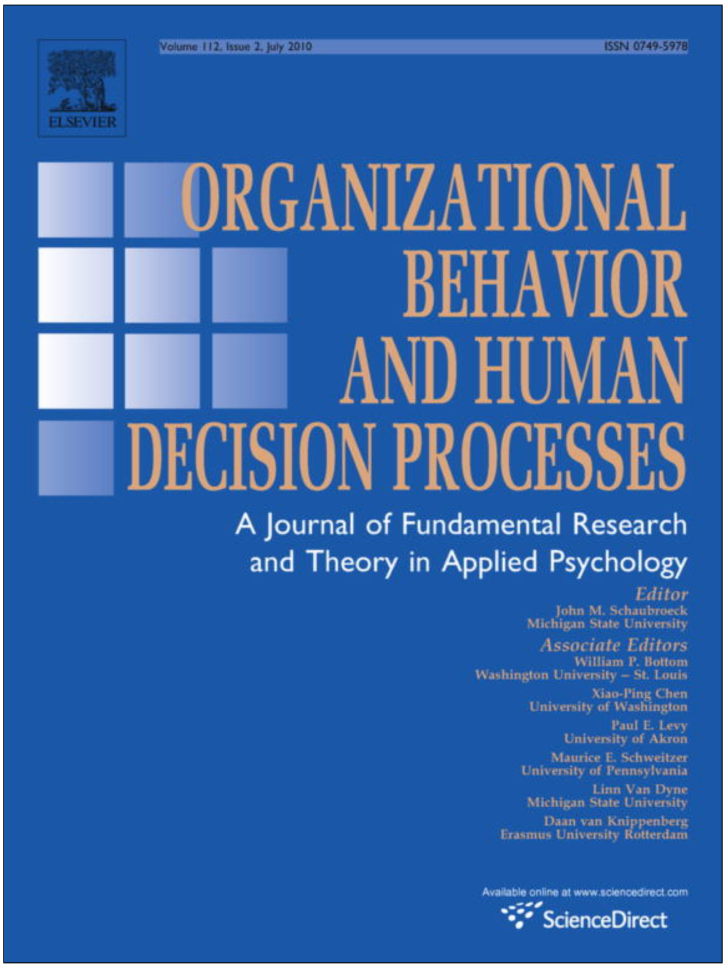
Volume 112, Issue 2, July 2010
ISSN 0749-5978
ELSEVIER
ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
A Journal of Fundamental Research and Theory in Applied Psychology
Editor
John M. Schaubroeck
Michigan State University
Associate Editors
William P. Bottom
Washington University – St. Louis
Xiao-Ping Chen
University of Washington
Paul E. Levy
University of Akron
Maurice E. Schweitzer
University of Pennsylvania
Linn Van Dyne
Michigan State University
Daan van Knippenberg
Erasmus University Rotterdam
Available online at www.sciencedirect.com
ScienceDirect

Organizational Behavior and Human Decision Processes 112 (2010) 112–125

Contents lists available at ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

The problem of empirical redundancy of constructs in organizational research: An empirical investigation

Huy Le [a,*], Frank L. Schmidt [b,c], James K. Harter [c], Kristy J. Lauver [d]

[a] Department of Psychology, University of Central Florida, Orlando, FL 32816-1390, USA
[b] Department of Management and Organizations, University of Iowa, Iowa City, IA, USA
[c] The Gallup Organization, Omaha, NE, USA
[d] Department of Management and Marketing, University of Wisconsin – Eau Claire, Eau Claire, WI, USA

ARTICLE INFO

Article history:
Received 30 December 2007
Accepted 22 February 2010
Available online 19 March 2010

Accepted by Paul Levy

Keywords:
Construct proliferation
Construct empirical redundancy
Discriminant validity
Job attitude
Measurement error correction
Disattenuation
Measurement artifacts
Reliability
Generalizability theory
Structural equation modeling

ABSTRACT

Construct empirical redundancy may be a major problem in organizational research today. In this paper, we explain and empirically illustrate a method for investigating this potential problem. We applied the method to examine the empirical redundancy of job satisfaction (JS) and organizational commitment (OC), two well-established organizational constructs. Analysis based on responses from a sample of 292 employees collected at two occasions showed that: (a) the construct-level correlation between JS and OC was very high (.91) and (b) both JS and OC are similarly related to positive affectivity and negative affectivity. These results suggest that the constructs may be empirically indistinguishable, despite their well-established conceptual distinction. These findings illustrate the problem of empirical redundancy of organizational constructs and provide a basis for a possible movement towards parsimony in the realm of constructs that could open the way to more rapid advances in knowledge in organizational research.

Introduction

Construct proliferation and construct redundancy are major problems today in industrial/organizational psychology, organizational behavior, and other social science areas. At any given time there are numerous constructs—for example, job satisfaction, organizational commitment, and job involvement—that appear to be similar from both a theoretical and an empirical point of view. That is, their theoretical definitions are similar and their observed correlations with each other are substantial. In addition, new constructs similar to existing ones are frequently proposed (the "old wine in new wineskins" phenomenon). Many such constructs may lack discriminant validity relative to other constructs; that is, they may be redundant with existing constructs and thus be examples of construct proliferation. This situation has been a cause of considerable concern (Morrow, 1983; Rousseau, 2007; Schwab, 1980) and can be viewed as a major failure to adhere to the canon of parsimony in science (i.e., failure to apply Occam's Razor). The problem is a serious one because a science that ignores the mandate for parsimony cannot advance its knowledge base and achieve cumulative knowledge. The purpose of science is to uncover the relatively simple deep structure principles or causes that underlie the apparent complexity observed at the surface structure level (Toulmin, 1961), and this is essentially impossible if the mandate for parsimony is not observed.

