Respondents			Careful Respondents		
	One-Factor	Two-Factor		One-Factor	Two-Factor	
		First Factor	Second Factor		First Factor	Second Factor
Item 1	**.76**	**.76**	-.06	**.73**	**.85**	.08
Item 2	**.75**	**.76**	-.03	**.60**	**.71**	.08
Item 3	**.84**	**.85**	-.02	**.81**	**.66**	-.24
Item 4	**.76**	**.81**	.06	**.85**	**.70**	-.24
Item 5	**.85**	**.87**	-.03	**.85**	**.78**	-.16
Item 6	**.80**	**.85**	.07	**.80**	**.80**	-.07
Item 7	**.72**	**.70**	-.14	**.69**	**.85**	.13
Item 8	**.76**	**.80**	.06	**.73**	.58	-.24
Item 9*	-.37	-.12	**.81**	**-.86**	-.49	.50
Item 10*	-.30	-.05	**.78**	**-.75**	-.27	**.62**
Item 11*	-.13	.08	**.65**	-.49	-.01	**.60**
Item 12*	-.27	-.05	**.69**	-.58	-.11	**.60**
Item 13*	.10	.31	**.63**	-.49	.21	**.85**
Item 14*	-.35	-.12	**.74**	**-.71**	-.28	.56
Item 15*	-.34	-.10	**.77**	**-.74**	-.30	.58

Note. * = work dissatisfaction item. Principal Axis Factoring Analysis with Direct Oblimin Rotation was conducted. Factor loadings at or above .30 are underlined. Factor loadings at or above .60 are underlined and bolded.

Table 21

Factor Loadings for Balanced Inventory of Job Satisfaction for Careless Respondents

	3-factor solution			4-factor solution				6-factor solution					
	1st Factor	2nd Factor	3rd Factor	1st Factor	2nd Factor	3rd Factor	4th Factor	1st Factor	2nd Factor	3rd Factor	4th Factor	5th Factor	6th Factor
WS1	.00	**.80**	-.09	.03	**.72**	-.11	-.12	-.18	.11	-.10	**.62**	.18	-.11
WS2	-.06	**.77**	-.11	-.06	**.70**	.09	-.11	-.12	-.01	.07	**.67**	.20	-.24
WS3*	**.73**	-.13	.31	**.65**	-.16	.36	-.06	.06	.09	**.50**	-.16	.17	.33
WS4*	**.74**	-.10	.02	**.73**	-.04	.06	.07	-.01	.03	.03	-.08	.02	**.84**
SS1	-.08	**.83**	.05	-.13	.51	.05	-.44	**-.75**	-.02	-.02	.20	.05	.12
SS2	-.07	**.82**	.01	-.18	.34	.03	**-.69**	**-.78**	**.20**	**.12**	.09	.05	-.09
SS3*	**.86**	-.11	-.07	**.89**	.00	-.03	.14	.35	.26	.16	.08	.30	.36
SS4*	**.79**	-.14	.11	**.80**	.03	.14	.21	.38	.07	.22	.14	.32	.34
CS1	-.06	**.81**	.08	-.03	**.88**	.06	.05	-.08	.17	.13	**.76**	-.24	.01
CS2	-.12	**.76**	.02	-.09	**.80**	.00	.03	-.22	-.04	-.10	**.65**	-.17	.18
CS3*	**.74**	-.02	.03	**.68**	-.20	.08	-.25	-.09	.09	.17	-.21	**.61**	.18
CS4*	**.86**	.03	-.02	**.83**	-.06	.03	-.12	.09	.33	.28	-.05	.33	.29
PBS1	.42	.48	-.39	.45	.30	-.38	-.25	-.15	**.74**	.01	.13	.01	.05
PBS2	.36	.48	-.51	.42	.37	-.51	-.16	-.13	.29	-.44	.30	.35	.18
PBS3*	.38	.09	**.64**	.29	.12	**.68**	.02	-.01	-.37	.52	.14	.23	.22
PBS4*	.38	.15	.54	.28	.09	.58	-.09	-.03	.06	**.77**	.04	.00	.00

Note. * = dissatisfaction item. WS = Work Satisfaction; SS = Supervisor Satisfaction; CS = Coworker Satisfaction; PBS = Pay and Benefits Satisfaction. Principal Axis Factoring Analysis with Direct Oblimin Rotation was conducted. Factor loadings at or above .30 are underlined. Factor loadings at or above .60 are underlined and bolded.

Table 22

Factor Loadings for Balanced Inventory of Job Satisfaction for Careful Respondents

	3-factor solution			4-factor solution				6-factor solution					
	1st Factor	2nd Factor	3rd Factor	1st Factor	2nd Factor	3rd Factor	4th Factor	1st Factor	2nd Factor	3rd Factor	4th Factor	5th Factor	6th Factor
WS1	**.62**	.27	.04	.01	.03	.04	**.82**	.07	.06	**.82**	.05	.00	.17
WS2	.55	.19	-.01	-.02	-.05	.00	**.77**	.02	-.02	**.73**	.02	.00	.14
WS3*	**-.62**	-.25	-.09	-.07	-.05	-.11	**-.72**	-.15	-.07	**-.69**	-.06	.02	.13
WS4*	-.52	-.19	.08	-.05	-.03	.06	**-.60**	.04	-.01	**-.60**	-.01	.06	.27
SS1	-.02	.01	**.89**	.04	.01	**.88**	-.05	**.85**	.01	-.01	.06	-.02	.09
SS2	-.06	.07	**.86**	-.06	.03	**.86**	.02	**.85**	.02	.04	-.02	-.03	.21
SS3*	-.11	.04	**-.75**	-.14	.02	**-.75**	.03	**-.78**	.00	.03	-.08	.00	.32
SS4*	-.05	-.06	**-.86**	.03	.00	**-.88**	-.09	**-.84**	-.01	-.11	.02	.02	.07
CS1	**.67**	-.16	.09	**.76**	-.03	.03	.00	.04	.04	.02	**.81**	.07	.22
CS2	**.69**	.10	.06	**.79**	.04	-.01	.00	.01	-.01	.03	**.79**	-.06	.14
CS3*	**-.68**	.11	-.05	**-.73**	-.02	.00	-.03	-.03	-.02	-.05	**-.66**	.03	.24
CS4*	**-.61**	.11	-.07	**-.62**	.01	-.04	-.05	-.06	.02	-.05	-.56	.03	.26
PBS1	-.04	**.78**	.02	-.01	**.83**	-.03	-.04	.01	**.88**	.02	-.03	-.04	.16
PBS2	.02	**.75**	.05	-.05	**.72**	.03	.09	-.04	.04	.01	.03	**-.91**	.16
PBS3*	-.05	**-.73**	-.01	-.11	**-.83**	.05	.07	.01	**-.92**	.01	-.03	.01	.16
PBS4*	-.02	**-.67**	-.11	.05	**-.63**	-.10	-.09	-.06	.00	.00	.02	**.85**	.12

Note. * = dissatisfaction item. WS = Work Satisfaction; SS = Supervisor Satisfaction; CS = Coworker Satisfaction; PBS = Pay and Benefits Satisfaction. Principal Axis Factoring Analysis with Direct Oblimin Rotation was conducted. Factor loadings at or above .30 are underlined. Factor loadings at or above .60 are underlined and bolded.

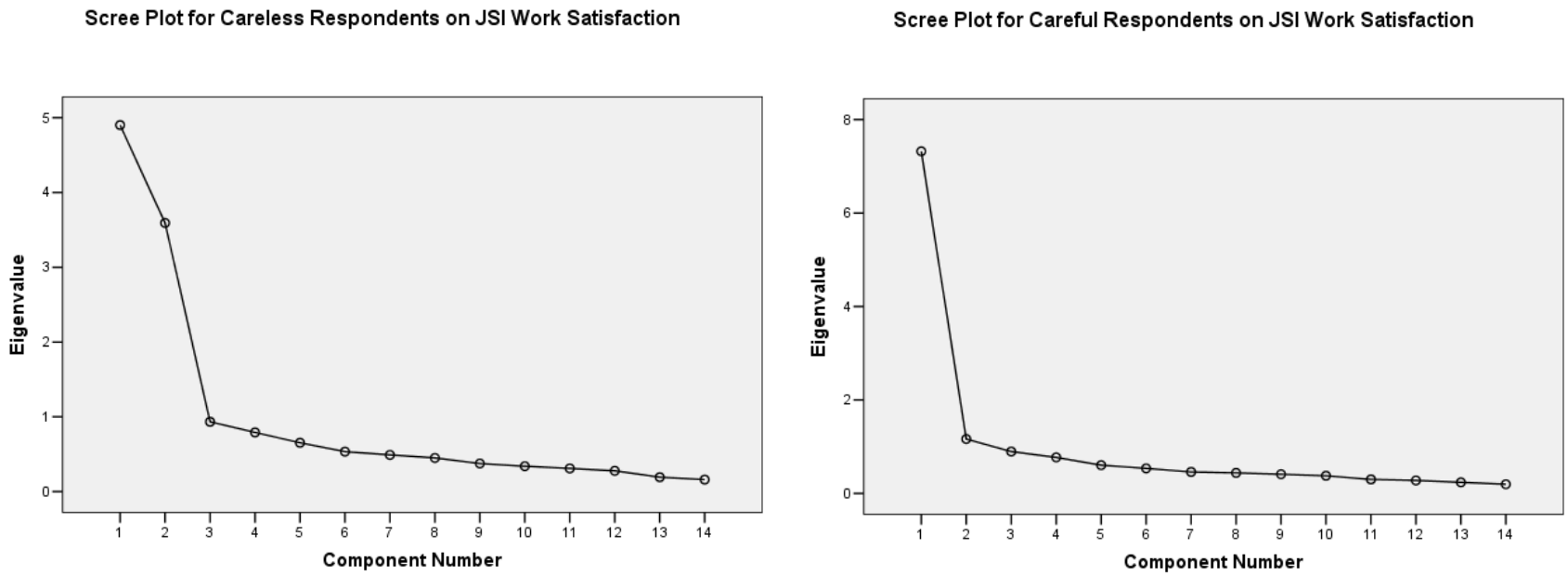

Figure 4. Plots of eigenvalues for the JSI Work Satisfaction scale in Study 2.