Schwab (1980) discussed the problem of redundancy of constructs in organizational research. He noted that many constructs hypothesized to be conceptually unique may in fact be empirically redundant, and pointed out that this situation "poses a problem if we take parsimony in scientific explanation seriously" (Schwab, 1980; p. 25). Morrow (1983) specifically highlighted the problem in connection with many forms of the construct "commitment" in the literature (continuance organizational commitment, affective organizational commitment, job involvement, work ethnic endorsement, and career commitment). In fact, it is generally agreed that failure to attend to the redundancy between constructs can result in the proliferation of constructs, hampering the process of systematic and cumulative research (Blalock, 1968; Singh, 1991; Tesser & Krauss, 1976). Nevertheless, the ever-increasing number of new constructs in the literature suggests that it is not simple to deal with this fundamental problem in organizational research.

* Corresponding author.
E-mail address: hale@mail.ucf.edu (H. Le).

0749-5978/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.obhdp.2010.02.003

To be considered distinct, any two constructs must meet two requirements (Singh, 1991). First, they must be conceptually and theoretically distinct. Because of the conceptual/theoretical fluency of researchers, this requirement is essentially a weak one and is usually easily met. For example, it is quite easy to posit a theoretical or conceptual distinction between job satisfaction and organizational commitment (Hulin, 1991; Locke, 1976; Mowday, Steers, & Porter, 1979; Wiener & Vardi, 1980). It is also possible to articulate a theoretical distinction between job satisfaction and job involvement (Lodahl & Kejner, 1965). In fact, the implicit assumption is often that if researchers can make a conceptual, theoretical, or logical distinction between constructs then this distinction will also exist in the minds of employees or survey respondents (Harter & Schmidt, 2008). This assumption may not hold.

The second requirement is that the constructs be empirically distinct. This requirement means that two supposedly distinct constructs should not correlate 1.00 or near 1.00. Since constructs are abstract concepts, they are operationalized via measures in empirical research and correlations between these measures are then used to infer construct-level relationships. However, it is well-known that correlations between measures do not perfectly reflect construct-level relationships because of the biasing effects of measurement artifacts (discussed next). Accordingly, the second requirement can be further explicated such that for two constructs to be considered distinct, their correlation *as estimated from their measures after the downward bias created by measurement artifacts is controlled for* should not be 1.00 (or close to 1.00). This requirement can be viewed as a test of the assumption that the distinction between the constructs exists in the minds of the respondents. That is, data collected from the respondents (i.e., their responses to measures of the constructs) should reflect the distinction between the constructs. If two constructs are correlated at 1.00 (or close to 1.00) they are not distinct in practice, or in other words, they are empirically redundant. This inference (from high construct-level correlation to construct redundancy) is based upon the notion that constructs derive their meanings from (or are defined by) the nomological networks in which they are embedded (Cronbach & Meehl, 1955). Different constructs are expected to be related differently with other constructs/variables in the relevant nomological network. Highly correlated constructs are likely to be similarly related to other variables in the nomological networks, which would mean that the constructs cannot be differentiated based upon their nomological networks. As such these constructs may be considered empirically redundant.

Because of its empirical nature, the second requirement is not easily examined. Use of appropriate methods to correct for the biases induced by measurement artifacts is critical for testing this requirement. Without such corrections, observed correlations between measures of different constructs may be modest (e.g., .60) when in fact the construct-level correlations (i.e., correlations between the constructs underlying the measures after the effect of measurement artifacts is taken into account) are really 1.00 or close to 1.00. Recent advances in models of measurement artifacts and procedures for correcting for the measurement artifacts allow for more accurate estimates of construct-level correlations.

This paper focuses on the second requirement for construct distinction noted above. Specifically, we discuss the problem of empirical redundancy of organizational constructs and describe a method to estimate construct-level relationships based on recent developments in artifact correction methods. We then illustrate an application of the method in examining the empirical redundancy of job satisfaction and organizational commitment, two well-established constructs in organizational research. As such, the method provides a tool needed to investigate the problem of construct proliferation by examining the empirical redundancy between constructs. At the very least we hope this paper will stimulate discussion and debate about the best methods for addressing the serious problem of construct proliferation in our research literatures.

The problem of construct empirical redundancy

As noted earlier, two constructs are considered distinct only if they are both conceptually and empirically non-redundant. Empirical non-redundancy simply means that the constructs can be distinguished based upon empirical data. More specifically, this requirement can be interpreted as saying that the constructs: (a) should not be perfectly (or very highly) correlated with each other and (b) should not have the same patterns of relationships with other variables. The rationale for the first condition is straightforward: constructs are empirically indistinguishable if all (or most) of their variances are common, meaning that people in the population can be ranked similarly on the constructs. As discussed earlier, the second condition is based upon the notion that constructs are defined by the nomological networks to which they belong (Cronbach & Meehl, 1955). If the two constructs are similarly correlated with other variables in a nomological network, their positions/roles in the network cannot be empirically distinguished. Conceivably, two constructs can be defined very differently and clearly specified to hold different positions in a certain nomological network (e.g., job satisfaction may be expected to be causally related to organizational commitment) but they can still be redundant to all intents and purposes if only one construct is sufficient: (a) to capture all the variation attributable to the other construct in the population of interest and (b) to examine the relationship of either construct with other relevant variables in empirical data. In other words, the constructs can be conceptually distinct but empirically redundant. This is a serious problem as there would be no way to empirically disentangle one construct from the other to examine them separately.

It can be seen that the issue discussed here underlies the well-known multitrait–multimethod approach (Campbell & Fiske, 1959) for establishing construct validity of measures. If two measures developed to measure supposedly different constructs are highly correlated, they lack discriminant validity. This would mean that either the measures actually reflect the same construct or the constructs underlying the measures cannot be distinguished by empirical data. Both of these possibilities indicate the problem of empirical redundancy. The corollary of this is that it is not possible to empirically differentiate the roles of these constructs in the relevant nomological networks by using these measures (although they may be conceptually differentiated). Consequently, empirical evidence deemed as supporting effects (or causes) of one construct can also be similarly attributed to the other. The problem may be even more serious if the measures and the constructs in question are well-established. If this is the case, it may be more parsimonious to posit one construct underlying these phenomena instead of using two or more empirically redundant constructs. As such, empirical redundancy has implications on the problem of construct proliferation in research.

The extant literature abounds with examples of constructs which are conceptually distinct but often found to be very highly related in empirical data. For example, Singh (1991) demonstrated that two important attitudinal constructs in research on consumer behavior, "consumer discontent" and "consumer alienation", though conceptually distinct, are actually empirically redundant, as their measures are found to be correlated at 1.00 after removal of biases created by measurement error. This finding advanced the principle of parsimony and resulted in a major restructuring of theories of consumer behavior. More recently, Unsworth and Engle

(2007) examined the constructs of working memory and short term memory, long thought to be distinct constructs based on different neural processes, and showed that their measures are actually perfectly collinear. This finding requires major changes in the understanding of cognitive processes, memory, and general intelligence. (As will be seen later, these findings are unusual in that they reached conclusions of co-linearity based on methods of correcting for measurement error that are not as complete as those presented and illustrated in the present study).

Construct empirical redundancy is obviously an empirical research question and should be answered based upon data. This, however, is not an easy task as constructs are unobservable, so their relationships must be inferred via correlations between measures. It has long been known that there are many sources of variance contributing to the observed variance of measures, in addition to the constructs they are meant to measure (Cronbach, 1947; Cronbach, Glesser, Nanda, & Rajaratnam, 1972; Thorndike, 1949, 1951). Variance due to these sources is measurement artifact variance and its biasing effects must be removed to accurately estimate relationships between constructs based on the measures. To date, commonly accepted solutions for this problem involve using structural equation modeling (*SEM*) and/or confirmatory factor analysis (CFA; Cohen, Cohen, Teresi, Marchi, & Velez, 1990; DeShon, 1998; Hunter & Gerbing, 1982; Marsh & Hocevar, 1988) or the disattenuation formula with reliability estimates (Schmidt & Hunter, 1996, 1999; Thorndike, 1951) to correct for the effects of measurement artifacts in observed correlations between measures. The recent resurgence of interest in the multiple sources of measurement artifacts suggests that current methods of estimating construct-level relationships are deficient because several important sources of measurement artifacts that contribute to the variance of the measures are not taken into account (Becker, 2000; DeShon, 1998; Le, Schmidt, & Putka, 2009; Schmidt & Hunter, 1999; Schmidt, Le, & Ilies, 2003). As explained more fully later, transient measurement error and scale specific factor error are not accounted for by either coefficient alpha, the most frequently used reliability coefficient, or by the most common application of CFA and *SEM* (DeShon, 1998; Le et al., 2009; Schmidt et al., 2003). Transient error, resulting from moods, feelings, or mental states that are *specific to an occasion* (Cronbach, 1947; DeShon, 1998; Le et al., 2009; Schmidt et al., 2003; Thorndike, 1951), is inherent in measures of any construct that is defined to be temporally stable across even short time periods. Just as item specific factor error exists within a scale, scale specific factor error exists across different measures of a construct. If not taken into account, these measurement artifacts will create biases in the estimated relationships between constructs underlying the measures. Psychological and organizational theories are meant to explain relationships among constructs, not among measures, and so such biased estimates of construct-level relationships may have important consequences. In particular, constructs may be concluded to be distinct from one another when they are not and in fact are empirically redundant.

In this paper we apply the CFA-based procedure suggested by Le et al. (2009) that accounts for the effects of all major sources of measurement artifacts in self-report measures of organizational constructs to estimate the construct-level relationship between job satisfaction and organizational commitment. Since these are arguably two of the most well-established constructs in organizational research, there should be no question about their conceptual distinction. Thus, examining the relationship between job satisfaction and organizational commitment would allow us to focus solely on the issue of empirical redundancy. Given the role of the constructs in organizational research and practice, this investigation can potentially have important implications. Conceivably, if job satisfaction and organizational commitment are found to be empirically redundant, their distinction and consequently unique contributions to organizational research will be questioned despite the indisputable fact that they are conceptually distinct. That finding could potentially require revisiting our understanding of many organizational theories and practices involving the constructs.

Measurement artifacts and their effects on estimated construct-level relationships

In this section, we briefly review the major sources of measurement artifacts in self-report measures and discuss how they create bias in observed correlations between measures. We also provide an overview of the estimation procedures presented by Le et al. (2009). Further details of the procedure and how it is applied in the current research are described in "*Methods*" section.

Traditional approaches

According to classical measurement theory, the observed variance of scores on a measure is the sum of true score variance and measurement error variances (Lord & Novick, 1968). Measurement errors in self-report measures include random response error, item specific factor error, and transient error (Cronbach, 1947; Schmidt et al., 2003; Thorndike, 1949). The psychological processes that create these measurement errors have been explained at length in psychometric textbooks (e.g., Thorndike, 1949, 1951) and more recently, by Schmidt & Hunter (1999), Schmidt et al. (2003). These measurement errors create downward bias in the observed correlations between scores on the measure and measures of other variables (Ree & Carretta, 2006; Schmidt & Hunter, 1996; Thorndike, 1951).

Traditionally, correction for the bias is made either by applying *SEM* (or CFA) or by using a reliability coefficient with the disattenuation formula (DeShon, 1998; Schmidt et al., 2003). The former approach generally involves splitting each measurement scale into several subscales (item parcels) and using these subscales as indicators for the latent factor representing the construct of interest. For example, in several studies examining the discriminant validity of measures of job attitudes, researchers split each measure of job satisfaction, organizational commitment, and job involvement into three parts to represent the underlying constructs in their *SEM* and/or CFA models (e.g., see Brooke, Russell, & Price, 1988; Mathieu & Farr, 1991; Nystedt, Sjöberg, & Hägglund, 1999). As such, the constructs (latent factors) are defined as shared variance among the indicators (i.e., subscales or item parcels) measuring the same construct. The classical disattenuation approach typically uses coefficient alpha, the most frequently used index of reliability (Schmidt et al., 2003). Conceptually, the two approaches are equivalent because they both account for two major sources of measurement errors: random response error and item specific factor error (Le et al., 2009). Transient error, however, is ignored in these approaches (DeShon, 1998; Schmidt et al., 2003), as is scale specific factor error, resulting in underestimation of construct-level relationships.