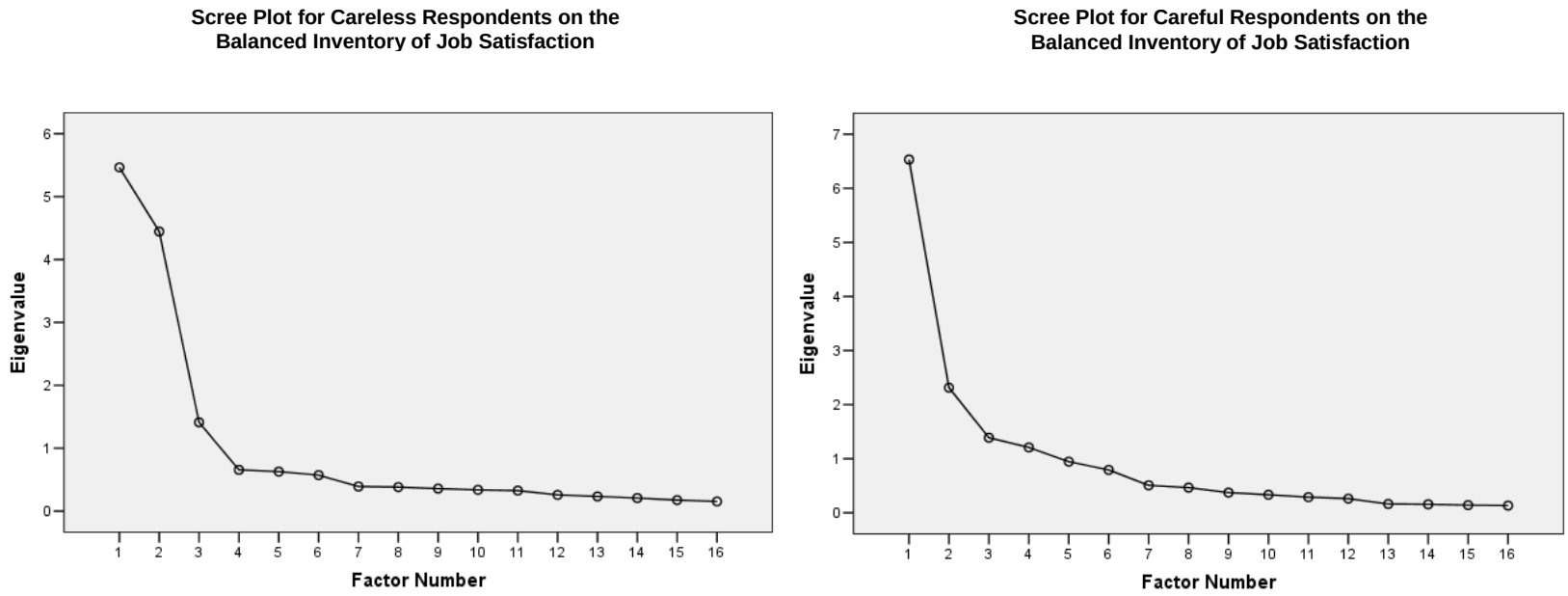

Figure 5. Plots of eigenvalues for the Balanced Inventory of Job Satisfaction in Study 2.

positive valence and negative valence occupied the first two of these factors. The six-factor solution for careless respondents was not interpretable.

In contrast to the data for careless respondents, the data from careful respondents showed a pattern of factor loadings that was consistent with item content rather than with item valence. For careful respondents, a three-factor solution showed that all four work satisfaction items and all four coworker satisfaction items loaded on the first factor, that all four pay and benefit satisfaction items loaded on the second factor, and that all supervisor satisfaction items loaded on the third factor. A four-factor solution showed that all four work satisfaction items, all four supervisor satisfaction items, all four coworker satisfaction items, and all four pay and benefit satisfaction items loaded on distinct factors. The 6-factor solution was similar to the 4-factor solution, except that pay and benefit satisfaction items were further split into two factors (i.e., one pay satisfaction and one pay dissatisfaction item load on one factor; one benefit satisfaction and one benefit dissatisfaction item load on another factor). In addition, one supervisor satisfaction item loaded on the supervisor satisfaction factor and weakly by itself (loading = .32). Overall, these results demonstrated that the item valence effect was stronger for careless respondents as compared to careful respondents. Although my MTMM analyses showed that the valence effect still exists for careful respondents, it does not appear to strongly affect the exploratory factor analytic (EFA) results.

Summary

Study 2 extended the major findings in Study 1. The first hypothesis of Study 2 was that item valence may induce job satisfaction items and job dissatisfaction items to load on separate factors in factor analyses. This hypothesis was supported. The second hypothesis of Study 2 was that job satisfaction will correlate stronger with other positively-valenced constructs (such as organizational citizenship behaviors) and that job dissatisfaction will

correlate stronger with other negatively-valenced constructs (such as counterproductive work behaviors). This hypothesis was also supported. My research suggested that item valence affects not only personality and social psychological constructs (in Study 1) but also an important industrial-organizational construct (i.e., job satisfaction in Study 2). The valence effect was found in both a student sample in Study 1 and a worker sample in Study 2.

Recall that objective 4 of the current dissertation was to examine whether the valence effect can be fully explained by careless responding. The corresponding hypothesis was that the effect of item valence is stronger for careless respondents than for careful respondents. This hypothesis was supported. My results demonstrated that the valence effect exists for both careful and careless respondents, although it is much stronger for the latter group. Interestingly, when I examined the results for careful respondents, the valence effect was found in both the confirmatory factor analysis and nomological network analysis but not the exploratory factor analysis. The former two analyses are perhaps more sensitive to item variance due to valence than is exploratory factor analysis.

My results do not lend strong support to the conclusion that job satisfaction is a two-factor construct. Among careful respondents, job dissatisfaction (as compared to job satisfaction) only showed a slight advantage in correlating with full-scale variables. In addition, these results were not replicated consistently across two job satisfaction measures. For example, dissatisfaction correlated higher with exhaustion than did satisfaction only in the JSI work satisfaction scale but not in the BIJS scale, and dissatisfaction correlated stronger with neuroticism only in the BIJS scale but not in the JSI work satisfaction scale. Furthermore, when I considered only careful respondents in my data, I did not replicate the findings by Credé et al. (2009) that job satisfaction and dissatisfaction showed differential nomological networks with extraversion or intention to leave. Therefore, researchers should

seriously consider the empirical value of treating job satisfaction and dissatisfaction separately.

CHAPTER 4 — GENERAL DISCUSSION

The main purpose of my dissertation was to examine the potential effect of item valence on the correlations among both items and variables. Item valence refers to the evaluative aspect of an item. In other words, it is the value of the attributes or the behaviors (as described in a survey item) to the respondent (Beauvois & Dubois, 2000; Hofstee, 1990; Peabody, 1967). For instance, when a survey respondent answers an item "How kind are you?", the respondent is not simply making an objective description of himself or herself but also an evaluation of himself or herself. In Study 1, I demonstrated the existence of valence effects in common psychological measures that I/O psychologists often use (e.g., extraversion, conscientiousness). In Study 2, I examined these same effects as they apply to another popular construct in I/O psychology (i.e., job satisfaction). In both Study 1 and 2, I addressed the questions of (a) whether the valence effect is simply the result of social desirability responding or careless responding and (b) how careless responding can affect the dimensionality of a specific construct in organizational psychology.

It is important to clarify the goal of the present research. It was not my intention to question or confirm the dimensionality of any particular constructs that I included in my empirical investigation. Rather, the main purpose of my MTMM analyses was to demonstrate that items of the same valence *within* a single measure and *across* different measures correlate more strongly with each other because they load on the same latent factor (which I labelled valence method factors). Similarly, the main purpose of my nomological network investigation was to show that construct scores will correlate more strongly together when they share the same valence. The overarching goal of the current dissertation is to demonstrate that item valence can bias researchers' interpretation of factor analytic results and nomological network results, *suggesting more dimensions than the actual dimensionality of the construct under investigation*.

Item valence can induce a unidimensional construct to appear two-dimensional or cause a two-dimensional construct to appear even more distinct in both types of analyses.

My dissertation contributed to the literature on construct dimensionality in at least five significant ways. First, my model comparison analyses in Study 1 and 2 demonstrated that item valence has a strong potential to influence item correlations across a variety of measures commonly used in psychology. Item keying direction matters mainly because items measuring opposite poles of a construct also differ in item valence. Item valence is thus an alternative explanation for evidence that seemingly bipolar constructs split into two factors in factor analyses. Previous researchers have taken this finding of two factors as evidence that the psychological instrument under investigation reflect two distinct constructs, but my research shows that item valence often contributes sufficient variance to cause a two-factor solution. Second, even more important, I showed that valence can affect the magnitude of the correlations among constructs in general. This novel finding has extremely important implications for the interpretation of nomological network analyses. If, as is often the case (e.g., Credé et al., 2009), the variables included in these analyses are measured with uniformly-valenced items, differences in the pattern of correlations cannot be unambiguously interpreted as evidence for the distinction between focal constructs (e.g., between introversion and extraversion). Third, I demonstrated that the observed effects of valence cannot be fully explained by social desirability response bias and therefore should be addressed as a separate issue in measure development and evaluation. Fourth, as shown in Study 2, even though the item valence effect is exaggerated with careless responding style, this response style cannot entirely explain the item valence effect. My findings thus suggest that there is a certain inherent nature within a participant to respond to the valence in addition to the content of an item. Finally, my results suggest that positively-valenced and negatively-valenced items are most likely to load on two separate factors among careless respondents

compared to careful respondents. Careless responding may thus bias the factor analytic and nomological network analytic results of an investigation.

One might question why valence can bias both factor analytic and nomological network investigations. Indeed, both analyses are correlation- or covariance-based techniques. In factor analyses, scale items that share the same valence correlate more strongly with each other and thus load on the same factor. Because construct scores are typically measured as aggregates of such measurement items, similarly-valenced constructs will also correlate more highly with each other than they do with oppositely-valenced constructs in nomological network analysis. Factor analysis is not identical to nomological network investigations. However, when items of opposite valence load on separate factors, the two factors are also likely to show differential correlations with other variables based on construct valence.

Practical Suggestions to Improve Construct Dimensionality Research

Understanding the dimensionality of psychological constructs like self-esteem and job satisfaction is important. To date, both academic and applied researchers in industrial-organizational psychology have commonly relied on EFA or simple CFA models (e.g., a two-factor CFA model) to guide their decision regarding the dimensionality of a construct. In my dissertation I have demonstrated an alternative analytic strategy that could be applied early in the measure development process. Marsh, Scalas, and Nagengast (2010) recently suggested that the use of appropriate quantitative techniques often comes at the cost of understanding how to conduct the analysis. However, my discussion shows that many of these techniques are manageable to most researchers. Below I will devote more attention to nomological network analyses than to factor analyses given that researchers appear to be less familiar with the role of valence in the former.