More complete conceptualization of measurement artifacts

Based on generalizability theory (Cronbach et al., 1972), Le et al. (2009) pointed out that apart from the construct that a measure (scale) is meant to capture, each measure contains specific factors that contribute to the variance of its true score. Such scale specific factors arise from the specific, idiosyncratic way that the measure operationalizes the theoretical construct. The factors can be sampling-based (e.g., idiosyncratic selection of items from content domains) or methodological (e.g., scale formats or measurement methods) in nature, but they are *not relevant to, or part of*, the construct measured by the scale. Different measures of the same construct contain different specific factors that are unrelated to the construct or to each other. These factors function the same way as do traditional measurement errors in that they create downward bias in observed correlations between measures. Specific factors in a measure (scale) are directly analogous to item specific factor error inherent in each item in classical measurement theory. The former is the specificity of a scale and is irrelevant to the construct of interest whereas the latter is the specificity of an item and is irrelevant to the true scores underlying that scale. Accordingly, Le et al. (2009) referred to these factors as scale specific factor error. Together with the measurement errors identified under classical measurement theory (i.e., random response error, transient error, and item specific factor error), scale specific factor error is a type of measurement artifact that biases observed correlations between measures and therefore should be accounted for so that construct-level relationships can be accurately calibrated.

This more complete conceptualization of measurement artifacts is conceptually identical to "best practice" applications of *SEM* and CFA, in which different measures of the same construct [instead of subscales (or item parcels) of the same measure] are used as indicators for the construct in *SEM* or CFA (Le et al., 2009). This usage of *SEM* and CFA should be familiar to organizational researchers because the literature includes studies that followed the *SEM* "best practice" and operationalized constructs using multiple measures (e.g., the multiple measures of job satisfaction used in Hom & Griffeth, 1991).

Procedures for estimating construct-level relationships

Le et al. (2009) presented two procedures for estimating the relationships between constructs based on the observed correlations between their measures. The first procedure is based on the generalized coefficient of equivalence and stability (GCES). This coefficient is analogous to the coefficient of equivalence and stability (CES), the most appropriate reliability coefficient for use in correcting for measurement error under classical measurement theory (Schmidt et al., 2003; Thorndike, 1949, 1951). The CES of a measure is estimated by correlating two classically parallel forms administered on two different occasions, while the GCES is estimated by correlating the measure with another measure (or other measures) of the same construct which is administered on a different occasion. As such, the GCES defines the construct as what is shared across occasions (times) by different *measures* developed to assess the same theoretical construct. This is analogous to the CES under classical measurement theory which defines the true score as what is shared by different *items* of a measure across different *times*. Like the CES which indicates the proportion of the observed variance of a measure due to the true score, the GCES reflects the proportion of the observed variance due to the construct in question. When the GCES is used in the disattenuation formula, it allows us to "partial out" all the effects of measurement artifacts on the observed correlation between measures, resulting in an unbiased estimate of the relationship between the constructs underlying these measures (Schmidt & Hunter, 1996). The second procedure presented by Le et al. (2009) is based on CFA. It requires that the multiple indicators for a latent variable (construct) be different measures of the same theoretical construct and that they be administered on different occasions. Construct-level relationships can then be obtained by allowing the correlation between latent factors representing the constructs to be freely estimated (cf. Marsh & Hocevar, 1988). Le et al. (2009) demonstrate that this procedure is conceptually equivalent to the GCES procedure and they present computer simulation studies showing that the two procedures produce the same estimates.

Measurement artifacts in measures of job attitudes

Measures of organizational constructs in general, and job attitudes in particular, are affected by a number of measurement artifacts. The existence of random response error and item specific factor error in these measures is well accepted, as seen in the fact that the widely used coefficient alpha takes these two forms of measurement error into account. Transient error and scale specific factor error, however, are less often recognized by organizational researchers. As noted earlier, transient error results from mental states that are specific to an occasion and thus exists in measures of any construct that is defined to be temporally stable. Job attitudes, stemming from employees' reactions to experiences at work, should be relatively stable as long as the job and the factors surrounding the job are stable (Harrison & Martocchio, 1998). Accordingly, individuals' standing on these constructs should not be affected by job-irrelevant sources specific to a certain occasion (such as transient mood or a temporary health-related issue such as a cold); any variation due to these sources in the measure should be treated as error. Scale specific factor error is also relevant to measures of job attitudes. As shown by Le et al. (2009), the concept of scale specific factor error is implicit in: (a) the logic of the multitrait–multimethod approach for establishing construct validity (Doty & Glick, 1998) and (b) the CFA and *SEM* practice of using different measures of the same theoretical construct as indicators for that construct (cf. Hom & Griffeth, 1991).

Traditional procedures for estimating construct-level relationships do not account for all the sources of measurement artifacts. Construct-level correlations among job attitude constructs estimated by these procedures are likely to be distorted, and conclusions in the literature based on such distorted estimates may be erroneous. The distortion is generally a downward bias, but can also be an upward bias in some situations due to correlated transient errors, as noted in Le et al. (2009). In the case of job attitudes, we hypothesize that the overall (net) bias is in the downward direction. That is, current estimates of construct-level relationships among job attitudes are likely to be lower than their actual values. Our expectation in this respect stems in part from the research on discriminant validity of job attitude constructs, typically job satisfaction, organizational commitment, and job involvement (Brooke et al., 1988; Mathieu & Farr, 1991). Previous studies empirically examining the question (Brooke et al., 1988; Mathieu & Farr, 1991; Nystedt et al., 1999) did not fully account for the effects of measurement artifacts. These studies concluded that the job attitudes examined are empirically distinct, but this conclusion should be critically re-examined. It is possible that the construct-level relationships between job attitude constructs are considerably higher when all measurement artifacts are appropriately accounted for.

Are organizational commitment and job satisfaction empirically redundant?

Organizational commitment and job satisfaction relationship

Locke (1976) defined job satisfaction as an emotional state resulting from the evaluation of one's job experiences. Roznowski and Hulin (1992) state that job satisfaction "accounts for variance in organizationally relevant responses far beyond the demonstrated usefulness of the newer and trendier constructs, notions and variables." (p. 124). Organizational commitment is generally defined as attitude toward, or loyalty to, the employing organization (Price, 1997). Both job satisfaction and organizational commitment therefore can be considered general affective responses to aspects of the work environment (Hulin, 1991). For the former, the target is the jobs, whereas it is the employing organizations for the latter. These two work-related attitudes reflect individuals' fundamental evaluation of their work experiences (Harrison, Newman, & Roth, 2006).

Though the constructs are conceptually distinct, empirical evidence indicates that measures of these constructs are highly correlated (Cooper-Hakim & Viswesvaran, 2005; Griffeth, Hom, & Gaertner, 2000; Meyer, Stanley, Herscovitch, & Topolnytsky, 2002). Further, both job satisfaction and organizational commitment appear to have similar dispositional determinants [e.g., positive affectivity and negative affectivity (Thoresen, Kaplan, Barsky, Warren, & Chermont, 2003); and affective disposition (Bowling, Beehr, & Lepisto, 2006)] and similar outcomes [e.g., turnover (Griffeth et al., 2000); and organizational citizenship behaviors (Organ & Ryan, 1995)]. Their patterns of relationships with other variables are also very similar (Brooke et al., 1988; Harrison et al., 2006). These findings have led to questions about the empirical redundancy of the constructs.

Also relevant here is the unsettled debate about the causal relationship between job satisfaction and organizational commitment. Most researchers believe that job satisfaction leads to organizational commitment because the former is considered a more immediate affective response to one's work which can be established shortly after joining an organization, whereas the latter is likely to develop more slowly over time since it is based not only on the job but also on other aspects of the organization, such as its goals and values (Cramer, 1996; Porter, Steers, Mowday, & Boulian, 1974). Others support the opposite order of causality based on self-perception theory which suggests that higher organizational commitment results in greater job satisfaction because organizational commitment may stimulate a rationalization process through which attitudes are made consistent with behavior (Bateman & Strasser, 1984; Cramer, 1996). Empirical evidence from studies directly examining the causal relationship is mixed. Some studies found support for the hypothesis that job satisfaction causes organizational commitment (Rusbult & Farrell, 1983; Williams & Hazer, 1986), whereas others supported the opposite causal ordering (Bateman & Strasser, 1984; Vandenberg & Lance, 1992). Yet other studies concluded that the relationship is spurious (Cramer, 1996; Curry, Wakefield, Price, & Mueller, 1986) or reciprocal (Farkas & Tetrick, 1989). These inconsistent findings further raised questions about the empirical redundancy of the constructs (Brooke et al., 1988; Harrison et al., 2006).

The studies that empirically investigated the discriminant validity of job satisfaction and organizational commitment have concluded that, while correlated, these constructs are not redundant (Brooke et al., 1988; Mathieu & Farr, 1991; Nystedt et al., 1999). Using confirmatory factor analysis to control for the effect of measurement error, Brooke et al. (1988) concluded that the latent constructs measured by a job satisfaction scale (Brayfield & Rothe, 1951) and a widely used organizational commitment scale (Organizational Commitment Questionnaire; Mowday et al., 1979) were empirically distinguishable, with the estimated correlation between the constructs being .55. This conclusion was also reached by Mathieu and Farr (1991) and Nystedt et al. (1999), who used different measures of job satisfaction as well as very different samples. Mathieu and Farr provided two estimates of the correlation between the constructs, .78 and .70, while Nystedt et al. produced an estimate of .59.

These studies failed to account for two important sources of measurement artifacts, transient error (Becker, 2000; DeShon, 1998; Schmidt et al., 2003) and scale specific factor error (Le et al., 2009; Schmidt & Hunter, 1999). It is therefore possible that a different conclusion would be reached if all the sources of measurement artifacts were appropriately taken into account. In the current study, we employ a special design (described below) that allows appropriate estimation of the construct-level relationship between measures of organizational commitment and job satisfaction to investigate this question. In addition, we are also interested in examining whether the constructs are differentially related with other variables. As discussed earlier, the patterns of correlations between the constructs and other variables in their nomological network can help determine whether the problem of construct empirical redundancy exists. Accordingly, in the current study we further examine how the constructs underlying measures of job satisfaction and organizational commitment are related with the constructs of affectivity, which have generally been found to be important determinants of organizational attitudes.

Affectivity and organizational attitudes

Trait negative affectivity (NA) and trait positive affectivity (PA) are dispositional constructs (traits) that reflect a person's general tendency to experience certain affective states (positive or negative; Judge & Larsen, 2001; Thoresen et al., 2003; Watson, Clark, & Tellegen, 1988). More specifically, PA is characterized by high energy, enthusiasm, and pleasurable engagement, whereas NA is characterized by distress, unpleasurable engagement, and nervousness (Judge & Larsen, 2001; Watson et al., 1988). These two constructs are considered dispositional determinants of organizational attitudes because they influence the way people respond to organizational and job stimuli (Judge & Larsen, 2001; Thoresen et al., 2003). As such, PA and NA hold important positions in the nomological network around organizational attitudes. It is thus important to accurately determine the construct-level relationships among these dispositional and other organizational constructs.

Recent meta-analysis shows that PA and NA are similarly related to both job satisfaction and organizational commitment (Thoresen et al., 2003): the correlations between PA and job satisfaction and organizational commitment are estimated to be .34 and .35, respectively; correlations between NA and these two organizational constructs are −.34 and −.27, respectively. These estimates, however, are likely to be biased because the effects of measurement artifacts, specifically those of transient error and scale specific factor error, have not been fully accounted for. In the current study, we apply the method introduced by Le et al. (2009) to estimate the construct-level relationships among PA, NA, job satisfaction, and organizational commitment. As discussed earlier, these estimates will allow us to critically examine the empirical distinction between the two organizational constructs.

Methods

The two approaches (Le et al., 2009) described earlier were applied to estimate the relationships between the constructs underlying widely used measures of PA, NA, job satisfaction, and organizational commitment. To save space, we only describe the CFA-based approach here (both approaches yielded the same results; details regarding the GCES approach are available from the authors upon request). As demonstrated later, this approach allows simpler estimation of different sources contributing to the variance of an item in a measure. The approach requires: (a) that there be different measures for the same construct and (b) that the measures be administered to the same sample of subjects on different occasions with relatively short intervals (so that any changes in the subjects' responses to the measures are due to transient error and not to real changes in construct scores).