Factor analysis. Factor analysis is an extremely useful tool to evaluate item-level data in terms of a few number of dimensions. When used to investigate a construct's

dimensionality, this statistical technique is based on the assumption that respondents will give similar answers to items that measure a common underlying factor (e.g., extraversion), but this analysis does not account for other non-substantive factors that affect item correlations (e.g., restriction of range, item distribution properties, and item extremity; see Bernstein, Garbin, & Teng, 1988; Guilford & Fruchter, 1978; McPherson & Mohr, 2005). The current research has demonstrated the prevalence of item valence among psychological instruments and shows its effect on research conclusions that rely primarily on simple factor analytic models (including EFA or basic CFA models without method factors). A more advanced set of techniques, MTMM CFAs, is surprisingly underutilized (Marsh et al., 2010). MTMM CFAs are simple extensions to the common CFA model and can be used in popular structural equation modeling (SEM) programs (e.g., Amos, Lisrel, EQS, Mplus) to evaluate dimensionality. I strongly recommend that researchers use this technique to check for the valence effect in their measurement instruments (refer to Eid and Diener [2006] for more detail).

Nomological network analysis. To ensure a fair comparison in correlation coefficients, researchers might consider using measures with a balanced set of opposite-valenced items. My empirical results demonstrate that a measurement instrument with a balanced set of positively-valenced and negatively-valenced items is least likely to show differential correlations with the opposite pole of a construct (e.g., extraversion vis-a-vis introversion). In practice, I realize that it is difficult to always use measures with a balanced set of oppositely-valenced items because many psychological scales consist of predominantly regular-keyed measurement items. One solution to this problem is to re-weight the scale items so the positively- and negatively-valenced items have the same overall contribution to a construct's final score. For instance, if extraversion is measured by six positively- and three negatively-valenced items, researchers can give twice as much weight to the negatively-

valenced items. However, this method cannot be implemented on measures without any positively-valenced or negatively-valenced items.

Some researchers may question whether a measurement instrument has to be exactly balanced (i.e., including the same number of regular-keyed and reverse-keyed items, or reweighting the overall contribution of positively- and negatively-valenced items) in order to minimize the valence effect that is discussed in the current paper. As a quick investigation into this question, I reweighted the overall contribution of regular-keyed items versus reverse-keyed items on the calculation of a construct score (e.g,. agreeableness, openness to experience, SDO) in Sample 1A to the following ratios: 50/50, 60/40, 70/30, 80/20, 90/10 and 100/0). I then looked at how the reweighted values correlated with extraversion, which is measured entirely by regular-keyed items. I was surprised by the possible impact of valence effect in these results. The correlation of extraversion with agreeableness rose substantially as I stepped up the weight of the regularly-keyed agreeableness items (.17 at 50/50, .22 at 60/40, .26 at 70/30, .30 at 80/20, .34 at 90/10, and .36 at 100/0). In this case, having one or two reverse-keyed items (out of a total of ten items) corresponds to 90/10 or 80/20 usage of regular-keyed items, respectively, and it is unlikely that this would help to eliminate the valence effect. Valence also caused a noticeable difference in the correlation coefficients associated with Openness to Experience (.17 at 50/50, .19 at 60/40, .21 at 70/30, .22 at 80/20, .24 at 90/10, and .25 at 100/0). In addition, I found that the effect that valence had on correlation coefficients ranged from virtually no impact (for BDW and self-esteem) to a glaring impact (for agreeableness). Thus, because researchers rarely have knowledge of how item valence will affect correlation coefficients before they begin data collection, I believe that researchers should be conservative by employing a balanced set of oppositely-valenced items in their nomological network studies.

Another way to prevent the item valence effect in studies of construct dimensions, according to Campbell and Fiske (1959), is to maximize method heterogeneity among measures used in a nomological network analysis. For example, to test for the dimensionality of job satisfaction, in addition to using a self-report measure of stress or well-being as potential correlates, researchers could include objective measures in the analysis. Physiological measures such as heart rate or cortisol level are not subject to the valence effect. Scores based on observable behaviors such as employees' absenteeism are also not prone to this bias. If differences in the correlations are observed for the self-report measures but not the physiological or behavioral measures, it would suggest that the former might be due to valence effects. Only in cases where differences are consistently observed in measures uncontaminated by the valence effect can a strong case be made for the bi-dimensionality of the focal construct.

A long-term solution to the item valence problem is to formulate items that are low in evaluative content (Bäckström, Björklund, & Larsson, 2009; Jackson, 1984), in addition to creating measurement instruments that contain a balanced set of oppositely-valenced items. Recently, a notable study by Bäckström et al. (2009) discovered that items in Big Five personality inventories are saturated with evaluative content that causes these personality factors to correlate with each other even though they are theoretically orthogonal. When these researchers minimized item valence by reframing the personality items to be more neutral in meaning, the inter-correlations among personality factors were substantially weaker, although they did not disappear entirely. These results suggest that the common variance among personality factors comes partially from participants' sensitivity to item valence. Their research did not investigate the ramifications of item valence on the uni- versus bi-dimensional debate of a construct as was done in the present dissertation. However, the implications of Bäckström et al.'s (2009) findings can easily apply here because items or

constructs may show inflated correlation with each other because they share the same valence. Therefore, by minimizing item valence during the scale development process, researchers can attenuate the inflated variance that is common among the items and among the constructs.

As shown in Study 1, both extraversion and introversion scores and positive and negative self-esteem scores showed differential correlations with a few full-scale balanced scores, even though there was no a priori theoretical explanations for these findings. I believe that these unexpected results might be due to non-substantive factors. One possible factor is the idiosyncratic characteristics of measurement instruments. When two scales share some common characteristics such as similar item wordings and rating scales, participants may respond to items in these two scales similarly regardless of the actual content in these scales. Another potential explanation is sampling errors. Sampling characteristics fluctuate from one study to another and thus the correlation results also fluctuate randomly from sample to sample (see Schmidt, 1992). Finally, other factors (e.g., range restriction of a score, scores' distributional characteristics of two scores, and item extremity) can also affect the magnitude of a correlation coefficient (Bernstein et al., 1988; Guilford & Fruchter, 1978; McPherson & Mohr, 2005). Therefore, it is more important to look at the overall pattern of findings rather than a particular correlation comparison, because any particular correlation comparison is too sensitive to the influence of any one non-substantive, construct-irrelevant factor (e.g., sampling errors).

Use of other strategies. The evidence provided in my dissertation shows that valence effects can contribute to evidence for bi-dimensionality. It does not preclude the possibility that a construct under a dimensionality debate is truly two dimensional. However, it is difficult to test this unambiguously with factor analysis and nomological network analyses because both analyses are correlation-based techniques that are subject to the item valence

effect described in the current article. Researchers should consider other strategies. These strategies include experimental techniques and group comparison techniques (Borsboom & Mellenbergh, 2007).

In *experimental techniques*, researchers use a controlled setting to manipulate two focal constructs in distinct experimental conditions (e.g., job satisfaction in one experimental condition and job dissatisfaction in another experimental condition). For example, respondents may be subliminally primed with the word job satisfaction or job dissatisfaction before they answer survey items. Non-equivalent impact of an experimental manipulation on job satisfaction items and job dissatisfaction items will strengthen evidence for the bi-dimensionality of a construct. With this technique, however, a researcher needs to be careful about how to manipulate the focal construct (job satisfaction and dissatisfaction) without also manipulating the valence (positive and negative). One convenient way to check this is to include positively-valenced items and negatively-valenced items of other constructs, such as optimism items and pessimism items. If a manipulation has an effect on the focal construct (job satisfaction or job dissatisfaction) but not the other constructs (e.g., optimism or pessimism), then one can conclude that one is manipulating something substantive rather than simply valence.

In the *group comparison technique*, researchers first develop a theory to explain how and why specific samples of individuals with varying characteristics will differ in the mean level of the focal constructs that are antipodal to each other (e.g., job satisfaction and dissatisfaction). If the subsequent results confirm the researcher's a priori theory that a sample with a particular set of characteristics (identified a priori) does indeed score high on both poles of a construct (e.g., job satisfaction and dissatisfaction), it adds further credence to the two-dimensional interpretation of the construct. Note that this is methodologically less rigorous than the experimental-based technique because it is subject to alternative

explanations that researchers must be careful to rule out. For example, a sample group with particularly strong acquiescence bias will show high means on both positively- and negatively-valenced items.

Both of these two strategies require both careful planning and ingenious designs from researchers. For example, the experimental technique requires researchers to consider how to manipulate one pole of a construct (optimism) without also manipulating its alleged antipode (pessimism). A detailed discussion of these techniques is beyond the scope of the current paper but is reviewed elsewhere (Borsboom & Mellenbergh, 2007; Borsboom et al., 2004; Embretson, 1983).

Controlling for careless responding. Study 2 compared the effect of careful vis-à-vis careless responding on test item responses, and showed that a construct is less likely to show a two-factor structure among careful respondents as compared to careless respondents. Careless respondents may give a similar answer to every survey item (e.g., 4 out of a 5-point Likert scale) regardless of the actual content of these items. Because the scores for reverse-keyed items are usually re-coded before calculation (e.g., changing the score from 4 to 2 in a 5-point Likert scale), the scores for regular-keyed items and reverse-keyed items will not agree well with each other. In contrast, careful respondents are attentive to item wordings. Researchers are strongly recommended to routinely check for careless responding in their scales and to rely on data only from careful respondents.

Further discussion on the practice of excluding reverse-keyed items

The current findings also have implications for the general issue of whether or not to include reverse-keyed items in survey instruments. As early as the first half of the 20th century, acclaimed researchers began to inquire whether reverse-keyed items should be included in surveys (Cronbach, 1942). Although decades of rigorous investigation have not led to a consensus, many researchers are excluding reverse-keyed items (e.g., Magazine,

Williams, & Williams, 1996; Schriesheim & Eisenbach, 1995; Schriesheim, Eisenbach, & Hill, 1991; Sweeney, Pillitteri, & Kozlowski, 1996). My findings seriously question the soundness of this practice in both individual research and meta-analyses.