Procedure and sample

Data for the current study were obtained through the Gallup Organization. A random sample of employed adults from the

Gallup Panel (a probability-based, nationally representative panel of US households) was invited to participate in the study, which required responding to two online surveys with a 1-week interval intervening. The surveys include measures of job attitudes and affectivity (described next) and other demographic information. The order of the measures in the surveys was rotated to create two forms. The forms were administered to the participants such that no participant received the same form on both occasions. Responses from 399 participants were available and could be matched across two occasions. This sample includes 49.0% females (195) and 51.0% males (203) with the mean age of 48.02 (*SD* = 10.54). Most of the participants hold professional (46.3%) or managerial positions (17.0%) in their organizations. The remaining participants hold either clerical (7.6%), service (6.6%), or sales (5.1%) jobs. As described later, we only used a subset of this sample in our study.

Measures

Job satisfaction (JS)

The calibration for scale specific factor measurement artifact requires agreement about the theoretical meaning of a construct underlying different measures developed to operationalize that construct. With its long history in the research literature, job satisfaction is a very well-established construct, and there does appear to be such agreement. Despite some disagreements about nuances in the theoretical underpinnings of different job satisfaction measures (e.g., Brief & Roberson, 1989; Scarpello & Campbell, 1983), it is likely that the same general construct of job satisfaction underlies all these measures. Empirical support for this conclusion is presented in Le et al. (2009). In the present study, we measured job satisfaction using two scales. The first scale was the Hoppock's job satisfaction scale (Hoppock, 1935), which is an established measure of job satisfaction frequently used in organizational research (cf. Cook, Hepworth, Wall, & Warr, 1981). The scale consists of four items, each with seven response options. The items ask employees' feelings about their jobs *in general*. As such, Hoppock's scale is a global measure of overall job satisfaction. The second scale was the Minnesota Satisfaction Questionnaire (MSQ; Weiss, Dawis, England, & Loftquist, 1967). We used the MSQ short form with 20 items in the current study. Unlike the Hoppock's scale, the MSQ includes items requesting respondents to indicate how satisfied they are with different specific aspects of their jobs. These job aspects can be classified into "intrinsic" and "extrinsic", so MSQ items can be combined to measure intrinsic and extrinsic job satisfaction constructs, respectively. Overall job satisfaction is measured by combining all 20 items. This scale was previously used in Nystedt et al. (1999), which replicated the findings of Brooke et al. (1988).

Organizational commitment (OC)

There may be less conceptual agreement among different measures for OC. This construct has been conceptualized somewhat differently by various researchers. Porter and colleagues (1974) defined OC as the individual's affective response to the organization, including his/her identification and involvement with the organization. The Organizational Commitment Questionnaire (OCQ; Mowday et al., 1979) was developed to operationalize the construct and has been the most popular measure of OC in the literature. Later, however, Allen and Meyer (1990) suggested that OC entails three components: affective, continuance, and normative. Out of these, only affective commitment is conceptually similar to the OC construct underlying the OCQ. Accordingly, in the current study, we used the 9-item short form (Curry et al., 1986) of the OCQ (Mowday et al., 1979) and the 8-item scale of affective commitment from the Allen and Meyer's measure (1990). All the items of the measures were answered using a 5-point Likert scale response format.

Positive and negative affectivity

Negative affectivity (NA) and positive affectivity (PA) were measured by the Positive Affect Negative Affect Schedule (PANAS; Watson et al., 1988), which is one of the most popular measures for these constructs (Price, 1997). The PANAS includes 20 adjectives (10 for PA and 10 for NA) describing various affective states; respondents were asked to indicate how they *typically* experienced these states using a scale ranging from 1 ("*Very slightly or not at all*") to 5 ("*Very much*"). In addition, we used the Multidimensional Personality Index (MPI; Watson & Tellegen, 1985) as the second measure for the affectivity constructs. The MPI is based on the same conceptualization of the PA and NA constructs as the PANAS and has been used in a number of past studies (Agho, Price, & Mueller, 1992; Schmidt et al., 2003). It includes 22 statements (11 for each affectivity construct) with response options based on a 5-point Likert scale ranging from "*Not at all characteristic of me*" to "*Very much characteristic of me*".

Additional questions

Apart from the job attitudes and affectivity measures, the surveys also included questions on respondents' employment conditions, demographic information, and other details not directly related to the purpose of the current study. Among such questions, one item asked about the respondents' perceived change in work conditions: "*In the past week, did anything significant happen at work that affected how you view the quality of your work life, or not?*" Answers to the question are either "*Yes*" or "*No*". As described next, this item was used to select participants for our analysis.

Analysis

Selecting participants

As noted earlier, there is concern that any changes observed in people's responses to job attitudes measures, as compared to their earlier responses, could be due to either real changes in the job attitudes or transient error. Because job attitudes reflect people's reactions to job experiences, it is reasonable to believe that the constructs have not changed if job experiences remain unchanged. Accordingly, in the current study, we included only those respondents who responded "*No*" to the question directly asking about changes in their work environments during the interval between the two measurement administrations. As a result, we can be reasonably certain that the levels of job attitudes for these respondents have not changed during the period of the study. Out of 399 participants, 107 answered "Yes" to the question and were therefore excluded from the study. Thus, our data included 292 participants who indicated that there was no change in their work conditions during the period between the two measurement administrations. The final sample is very similar to the original sample of 399 participants in terms of demographic makeup [46.8% females (137) and 52.9% males (155); mean age = 48.12 (*SD* = 10.65)].

CFA procedure

As discussed earlier, Le et al. (2009) suggested that there are five sources that contribute to the observed variance of an item of a measure (Eq. (1), p. 167):

$$\mathrm{Var}(X) = \mathrm{Var}(p) + \mathrm{Var}(po) + \mathrm{Var}(pi:s) + \mathrm{Var}(ps) + \mathrm{Var}(e). \tag{1}$$

In the above equation, Var(X) is observed variance, Var(p) is variance due to the construct of interest, Var(po) is transient error variance, Var(pi:s) is item specific factor error variance, Var(ps) is scale specific factor error variance, and Var(e) is random response error variance (these notations follow the generalizability theory's conventions). These sources of variance can be specified as latent factors in CFA models (Le et al., 2009).