Correlations in an individual empirical study. My research shows that when reverse-keyed items are eliminated, the construct score will have a valence that has implications on the magnitude with which it correlates with other variables. If two construct scores are valenced in the same direction, the magnitude of their correlation will be inflated. Conversely, if two construct scores are valenced in an opposite direction, the magnitude of their correlation will be deflated. To reduce bias in study measurements, researchers should strive to use measures that are uncontaminated by the valence effect. This can be accomplished by re-weighting the positively- and negatively-valenced items in the calculation of a construct's score, using a scale with a balanced set of regular- and reverse-keyed items, and/or maximizing method heterogeneity in construct measurements.

Meta-analytic correlations. Researchers often assume that meta-analytic findings are less susceptible to bias than are individual studies because they believe that an aggregation of findings from multiple studies will enable the different sources of error to balance out. Unfortunately, this assumption does not hold true if most of the studies in a meta-analysis suffer from a similar problem (Bobko & Stone-Romero, 1998), such as eliminating reverse-keyed items that have a particular valence (e.g., the negatively-valenced items for the construct extraversion). When this occurs, the summarized findings based on these studies will similarly be distorted. In addition, systematic bias resulting from item valence will be further amplified with any unreliability corrections. Specifically, my empirical results demonstrate that the correction procedure widens the difference in correlations between positively- and negatively-valenced scores. For example, the raw (i.e., uncorrected) correlation of extraversion and introversion scores with negative-valenced

conscientiousness are .108 and .227, respectively (Table 10), a difference of .119. Following correction of unreliability, the correlation was adjusted to .141 and .303 (Table 11), a difference that is .162. Although I understand the advantage of correlation corrections (see Schmidt & Hunter, 1999), researchers should be extremely cautious in interpreting unattenuated correlations that are biased by the valence effect.

In summary, given the predominance of using measurement instruments that have unbalanced use of regular- and reverse-keyed items, I expect that the prevalent effect of item valence already exists in past correlational studies. However, I feel that it is time for researchers to seriously attend to the valence effect in their future research investigations.

Social Desirability Response Bias and Other Explanations for the Valence Effect

My results suggest that the social desirable response bias only partially explained the valence effect. Social desirability responding did not have a high correlation with the valence effect in Study 1. My findings thus suggest that social desirability and valence are theoretically distinct. According to Paulhus (1991), social desirability response bias reflects a desire to exaggerate one's positive cognitive attributes or create a positive impression. Therefore, it does not capture other substantive response styles unrelated to this self-serving bias. Below I suggest some other potential explanations for the valence effect.

One possibility is that the memory system is involved. Research by Showers (1992) has shown that some individuals have a tendency to organize positive and negative knowledge into separately valenced memory categories. Similarly, Credé et al. (2009) argued that, when confronted with positively-valenced items, respondents are likely to tap into positive memories that justify agreement. Similarly, when confronted with negatively-valenced items they tap into negative memories that can lead to agreement. This tendency to focus on valence-relevant memories leads to an increase in the correlations among similarly-valenced items and reduces the correlation among opposite-valenced items. Memory storage

might thus be another mechanism that helps to explain the item valence effects observed in my dissertation.

A second possibility is that participants' motivational systems affect their responses to positively- and negatively-valenced items. Gray (1981, 1987) postulated that individuals have both a motivation to approach and a motivation to avoid, and that these motivations are regulated by two separate biological systems. The first system, the Behavioral Activation System (BAS), results in a goal-oriented tendency to pursue potential rewards (i.e., approach motivation). The second system, the Behavioral Inhibition System (BIS), inhibits us from behaviors that may lead to punishment (i.e., avoidance motivation). Quilty, Oakman, and Risko (2006) showed empirically that BIS scores correlate with a method factor extracted from negatively-valenced items in a self-esteem measure. The stronger the respondent's avoidance motivation, the more likely he or she was to agree with a negatively-valenced self-esteem item (such as "All in all, I am inclined to feel that I am a failure"). This implies that a participants' avoidance motivation should affect ratings of a personality item. However, the empirical result by Quilty et al. (2006) has not been consistently found (DiStefano & Motl, 2006). Thus, I encourage further research to continue examining this potential mechanism for the valence effect.

Both explanations (i.e., the memory system and the motivational system) suggest the existence of individual differences in the valence effect. I do not believe that everyone shows the same magnitude of the valence effect. As previously elaborated, the memory system explanation suggested that individuals may tap into valence-relevant memories that justify agreement when reading a positively- or a negatively-valenced item. Research by Showers (1992) suggested that this response pattern will only happen for individuals who compartmentalize positive and negative memories. To this group of individuals, the recollection of a positive memory will not activate a relevant negative memory and the recall

of a negative memory will not activate a relevant positive memory, because positive and negative memories are stored separately. This cognitive pattern will cause an individual to dissociate positively-valenced memories (such as job satisfaction) from negatively-valenced memories (such as job dissatisfaction), resulting in a two-factor solution of a construct. Showers (1992) also demonstrated that other individuals integrate positive and negative memory together. To them, the recall of positive information may simultaneously lead to the activation of relevant negative information. This latter group of individuals, therefore, are less likely to show a two-dimensional structure of a construct, because they are likely to recall both positive and negative memories before responding to a survey item. Showers' research (1992) thus suggested that individual difference exist in the valence effect. Similarly, the aforementioned motivation system explanation by Gary (1987) postulated that some individuals are more sensitive to positive stimuli while others are more responsive to negative stimuli. This differential sensitivity may have implications regarding how participants answer positively-valenced versus negatively-valenced questions. Future research should explore these and other variables that predict individual differences in responding to positively- and negatively-valenced survey items.

Limitations and Future Directions

Like most psychological research, the two studies in my dissertation have limitations. First, because the current work used only university students and work samples as respondents I cannot necessarily generalize my results to other populations. Replication of my results with a sample of different age (e.g., aged adults, underage people) might be beneficial. Second, although I have included a wide range of measures in my empirical investigations, it would be beneficial to include more measures from other areas of psychology (e.g., developmental psychology, clinical psychology) to further examine the generalizability of my findings. Third, as mentioned

in the preceding section, the current research has not fully explored other potential mechanisms underlying the valence effect. This line of research is extremely important in advancing our current knowledge of item response mechanisms inside a respondent's mind.

Furthermore, in my two studies I did not alternate the order in which the items were presented to the respondents. The systematic line of research by Schwarz (1999) has shown that the order of item presentation can affect participants' responses. For example, when respondents were asked about their life satisfaction before being asked about their marital satisfaction, the correlation between the two kinds of satisfaction was substantively lower than when the order of the two questions was reversed (first marital satisfaction followed by life satisfaction). Apparently, respondents in the latter condition used their judgement of marital satisfaction to evaluate their life satisfaction, resulting in a stronger correlation between the items as compared to what was found using the former condition. However, in Study 2 of my dissertation, the survey items relating to job satisfaction, emotional exhaustion, and turnover intention were scrambled and then randomly placed among each other item. As such it is not clear how the unvaried order of item presentation in Study 2 could provide an alternative explanation to my thesis regarding the systematic effect of item valence.

Another limitation of the current study is that the valence of an item was determined by traditional assumptions from previous research. For example, extraversion, conscientiousness and agreeableness are generally considered more favourable than introversion, "un-conscientiousness" and disagreeableness. Therefore, I made the assumption that items measuring the former three (i.e., extraversion, conscientiousness and agreeableness) are positively-valenced and that items measuring the latter three (i.e., introversion, "un-conscientiousness" and disagreeableness) are

negatively-valenced. However, I am aware of the possibility that the valence of an item may differ in different situations. For example, introversion may be more favourable than extraversion during a job interview for a position as a book editor. Therefore, my analyses have not considered the situational-specificity of an item's valence. Future research should examine how differences in situations may affect the valence of an item, because it may have implications on how item and construct scores correlate with each other (i.e., nomological network of a scale score).

Similarly, my study has not considered the possibility that the factor structure of a construct may differ between individuals. Job satisfaction, for example, may be unidimensional for some individuals but bidimensional for other individuals. The analyses that I conducted thus reflected an aggregated investigation that grouped all individuals together in a single analysis, as opposed to a person-specific investigation that fully considered individual differences in construct dimensionality. In other words, my studies did not consider the possibility that a construct's dimensionality differs across individuals (Borsboom, Mellenbergh, & van Heerden, 2003). The factor structure of a construct at the aggregated level may not apply to all individuals within the data. Future research should thus address the issue regarding individual differences in a construct's dimensionality, and explored whether the impact of item valence on construct dimensionality exists among only some as opposed to all individuals.

The current dissertation did not investigate the effect of different types of reverse-keyed items on the valence effect. According to Holden and Fekken (1990) and Schriesheim and Eisenbach (1995), reverse-keyed items (e.g., for extraversion) include the use of polar opposites (e.g., "I am a reserved person"), negations (e.g., "I am not an extravert"), implicit negations (e.g., "I am an unextroverted person"), negative qualifiers (e.g,. "I am seldom extroverted") and double negation (e.g., "I am not an

unreserved person"). Does one type of reverse-keyed items have a stronger effect on valence than the other types? Intuitively, the more explicit a negation is, the more it will lead to interpretation difficulties among respondents because they will experience problems with agreement to a negated item (Holden & Fekken, 1990). According to this view, double negation items may cause the most severe interpretation difficulties and thus unsystematic errors among respondents. As such, the valence effect on double negation items may be smallest compared to other types of reverse-keyed items because most variance in that case will be attributed to random error. Future research should examine the implication of different types of reverse-keyed items on the valence effect.

The final limitation that I will discuss here is that my research used a cross-sectional design and did not examine the temporal stabilities of the valence effect. Although I speculate that this valence effect is caused by temporally stable factors such as memory storage and motivational systems, the exact determinants of the valence effect warrant further investigation. Marsh et al. (2010) recently conducted an analysis of this with a self-esteem measure and found that the valence factors correlate about .40-.60 across four waves of data collection separated by one year (see Motl & DiStefano, 2002, for similar longitudinal results). Those results suggested that there are certain temporal stabilities of the valence effect, but future research is needed to reproduce and expand on their findings for measures other than self-esteem.