In the current paper, we follow the procedure described above to model the sources of variances in measures of JS, OC, PA, and NA. However, instead of using items as indicators, we first created subscales (item parcels) for our analysis. This step is necessary to: (a) increase the ratio of sample size to number of parameter estimates and (b) create relatively interval and normally distributed indicators. These conditions (i.e., high ratio of sample size to parameter estimates and interval indicators) are needed to ensure the accuracy of the maximum likelihood estimation procedure used in CFA (Bargozzi & Edwards, 1998; Hau & Marsh, 2004; Jackson, 2003; Marsh & Hocevar, 1988). To create the subscales, we split each scale into several item parcels. Specifically, for the MSQ we created four subscales (two with six items each representing intrinsic satisfaction and two with four items each for extrinsic satisfaction). Two subscales (each with two items) were created for the Hoppock, three for the OCQ (each with three items), and three for the Allen and Meyer's scale (two with two items and one with four items; all the negatively worded items were combined together into one subscale). With the PANAS, we created six subscales, three for the PA (including two subscales with three items and one with four items) and three for the NA (also two with three and one with four items). Finally, the MPI was also split into six subscales with three for the PA (two with four items and one with three items) and three for the NA (same as the PA). In total, we created 24 subscales from the original eight measures. Scores for all the subscales were available for both occasions, so the total number of observed indicators for the JS, OC, PA, and NA constructs in our analysis is 48.

For each indicator, we specified three latent factors representing: (a) the construct it was meant to measure (JS, OC, PA, or NA), (b) scale specific factor error, and (c) transient error, respectively. The residual of each indicator includes both random response error and item specific factor error, with the latter being modeled as the correlation between two residuals of the same subscale across occasions (see Le et al. (2009) for details). In total, there are four latent variables representing the constructs (OC, JS, PA, and NA); eight latent variables representing eight scale specific factor errors for the measures (MSQ, Hoppock, OCQ, Allen and Meyer, the PA and NA scales of the PANAS, and the PA and NA scales of the MPI); and eight latent factors representing transient errors (at each occasion there were four transient errors, one for each construct). The latent variable representing the JS construct was specified to underlie (i.e., to be causally related to) all 12 subscales for measures of JS in both occasions. Similarly, the latent variables for OC, PA, and NA were specified to underlie the subscales for measures of OC, PA, and NA, respectively. All subscales belonging to a measure (e.g., the four subscales of the MSQ) at both occasions were specified to have the same latent variable representing the scale specific factor error of that measure. Finally, all subscales for a construct at one occasion shared the same latent factor representing the transient error for that occasion [e.g., T1(JS) is the transient error for all six subscales of JS at Time 1]. All loadings of a subscale in one occasion are constrained to be equal to the corresponding loadings of the same subscale on the other occasion (because it is theoretically expected that the psychometric properties of a measure do not change across occasions).

Fig. 1 shows the latent factors and how they are related to the subscales (indicators) and to one another. As can be seen, we allowed the four latent factors representing the constructs (JS, OC, PA, and NA) to be correlated. These correlations provide the estimates of the construct-level relationships among these constructs, after the biasing effects of measurement artifacts were removed. Finally, latent factors representing transient errors for measures of the constructs at the same occasion (e.g., Time 1) were allowed to be correlated with each other. These specifications were needed to address the potential problem of correlated transient errors (Le et al., 2009).

To investigate the potential problem of construct empirical redundancy, we examined three hierarchically nested models. These models are different in the extent to which the correlations among the latent variables representing the constructs were constrained. In the first model (Model 1), all the correlations were allowed to be freely estimated (i.e., no constraint). In the second model (Model 2), the correlation between JS and PA was constrained to be the same as the correlation between OC and PA. Similarly, correlations between JS and NA and between OC and NA were constrained to be the same. As such, Model 2 specifies that JS and OC have the same pattern of relationships with PA and NA, signifying that the constructs may be empirically redundant (the plausibility of this conclusion would also depend on how high the correlation between JS and OC is estimated by the model). Finally, the third model (Model 3) is the same as Model 2 except the correlation between JS and OC was constrained to be 1.00. Model 3 thus specifies that JS and OC are perfectly indistinguishable, indicating the problem of empirical redundancy. To compare the models, we looked at differences in fit indexes. Although the χ^2 difference is commonly used to compare hierarchically nested models, it is susceptible to the same problem as the χ^2 used in examining CFA and *SEM* model fit in general (that is, it is heavily influenced by sample size; Brannick, 1995). Accordingly, we also examined another index, difference in the comparative fit index (CFI), to compare these models. In the context of testing measurement invariance in multi-group CFA and *SEM*, several researchers have suggested using the difference in CFI between hierarchically nested models for model selection (Cheung & Rensvold, 2002; Meade, Johnson, & Braddy, 2008). Although the hierarchically nested models examined in the current study do not directly pertain to the issue of measurement invariance, we employ the cut-off value of $-.002$ suggested by Meade et al. (2008) to aid our model comparison process. These authors suggest that the difference between the CFI of the nested model and the original model should not be smaller than $-.002$; that is, the absolute value of the negative difference should not exceed .002. We used SAS Proc CALIS (SAS 9.1) with the maximum likelihood estimation method to analyze data for the models.

Results

Table 1 shows the observed correlations among the measures used in the study and their reliabilities (internal consistency as indexed by coefficient alpha). As can be seen, all the measures have high internal consistencies, ranging from .87 to .93. Traditionally, these values would indicate that the measures were highly reliable and thus correlations among the constructs underlying the measures would not be seriously attenuated due to measurement error. However, this traditional view of measurement error is inadequate and likely to lead to underestimation of construct-level relationships. As noted earlier, results from the CFA models, which take into account all major sources of measurement artifacts, provide more accurate estimates of the relationships among the constructs.