Conclusion

The major goal of my dissertation was to demonstrate that item valence has a strong potential to influence decisions with regard to the dimensionality of a construct domain. Valence affects not only factor analytic results but also the results of nomological network analyses. It is my hope that my results will emphasize the potential problems associated with the interpretation of exploratory factor analyses and nomological network analyses in

corresponding empirical assessments. I encourage researchers to employ my suggested methods to improve their factor analytic and nomological network investigations into construct dimensionality.

REFERENCES

Alhija, F. N.-A. and Wisenbaker, J. (2006). A Monte Carlo Study Investigating the Impact of Item Parceling Strategies on Parameter Estimates and Their Standard Errors in CFA. *Structural Equation Modeling, 13*, 204 - 228.

Altemeyer, R. A. (1981). *Right-wing authoritarianism*. Winnipeg: University of Manitoba Press.

Bäckström, M., Björklund, F., & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality, 43*, 335–344.

Baken, D. (1966). *The duality of human existence: An essay on Psychology and Religion*. Oxford, England: Rand Mcnally.

Beauvois, J. L., & Dubois, N. (2000). Affordances in social judgment: Experimental proof of why it is a mistake to ignore how others behave towards a target and look solely at how the target behaves. *Swiss Journal of Psychology, 59*, 16–33.

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16*, 78-117.

Bernstein , I. H., Garbin, C., & Teng, G. (1988). *Applied Multivariate Analysis.* New York: Springer-Verlag.

Bobko, P., & Stone-Romero, E. (1998). Meta-analysis is another useful research tool but it is not a panacea. In G. Ferris (Ed.), Research in personnel and human resources management, Vol. 16, pp. 359–397, Greenwich, CT: JAI Press.

Borsboom, D., & Mellenbergh, G. J. (2007). Test validity in cognitive assessment. *Cognitive diagnostic assessment for education: Theory and applications*. Cambridge University Press.

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2004). The concept of validity. *Psychological Review, 111*, 1061–1071.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage.

Chernyshenko, O., Stark, S., Credé, M., Wadlington, P. L., & Lee, W. C. (2003, April). Improving the Measurement of Job Attitudes: The Development of the JSI. In Rupp, D. E. (Chair), New Frontiers in Job Satisfaction, Job Performance, and Their Linkages. Symposium conducted at the 18th Annual Meeting of the Society for Industrial and Organizational Psychology, Orlando, Florida.

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin, 136*, 1092-1122.

Credé, M., Oleksandr, C. S., Bagraim, J., & Sully, M. (2009). Contextual performance and job satisfaction-dissatisfaction distinction: Examining artifacts and utility. *Human Performance, 22*, 246-272.

Cronbach, L. J. (1942). Studies of acquiescence as a factor in the true-false test. *Journal of Educational Psychology, 33*, 401-415.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281-302.

Dalbert, C., Lipkus, I. M., Sallay, H., & Goch, I. (2001). A just and an unjust world: Structure and validity of different world beliefs. *Personality and Individual Differences, 30*, 561-577.

Demerouti, E., Karina, M., & Bakker, A. B. (2010). Burnout and work engagement: A thorough investigation of the independency of both constructs. *Journal of Occupational Health Psychology, 15*, 209-222.

Diener, E., Wirtz, D., Tov, W., Chu, K.-P., Choi, D.-w., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97*, 143-156.

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling, 13*, 440-464.

Donovan, J. J., Dwight, S. A., & Hurtz, G. M. (2003). An assessment of the prevalence, severity, and verifiability of entry-level applicant faking using the randomized response technique. *Human Performance, 16*, 81–106.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of the relationship between constructs and measures. *Psychological Methods, 5*, 155-174.

Eid, M., & Diener, E. (Eds.). (2006). *Handbook of multimethod measurement in psychology.* Washington, DC: American Psychological Association.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93*, 179-197.

Esses, V.M., Jackson, L.M., & Armstrong, T.L. (1998). Intergroup competition and attitudes toward immigrants and immigration: An instrumental model of group conflict. *Journal of Social Issues, 54*, 699-724.

Fincham, F. D., & Linfield, K. J. (1997). A new look at marital quality: Can spouses feel positive and negative about their marriage? *Journal of Family Psychology, 11*, 489-502.

Fox, S., Spector, P. E., Bruursema, K., Kessler, S., & Goh, A. (2007, August). *Necessity is the mother of behavior: Organizational constraints, CWB and OCB.* Paper presented at the meeting of the Academy of Management, Philadelphia, PA.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652-670.

Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. London, UK: Sage Publication.

Goffin. R.D., & Boyd, A.C. (2009). Faking and personality assessment in personnel selection: Advancing models of faking. *Canadian Psychology, 50*, 151-160.

Gray, J. A. (1981). A critique of Eysenck's theory of personality. In H. J. Eysenck (Ed.), *A model for personality* (pp. 246-276). New York: Springer.

Gray, J. A. (1987). Perspectives on anxiety and impulsivity: A commentary. *Journal of Research in Personality, 21*, 493-509.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*, 26-42.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96.

Greenberger, E., Chen, C. Dmitrieva, J., Farruggia, S. P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences, 35*, 1241-1254.

Guilford, J. P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education.* New York: McGraw-Hill.

Hall, E.T. (1983). *The dance of life*. Garden City, NY: Anchor Press.

Herzberg, P. Y., Glaesmer, H., & Hoyer, J. (2006). Separating Optimism and Pessimism: A Robust Psychometric Analysis of the Revised Life Orientation Test (LOT-R). *Psychological Assessment, 18*, 433–438.

Hinkins, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*, 967-988.

Hofstede, G. (1980). *Culture's consequences.* Thousand Oaks, CA: Sage.

Hofstee, W. K. (1990). The use of everyday personality language for scientific purposes. *European Journal of Personality, 4*, 77-88.

Holden, R., & Fekken, G. (1990). Structured psychopathological test item characteristics and validity. *Psychological Assessment, 2*, 35-40.

Holden, R. R., Fekken, G. C., & Cotton, D. H. (1991). Assessing psychopathology using structured test-item response latencies. *Psychological Assessment, 3*, 111-118.

Horan, P.M., DiStefano, C., & Motl, R. W. (2003). Wording effects in self esteem scales: Methodological artifact or response style? *Structural Equation Modeling, 10*, 444–455.

Jackson, D. N. (1984). *Personality research form manual* (3rd ed.). Port Huron, MI: Research Psychologists Press.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford.

Kwan, V. S. Y., Bond, M. H., & Singelis, T. M. (1997). Pancultural explanations to life satisfaction: Adding relationship harmony to self-esteem. *Journal of Personality and Social Psychology, 5*, 1038-1051.

Lance, C. E., Dawson, B., Birkelbach, D., & Hoffman, B. J. (2010). Method effects, measurement error, and substantive conclusions. *Organizational Research Methods, 13*, 435-455.

Li, A., & Bagger, J. (2007). The Balanced Inventory of Desirable Responding (BIDR): A reliability generalization study. *Educational and Psychological Measurement, 67,* 525-544.

Locke, E. A. (1976). The nature and causes of job satisfaction. In M. D. Dunnette (Ed.), Handbook of industrial and organizational psychology (pp. 1297-1343). Chicago: Rand McNally.

Locke, H. J., & Wallace, K. M. (1959). Short marital adjustment prediction tests: Their reliability and validity. *Marriage and Family Living, 21,* 251-255.

MacCorguodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review, 55,* 95-107.

Magazine, S. L., Williams, L. J., & Williams, M. L. (1996). A confirmatory factor analysis examination of reverse coding effects in Meyer and Allen's affective and continuance commitment scales. *Educational and Psychological Measurement, 56,* 241-250.

Marsh, H. W. (1989). Confirmatory factor analyses of multitrait–multimethod data: Many problems and a few solutions. *Applied Psychological Measurement, 13,* 335–361.

Marsh, H. W. (1996). Positive and negative self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70,* 810-819.

Marsh, H. W., & Bailey, M. (1991). Confirmatory factor analysis of multitrait–multimethod data: A comparison of the behavior of alternative models. *Applied Psychological Measurement, 15,* 47–70.

Marsh, H W., Scalas, L. F., & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem scale: Traits, ephemeral artifacts, and stable response styles. *Psychological Assessment, 22,* 366-381.

Marshall, G. N., Wortman, C. B., Kusulas, J. W, Hervig, L. K., & Vickers, R. R, Jr. (1992). Distinguishing optimism from pessimism: Relations to fundamental dimensions of mood and personality. *Journal of Personality and Social Psychology. 62*, 1067-1074.

McPherson, J., & Mohr, P. (2005). The role of item extremity in the emergence of keying-related factors: An exploration with Life Orientation Test. *Psychological Methods, 10*, 120-131.

Meade, A. W., & Craig, S. B. (in press). Identifying careless responses in survey data. *Psychological Methods*.

Meeker, W. Q., & Escobar, L. (1998). *Statistical methods for reliability data*. New York: John Wiley.

Meng, X.-L., Rosenthal, R., & Rubin, S. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111*,172-175.

Meyer, J. P. & Allen, N. J. (1984). Testing the "side-bet theory" of organizational commitment: Some methodological considerations. *Journal of Applied Psychology, 69*, 372-378.

Miller, T. R., & Cleary, T. A. (1993). Direction of wording effects in balanced scales. *Educational and Psychological Measurement, 53*, 51-60.

Motl, R. W., & DiStefano, C. (2002). Longitudinal invariance of self-esteem and method effects associated with negatively worded items. *Structural Equation Modeling, 9*, 562–578.

Muthén, L.K. and Muthén, B.O. (1998-2010). *Mplus User's Guide. Sixth Edition*. Los Angeles, CA: Muthén & Muthén.

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement, 63*, 729-757.

Oh, I.-S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the Five-Factor Model of personality: A meta-analysis. *Journal of Applied Psychology, 96*, 762-773.

Owens, T. J. (1994). Two dimensions of self-esteem: Reciprocal effects of positive self-worth and self-deprecation on adolescent problems. *American Sociological Review, 59*, 391-407.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59) San Diego, CA: Academic Press.

Paunonen, S.V., & Ashton, M.C. (2001). Big Five predictors of academic achievement. *Journal of Research in Personality, 35*, 78-90.

Paunonen, S. V., & Hong, R. Y. (in press). In defense of personality traits. In P. R. Shaver and M. Mikulincer (Eds.), *Handbook of personality and social psychology*. New York: APA.

Paunonen, S. V., & Jackson, D. N. (1979). Nonverbal trait inference. *Journal of Personality and Social Psychology, 37*, 1645-1659.

Peabody, D. (1967). Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology Monographs*, 7(Whole No. 644).

Quilty, L. C., Oakman, J. M., & Risko, E. (2006). Correlates of the Rosenberg Self-Esteem Scale Method Effects. *Structural Equation Modeling, 13*, 99-117.

Rauch, W. A., Schweizer, K., & Moosbrugger, H. (2007). Method effects due to social desirability as a parsimonious explanation of the deviation from unidimensionality in LOT-R scores. *Personality and Individual Differences, 42*, 1597−1607.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rubin, Z., & Peplau, L.A. (1975). Who believes in a just world? *Journal of Social Issues, 31*, 65-90.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika,* 66, 507-514.

Saucier, G. (1994). Separating description and evaluation in the structure of personality attributes. *Journal of Personality and Social Psychology, 66,* 141-154.

Scheier, M. E, & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology,* 4, 219-247.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of Life Orientation Test. *Journal of Personality and Social Psychology, 67,* 1063-1078.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist, 47*, 1173-1181.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1,* 199-223.

Schmidt, F. L., & Hunter, J. E. (1999). Theory testing and measurement error. *Intelligence, 27*, 183–198.

Schmitt, N., & Stuits, D. M. (1985). Factor defined by negatively keyed items: The result of careless respondents. *Applied Psychological Measurement, 9*, 367-373.

Schriesheim, C. A., & Eisenbach, R. (1995). An exploratory and confirmatory factor-analytic investigation of item wording effects on the obtained factor structures of survey questionnaire measures. *Journal of Management, 21*, 1177-1193.

Schriesheim, C. A., Eisenbach, R. J., & Hill, K. D. (1991). The effect of negation and polar opposite item reversals on questionnaire reliability and validity: An experimental investigation. *Educational and Psychological Measurement, 51*, 67-78.

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologists, 54*, 93-105.

Showers, C. (1992). Compartmentalization of positive and negative self-knowledge: Keeping bad apples out of the bunch. *Journal of Personality and Social Psychology, 62*, 1036-1049.

Sidanius, J. & Pratto, F. (2001). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Spector, P. E., Bauer, J. A., & Fox, S. (2010). Measurement artifacts in the assessment of counterproductive work behavior and organizational citizenship behavior: Do we know what we think we know? *Journal of Applied Psychology, 95*, 781-790.

Spector, P. E., Fox, S., Penney, L. M., Bruursema, K., Goh, A., & Kessler, S. (2006). The dimensionality of counterproductivity: Are all counterproductive behaviors created equal? *Journal of Vocational Behavior, 68*, 446–460.

Stöber, J., Dette, D. E., & Musch, J. (2002). Comparing continuous and dichotomous scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment, 78*, 370-389.

Sweeney, C. T., Pillitteri, J. L., & Kozlowski, L. T. (1996). Measuring drug urges by questionnaire: Do not balance scales. *Addictive Behaviors, 21*, 199-204.

Syracuse University (2005). The StudyResponse Project: An online social science research resource. Retrieved from http://studyresponse.syr.edu/

Tellegen, A. (1993). Folk concepts and psychological concepts of personality and personality disorder. *Psychological Inquiry, 4*, 122-130.

Tomás, J. M., & Oliver, A. (1999) Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling, 6*, 84-98.

Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: A reference dependent model. *Quarterly Journal of Economics, 107*, 1039-1061.

Vagg, P. R., Spielberger, C. D., & O'Hearn, T. P., Jr. (1980). Is the State-Trait Anxiety Inventory multidimensional? *Personality and Individual Differences, 1*, 207-214.

Vautier, S., Callahan, S., Moncany, D., & Sztulman, H. (2004). A bistable view of single constructs measured using balanced questionnaires: Application to trait anxiety. *Structural Equation Modeling, 11*, 261-271.

Vautier, S., & Pohl, S. (2009). Do balanced scales assess bipolar constructs? The case of the STAI scales. *Psychological Assessment, 21*, 187-193.

Verkuyten, M. (2003). Positive and negative self-esteem among ethnic minority early adolescents: Social and cultural sources and threats. *Journal of Youth and Adolescence, 32*, 267-277.

Vigneau, F. & Cormier, S. (2008). The factor structure of the State-Trait Anxiety Inventory: An alternative view. *Journal of Personality Assessment, 90*, 280-285.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063–1070.

William, L. J., Cote, J. A., & Buckley, M. R. (1989). Lack of method variance in self-reported affect and perceptions at work: Reality or artifact? *Journal of Applied Psychology, 74*, 462-468.

Wood, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Asssessment, 28*, 186-191.

Yuan, K.-H., Bentler, P. M., & Kano, Y. (1997). On averaging variables in a CFA model. *Behaviormetrika, 24*, 71–83.

APPENDIX A

QUESTIONS IN STUDY 1

(+) = positively-valenced item
(-) = negatively-valenced item

I. Big 5 Personality and Social Desirability Response Bias Questions

In a 5-point scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*)

i. Openness to Experience

1. I believe in the importance of art. (+)
2. I have a vivid imagination. (+)
3. I tend to vote for liberal political candidates. (+)
4. I carry the conversation to a higher level. (+)
5. I enjoy hearing new ideas. (+)
6. I am not interested in abstract ideas. (-)
7. I do not like art. (-)
8. I avoid philosophical discussions. (-)
9. I do not enjoy going to art museums. (-)
10. I tend to vote for conservative political candidates. (-)

ii. Conscientiousness

1. I am always prepared. (+)
2. I pay attention to details. (+)
3. I get chores done right away. (+)
4. I carry out my plans. (+)
5. I make plans and stick to them. (+)
6. I waste my time. (-)
7. I find it difficult to get down to work. (-)
8. I do just enough work to get by. (-)
9. I don’t see things through. (-)
10. I shirk my duties. (-)

iii. Extraversion

1. I feel comfortable around people. (+)
2. I make friends easily. (+)
3. I am skilled in handling social situations. (+)
4. I am the life of the party. (+)
5. I know how to captivate people. (+)
6. I have little to say. (-)

7. I keep in the background. (-)
8. I would describe my experiences as somewhat dull. (-)
9. I don't like to draw attention to myself. (-)
10. I don't talk a lot. (-)

iv. Agreeableness

1. I have a good word for everyone. (+)
2. I believe that others have good intentions. (+)
3. I respect others. (+)
4. I accept people as they are. (+)
5. I make people feel at ease. (+)
6. I have a sharp tongue. (-)
7. I cut others to pieces. (-)
8. I suspect hidden motives in others. (-)
9. I get back at others. (-)
10. I insult people. (-)

v. Neuroticism

1. I rarely get irritated. (+)
2. I seldom feel blue. (+)
3. I feel comfortable with myself. (+)
4. I am not easily bothered by things. (+)
5. I am very pleased with myself. (+)
6. I often feel blue. (-)
7. I dislike myself. (-)
8. I am often down in the dumps. (-)
9. I have frequent mood swings. (-)
10. I panic easily. (-)

vi. Self-Deception

1. I always know why I like things. (+)
2. Once I've made up my mind, other people can seldom change my opinion. (+)
3. My first impressions of people usually turn out to be right. (+)
4. I am fully in control of my own fate. (+)
5. I never regret my decisions. (+)
6. The reason I vote is because my vote can make a difference. (+)
7. I am a completely rational person. (+)
8. I am very confident of my judgments. (+)

9. It’s all right with me if some people happen to dislike me. (+)
10. I don’t care to know what other people really think of me. (+)
11. I have not always been honest with myself. (-)
12. I am not a safe driver when I exceed the speed limit. (-)
13. It’s hard for me to shut off a disturbing thought. (-)
14. I sometimes lose out on things because I can’t make up my mind soon enough. (-)
15. My parents were not always fair when they punished me. (-)
16. I rarely appreciate criticism. (-)
17. I don’t always know the reasons why I do the things I do. (-)
18. It would be hard for me to break any of my bad habits. (-)
19. When my emotions are aroused, it biases my thinking. (-)

vii. Impression Management

1. I never cover up my mistakes. (+)
2. I never swear. (+)
3. I always obey laws, even if I’m unlikely to get caught. (+)
4. When I hear people talking privately, I avoid listening. (+)
5. I always declare everything at customs. (+)
6. I have never dropped litter on the street. (+)
7. I never take things that don’t belong to me. (+)
8. I have never damaged a library book or store merchandise without reporting it. (+)
9. I don’t gossip about other people’s business. (+)
10. I have some pretty awful habits. (-)
11. There have been occasions when I have taken advantage of someone. (-)
12. I sometimes try to get even rather than forgive and forget. (-)
13. I have said something bad about a friend behind his or her back. (-)
14. I have received too much change from a salesperson without telling him or her. (-)
15. When I was young I sometimes stole things. (-)
16. I sometimes drive faster than the speed limit. (-)
17. I have done things that I don’t tell other people about. (-)
18. I have taken sick-leave from work or school even though I wasn’t really sick. (-)
19. I sometimes tell lies if I have to. (-)

II. Rosenberg Self-Esteem Scale

In a 4-point scale from 1 (*Strongly Disagree*) to 4 (*Strongly Agree*)

1. On the whole, I am satisfied with myself. (+)
2. I feel that I have a number of good qualities. (+)
3. I am able to do things as well as most other people. (+)
4. I feel that I'm a person of worth, at least on an equal plane with others. (+)
5. I take a positive attitude toward myself. (+)
6. At times, I think I am no good at all. (-)
7. I feel I do not have much to be proud of. (-)
8. I certainly feel useless at times. (-)
9. I wish I could have more respect for myself. (-)
10. All in all, I am inclined to feel that I am a failure. (-)

III. Social Dominance Orientation Scale

In a 7-point scale from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*)

1. We should do what we can to equalize conditions for different groups. (+)
2. We would have fewer problems if we treated different groups more equally. (+)
3. No one group should dominate in society. (+)
4. Group equality should be our ideal. (+)
5. All groups should be given an equal chance in life. (+)
6. We must increase social equality. (+)
7. We must strive to make incomes more equal. (+)
8. It would be good if all groups could be equal. (+)
9. Some groups of people are just more worthy than others. (-)
10. In getting what your group wants, it is sometimes necessary to use force against other groups. (-)
11. If certain groups of people stayed in their place, we would have fewer problems. (-)
12. To get ahead in life, it is sometimes necessary to step on other groups. (-)
13. Superior groups should dominate inferior groups. (-)
14. It's probably a good thing that certain groups are at the top and other groups are at the bottom. (-)
15. Sometimes other groups must be kept in their place. (-)
16. Inferior groups should stay in their place. (-)

IV. Zero-Sum Beliefs About Immigrants Scale

In a 7-point scale from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*)

1. When immigrants make economic gains, it is not at the expense of Canadians already living here. (+)
2. More good jobs for immigrants means fewer good jobs for Canadians already living here. (-)
3. Canadians already living here lose out when immigrants make political and economic gains. (-)
4. The more business opportunities are made available for immigrants, the fewer business opportunities are available for Canadians already living here.(-)
5. More tax dollars spent on immigrants means fewer tax dollars spent on Canadians already living here. (-)
6. Money spent on social services for immigrants means less money for services for Canadians already living here.(-)

V. Belief in a Dangerous World Scale

In a 9-point Likert scale from -4 (*Very Strongly Disagree*) to +4 (*Very Strongly Agree*)

1. Although it may appear that things are constantly getting more dangerous and chaotic, it really isn't so. Every era has its problems, and a person's chances of living a safe, untroubled life are better today than ever before. (+)
2. Our society is not full of immoral and degenerate people who prey on decent people. News reports of such cases are grossly exaggerating and misleading. (+)
3. The 'end' is not near. People who think that earthquakes, wars, and famines mean God might be about to destroy the world are being foolish. (+)
4. Despite what one hears about "crime in the street", there probably isn't any more now than there has been. (+)
5. If a person takes a few sensible precautions, nothing bad will happen to them. We do not live in a dangerous world. (+)
6. Our country is not falling apart or rotting from within. (+)
7. It seems that every year there are fewer and fewer truly respectable people, and more and more persons with no morals at all who threaten everyone else. (-)
8. If our society keeps degenerating the way it has been lately, it's liable to collapse like a rotten log and everything will be chaos. (-)
9. There are many dangerous people in our society who will attack someone out of pure meanness, for no reason at all. (-)
10. Every day, as our society becomes more lawless and bestial, a person's chances of being robbed, assaulted, and even murdered go up and up. (-)
11. Things are getting so bad, even a decent law-abiding person who takes sensible precautions can still become a victim of violence and crime. (-)
12. Any day now, chaos and anarchy could erupt around us. All the signs are pointing to it. (-)

APPENDIX B

QUESTIONS IN STUDY 2

(+) = positively-valenced item
(-) = negatively-valenced item

Instructions: Please read each of the following statements and indicate how strongly you agree or disagree with that statement.

1 = Strongly Disagree; 2 = Disagree; 3 = Neither Disagree or Agree; 4 = Agree; 5 = Strongly Agree

I. Satisfaction with Work Subscale in Illinois Job Satisfaction Index (Chernyshenko, Stark, Crede, Wadlington, & Lee, 2003)

1. My work is meaningful. (+)
2. I am better off working in my organization than anywhere else. (+)
3. I look forward to coming to work. (+)
4. I enjoy most of what I do at work. (+)
5. I love my job. (+)
6. My job gives me a sense of dignity. (+)
7. I am proud of my work. (+)
8. I would recommend my job to others. (+)
9. I don't like my work. (-)
10. This job is terrible. (-)
11. My work is too repetitive. (-)
12. I can't wait to leave work each day. (-)
13. My work tires me out very quickly. (-)
14. This job is frustrating. (-)
15. There are a lot of things I do not like about my work. (-)

II. Balanced Inventory of Job Satisfaction

i. Satisfaction with Work

1. I like my work. (+)
2. My work is meaningful. (+) **
3. I dislike my work. (-)
4. My work is trivial. (-)

ii. Satisfaction with Immediate Supervisor

5. My supervisor treats me well. (+)
6. I am satisfied with my supervisor. (+)
7. My supervisor treats me badly. (-)
8. I am dissatisfied with my supervisor. (-)

iii. Satisfaction with Coworker

9. I like my coworkers. (+)
10. My coworkers are friendly. (+)
11. I dislike my coworkers. (-)
12. My coworkers are distant. (-)

iv. Satisfaction with Pay and Benefits

13. I am satisfied with my pay. (+)
14. I am satisfied with the benefits I receive. (+)
15. I am dissatisfied with my pay. (-)
16. I am dissatisfied with the benefits provided by my company. (-)

** This item duplicated with item 1 in Satisfaction with Work Subscale from Illinois Job Satisfaction Index

III. Turnover Intention Scale

1. I intend to remain my organization for at least the next three years. (+)
2. I intend to leave my organization in the near future. (-)

IV. Exhaustion Items from Oldenberg Burnout Inventory (Demerouti, Mostert, & Bakker, 2010)

1. I can tolerate the pressure of my work very well. (+)
2. After working, I have enough energy for my leisure activities. (+)
3. Usually, I can manage the amount of my work well. (+)
4. When I work, I usually feel energized. (+)
5. There are days when I feel tired before I arrive at work. (-)
6. After work, I tend to need more time than in the past in order to relax and feel better. (-)
7. During my work, I often feel emotionally drained. (-)

8. After my work, I usually feel worn out and weary. (-)

How often have you done each of the following things in your present job over the past year?

1 = Never; 2 = Once or twice; 3 = Once or twice per month; 4 = Once or twice per week; 5 = Everyday

V. Organizational Citizenship Behavior Items (Spector, Bauer, & Fox, 2010)

1. Took time to advise, coach, or mentor a co-worker. (+)
2. Helped a co-worker learn new skills or shared job knowledge. (+)
3. Helped new employees get oriented to the job. (+)
4. Lent a compassionate ear when someone had a work problem. (+)
5. Offered suggestions to improve how work is done. (+)
6. Helped a co-worker who had too much to do.(+)
7. Volunteered for extra work assignments. (+)
8. Worked weekends or other days off to complete a project or task. (+)
9. Volunteered to attend meetings or work on committees on your own time. (+)
10. Gave up meal and other breaks to complete work. (+)

VI. Counterproductive Work Behavior Items (Spector, Bauer, & Fox, 2010)

1. Purposely wasted your employer's materials/supplies. (-)
2. Complained about insignificant things at work. (-)
3. Told people outside the job what a lousy place you work for. (-)
4. Came to work late without permission. (-)
5. Stayed home from work and said you were sick when you weren't. (-)
6. Insulted someone about their job performance. (-)
7. Made fun of someone's personal life. (-)
8. Ignored someone at work. (-)
9. Started an argument with someone at work. (-)
10. Insulted or made fun of someone at work. (-)
11. Took supplies or tools home without permission. (-)

VII. Scale of Positive and Negative Experience (SPANE; Diener et al., 2010)

Please think about what you have being doing and experiencing during the past month at work. Then report how much you experienced each of the following feelings.

1 = Very rarely or never; 2 = Rarely; 3 = Sometimes; 4 = Often; 5 = Very often or always

i. Positive Affect

1. Positive (+)
2. Good (+)
3. Pleasant (+)
4. Happy (+)
5. Joyful (+)
6. Contented (+)

ii. Negative Affect

1. Negative (-)
2. Bad (-)
3. Unpleasant (-)
4. Sad (-)
5. Afraid (-)
6. Angry (-)

VIII. Extraversion and Emotional Stability Scale (Goldberg, 1992)

Please describe yourself as you are <u>generally</u>, as compared with other persons you know of. Describe yourself <u>as you are</u>, not as you wish to be in the future.

1 = Very Inaccurate; 2 = Inaccurate; 3 = Neither Agree or Disagree; 4 = Accurate; 5 = Very Accurate

I am generally or typically:

i. Extraversion

1. extraverted (+)
2. energetic (+)
3. talkative (+)
4. bold (+)
5. active (+)
6. assertive (+)
7. adventurous (+)
8. introverted (-)
9. unenergetic (-)
10. silent (-)
11. timid (-)
12. inactive (-)
13. unassertive (-)
14. unadventurous (-)

ii. Emotional Stability

1. calm (+)
2. relaxed (+)
3. at ease (+)
4. not envious (+)
5. stable (+)
6. contented (+)
7. unemotional (+)
8. angry (-)

9. tense (-)
10. nervous (-)
11. envious (-)
12. unstable (-)
13. discontented (-)
14. emotional (-)

APPENDIX C

CORRELATION BETWEEN SOCIAL DESIRABILITY SCORES AND MEASUREMENT ITEMS IN SAMPLE 1B

	Averaged Scores Method			Paulhus Scoring Method		
	DESIR	IM	SDEP	DESIR	IM	SDEP
CONSCIENTIOUSNESS						
I am always prepared.	.28	.19	.28	.00	-.01	.01
I pay attention to details.	.15	.08	.17	.15	.09	.14
I get chores done right away.	.31	.26	.24	-.01	.01	-.02
I carry out my plans.	.26	.13	.30	.10	.07	.10
I make plans and stick to them.	.26	.15	.29	.07	.07	.03
I waste my time.	-.40	-.31	-.33	.13	.13	.07
I find it difficult to get down to work.	-.33	-.23	-.31	.12	.11	.08
I do just enough work to get by.	-.28	-.23	-.21	.04	.08	-.03
I don't see things through.	-.28	-.15	-.31	-.03	.02	-.08
I shirk my duties.	-.28	-.26	-.19	.02	.06	-.04
EXTRAVERSION						
I feel comfortable around people.	.15	-.01	.28	-.01	-.02	.00
I make friends easily.	.10	-.05	.24	.06	.02	.08
I am skilled in handling social situations.	.16	-.03	.32	.05	.01	.08
I am the life of the party.	-.06	-.20	.13	.05	.03	.05
I know how to captivate people.	.07	-.08	.23	.12	.05	.15
I have little to say.	-.05	.06	-.16	-.10	-.04	-.12
I keep in the background.	-.07	.08	-.23	-.01	.02	-.04
I would describe my experiences as somewhat dull.	-.16	-.02	-.27	-.06	-.02	-.08
I don't like to draw attention to myself.	.09	.17	-.05	.01	.06	-.04
I don't talk a lot.	.01	.12	-.13	-.06	-.01	-.07

Note. DESIR = Total Desirability Scores; IM = Impression Management Scores; SDEP = Self-Deception Scores. Calculation methods of these scores were explained in the method section of Study 1.

(continue to the next page)

(continue from the previous page)

	Averaged Scores Method			Paulhus Scoring Method		
	DESIR	IM	SDEP	DESIR	IM	SDEP
AGREEABLENESS						
I have a good word for everyone.	.21	.24	.08	.04	.02	.04
I believe that others have good intentions.	.15	.15	.08	-.02	.02	-.06
I respect others.	.24	.21	.16	.08	.07	.06
I accept people as they are.	.21	.18	.16	.08	.06	.06
I make people feel at ease.	.19	.12	.20	.10	.07	.09
I have a sharp tongue.	-.20	-.30	.00	.18	.09	.20
I cut others to pieces.	-.27	-.32	-.08	.13	.08	.12
I suspect hidden motives in others.	-.26	-.25	-.16	.22	.15	.20
I get back at others.	-.40	-.49	-.12	.17	.14	.12
I insult people.	-.45	-.49	-.19	.12	.11	.09
OPENNESS						
I believe in the importance of art.	.07	.09	.01	.12	.09	.10
I have a vivid imagination.	.04	.00	.06	.24	.14	.24
I tend to vote for liberal political candidates.	-.08	-.06	-.07	.09	.04	.10
I carry the conversation to a higher level.	.11	-.01	.20	.16	.10	.16
I enjoy hearing new ideas.	.12	.07	.12	.21	.14	.18
I am not interested in abstract ideas.	-.01	.00	-.01	-.06	-.01	-.09
I do not like art.	-.06	-.07	-.02	-.07	-.05	-.06
I avoid philosophical discussions.	-.06	-.02	-.08	-.10	-.03	-.13
I do not enjoy going to art museums.	-.11	-.12	-.06	-.03	-.01	-.04
I tend to vote for conservative political candidates.	.09	.05	.11	-.04	-.01	-.06

Note. DESIR = Total Desirability Scores; IM = Impression Management Scores; SDEP = Self-Deception Scores. Calculation methods of these scores were explained in the method section of Study 1.

(continue to the next page)

(continue from the previous page)

	Averaged Scores Method			Paulhus Scoring Method		
	DESIR	IM	SDEP	DESIR	IM	SDEP
NEUROTICISM						
I rarely get irritated.	.24	.21	.17	-.12	-.06	-.12
I seldom feel blue.	.24	.09	.31	-.11	-.06	-.10
I feel comfortable with myself.	.33	.11	.46	.04	.01	.06
I am not easily bothered by things.	.26	.13	.31	-.07	-.03	-.08
I am very pleased with myself.	.29	.07	.43	.03	.01	.03
I often feel blue.	-.28	-.11	-.36	.12	.09	.10
I dislike myself.	-.30	-.09	-.42	.04	.03	.04
I am often down in the dumps.	-.31	-.13	-.39	.08	.05	.08
I have frequent mood swings.	-.35	-.22	-.36	.17	.14	.12
I panic easily.	-.21	-.01	-.37	.07	.08	.03
SELF-ESTEEM						
On the whole, I am satisfied with myself.	.27	.07	.40	.03	-.01	.05
I feel that I have a number of good qualities.	.19	.01	.33	.14	.06	.17
I am able to do things as well as most other people.	.21	.04	.33	.13	.07	.14
I feel that I'm a person of worth, at least on an equal plane with others.	.20	.05	.31	.11	.03	.14
I take a positive attitude toward myself.	.31	.10	.43	.00	-.03	.03
At times, I think I am no good at all.	-.36	-.15	-.46	.11	.16	.01
I feel I do not have much to be proud of.	-.23	-.05	-.35	-.03	.00	-.05
I certainly feel useless at times.	-.34	-.14	-.44	.06	.11	-.01
I wish I could have more respect for myself.	-.34	-.14	-.45	.03	.05	.00
All in all, I am inclined to feel that I am a failure.	-.29	-.09	-.40	.02	.03	.00

Note. DESIR = Total Desirability Scores; IM = Impression Management Scores; SDEP = Self-Deception Scores. Calculation methods of these scores were explained in the method section of Study 1.

(continue to the next page)

(continue from the previous page)

	Averaged Scores Method			Paulhus Scoring Method		
	DESIR	IM	SDEP	DESIR	IM	SDEP
SOCIAL DOMINANCE ORIENTATION						
We should do what we can to equalize conditions for different groups.	.08	.14	-.02	.08	.09	.04
We would have fewer problems if we treated different groups more equally.	.08	.12	.01	.11	.10	.07
No one group should dominate in society.	.08	.13	-.01	.03	.03	.02
Group equality should be our ideal.	.14	.16	.05	.07	.08	.04
All groups should be given an equal chance in life.	.11	.14	.03	.11	.11	.06
We must increase social equality.	.09	.13	.01	.05	.05	.02
We must strive to make incomes more equal.	.08	.13	-.02	.00	.04	-.05
It would be good if all groups could be equal.	.12	.14	.04	.06	.06	.04
Some groups of people are just more worthy than others.	-.14	-.19	-.01	.02	.02	.00
In getting what your group wants, it is sometimes necessary to use force against other groups.	-.14	-.21	.01	.01	.00	.02
If certain groups of people stayed in their place, we would have fewer problems.	-.16	-.20	-.03	.07	.03	.07
To get ahead in life, it is sometimes necessary to step on other groups.	-.24	-.30	-.06	.05	.04	.05
Superior groups should dominate inferior groups.	-.06	-.13	.05	-.05	-.05	-.03
It's probably a good thing that certain groups are at the top and other groups are at the bottom.	-.08	-.15	.03	.03	.02	.03
Sometimes other groups must be kept in their place.	-.12	-.18	.00	.04	.01	.05
Inferior groups should stay in their place.	-.09	-.14	.00	-.05	-.05	-.03

Note. DESIR = Total Desirability Scores; IM = Impression Management Scores; SDEP = Self-Deception Scores. Calculation methods of these scores were explained in the method section of Study 1.

APPENDIX D

STUDY 2 ETHICS APPROVAL

Department of Psychology [redacted]

Use of Human Subjects - Ethics Approval Notice

Review Number	**11 04 06**	**Approval Date**	**11 04 25**
Principal Investigator	**John Meyer/Chester Kam**	**End Date**	**11 08 31**
Protocol Title	**Work attitudes and behavior**		
Sponsor	**n/a**		

This is to notify you that The University of Western Ontario Department of Psychology Research Ethics Board (PREB) has granted expedited ethics approval to the above named research study on the date noted above.

The PREB is a sub-REB of The University of Western Ontario's Research Ethics Board for Non-Medical Research Involving Human Subjects (NMREB) which is organized and operates according to the Tri-Council Policy Statement and the applicable laws and regulations of Ontario. (See Office of Research Ethics web site: http://www.uwo.ca/research/ethics/)

This approval shall remain valid until end date noted above assuming timely and acceptable responses to the University's periodic requests for surveillance and monitoring information.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the PREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of research assistant, telephone number etc). Subjects must receive a copy of the information/consent documentation.

Investigators must promptly also report to the PREB:
a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) all adverse and unexpected experiences or events that are both serious and unexpected;
c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to the PREB for approval.

Members of the PREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the PREB.

[redacted]
Clive Seligman Ph.D.

Chair, Psychology Expedited Research Ethics Board (PREB)

The other members of the 2010-2011 PREB are: Mike Atkinson (Introductory Psychology Coordinator), David Dozois, Vicki Esses, Riley Hinson Albert Katz (Department Chair), and Tom O'Neill (Graduate Student Representative)

CC: UWO Office of Research Ethics

This is an official document. Please retain the original in your files

CURRICULUM VITAE

CHESTER CHUN SENG KAM

EDUCATION AND DEGREES

Candidate for Doctor of Philosophy (Ph.D.) In Progress
Industrial-Organizational Psychology*
The University of Western Ontario
Supervisor: Prof. John P. Meyer

Master of Philosophy (M. Phil.) 2007
Social and Personality Psychology
The Chinese University of Hong Kong
Supervisor: Prof. Michael H. Bond

Bachelor of Arts Honours (B.A. Hons.) 2003
Psychology (General)
The University of Waterloo
Supervisor: Prof. Michael Ross

ACADEMIC AWARDS

Kam, C. C.-S., Wilson, A., Perunovic, E., Bond, M. H., Zhang, X., & Zhou, X. (2009). *Do Chinese self-enhance? Converging evidence from social comparison and temporal appraisal paradigms.* Paper presented at the 10th Annual Society for Personality and Social Psychology, Tampa, Florida. Won the Student Travel Award Competition. Decisions based on the excellence of the work and on the strength of the applicant's scholarly record.
(Award Value: US$500)

Kam, C. C.-S., & Bond, M. H. (2007). *Do Chinese self-enhance? Evidence from two different approaches.* Poster presented at the Annual Conference of the Hong Kong Psychological Society 2007. Won the J. P. Leung Memorial Award for Best Student Papers, runner-up.
(Award Value: HK$1000 ~ US$128.21)

PUBLICATIONS

Yeung, D. Y., Fung, H. H., & Kam, C. (in press). Age differences in problem solving strategies: The mediating role of future time perspective. *Personality and Individual Differences*.

Kam, C., & Meyer, J. P. (2012). Do optimism and pessimism have different relationships with personality dimensions? A re-examination. *Personality and Individual Differences, 52*, 123-127.

Kam, C. C.-S., & Bond, M. H. (2009). Emotional reactions of anger and shame to the norm violation characterizing episodes of interpersonal harm. *British Journal of Social Psychology, 48*, 203-219.

Kam, C. C.-S., & Bond, M. H. (2008). Role of emotions and behavioural responses in mediating the impact of face loss on relationship deterioration: Are Chinese more face-sensitive than Americans? *Asian Journal of Social Psychology, 11*, 175-184.

Hui, C. M., Lo, I. Y. M., Bond, M. H., & Kam, C. C.-S. (2008). Which aspects of interpersonal experience count in judgments of well-being? *Personality and Individual Differences, 44*, 501-511.

Lee, Y.-Y., Kam, C. C.-S., & Bond, M. H. (2007). Predicting emotional reactions after being harmed by another. *Asian Journal of Social Psychology, 10*, 85-92.

RELEVANT WORK EXPERIENCE

Course Instructor	
Research Methods in Human Resources	Fall 2011
Introduction to Industrial-Organizational Psychology	Summer 2011
Honours Thesis Advisor	
Lisa Plant	2009 - 2010
Lab Instructor	
Psychological Statistics Using Computers	2010 - 2011
Research Methods in Psychology	2009 - 2010
Teaching Assistant	
Introduction to Test and Measurement	2008
Introduction to Personality Psychology	2005-2006