Model 1 showed reasonably good fit: $CFI = .948$, $RMSEA = .051$, $SRMR = .052$, $\chi^2 = 1726.49$, $df = 1042$ ($p < .01$). In Table 2, the values above the diagonal and outside the parentheses are the construct-level correlations estimated by the model. Compared to the observed correlations (the values below the diagonal) and correlations corrected by using the coefficients alpha (above the

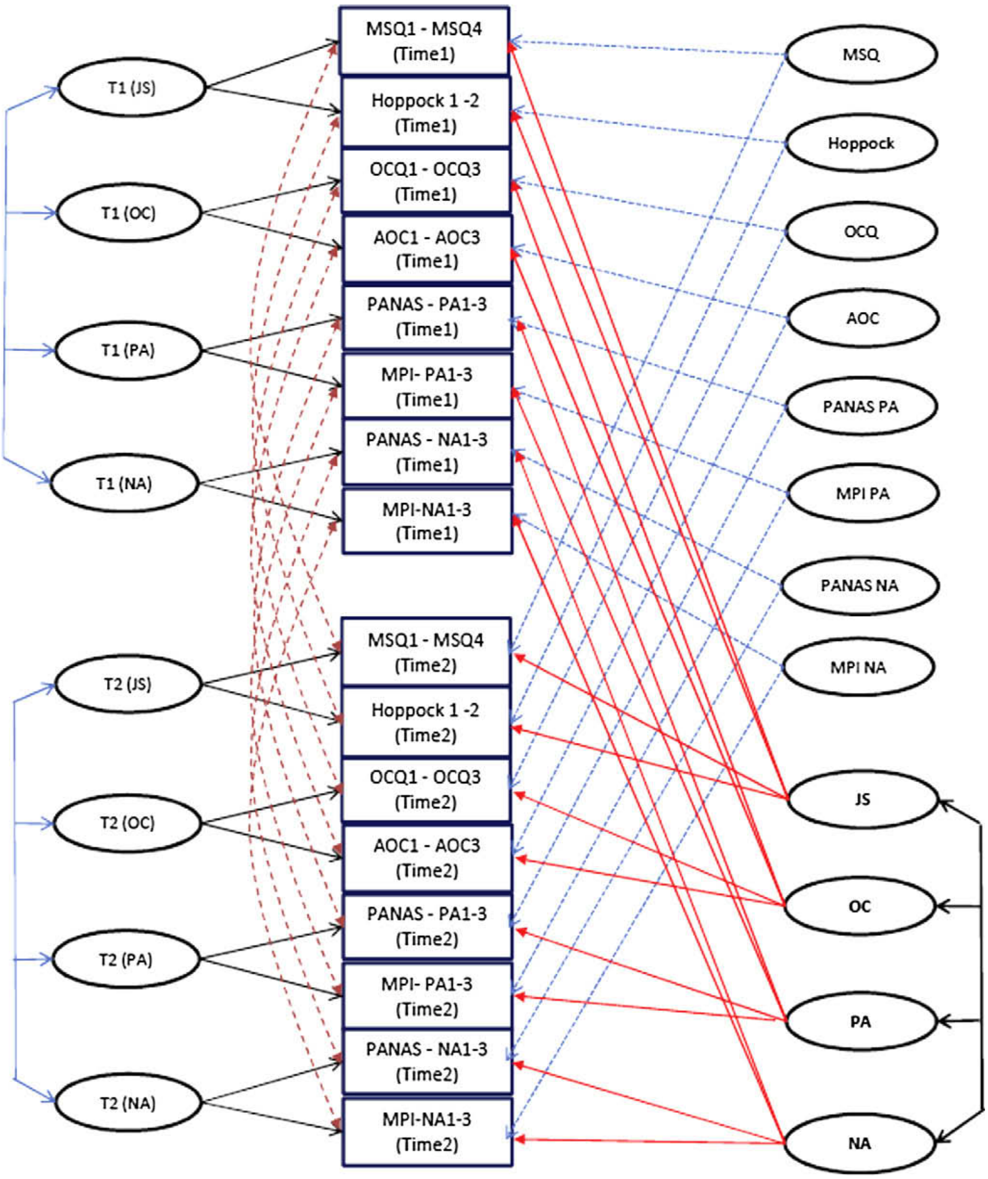

Fig. 1. The CFA model for estimating the construct-level relationships among JS, OC, PA, and NA. *Figure notes*. MSQ1–MSQ4 (Time 1) = four subscales of the MSQ at Time 1. Hoppock 1–2 (Time 1) = two subscales of the Hoppock's scale at Time 1. OCQ1–OCQ3 (Time 1) = three subscales of the OCQ at Time 1. AOC1–AOC3 (Time 1) = three subscales of the Allen and Meyer's Affective Commitment at Time 1. PANAS–PA1–3 (Time 1) = three subscales of the PANAS for positive affectivity at Time 1. MPI–PA1–3 (Time 1) = three subscales of the MPI for positive affectivity at Time 1. PANAS–NA1–3 (Time 1) = three subscales of the PANAS for negative affectivity at Time 1. MPI–NA1–3 (Time 1) = three subscales of the MPI for negative affectivity at Time 1. MSQ1–MSQ4 (Time 2) = four subscales of the MSQ at Time 2. Hoppock 1–2 (Time 2) = two subscales of the Hoppock's scale at Time 2. OCQ1–OCQ3 (Time 2) = three subscales of the OCQ at Time 2. AOC1–AOC3 (Time 2) = three subscales of the Allen and Meyer's Affective Commitment at Time 2. PANAS–PA1–3 (Time 2) = three subscales of the PANAS for positive affectivity at Time 2. MPI–PA1–3 (Time 2) = three subscales of the MPI for positive affectivity at Time 2. PANAS–NA1–3 (Time 2) = three subscales of the PANAS for negative affectivity at Time 2. MPI–NA1–3 (Time 2) = three subscales of the MPI for negative affectivity at Time 2. = effect of a construct (JS, OC, PA, or NA) on a subscale. = effect of a scale specific factor error (e.g., MSQ, Hoppock, OCQ) on a subscale. = effect of a transient error on (Time 1 or Time 2) on a subscale. = correlation between residuals of the same subscale across occasions representing the item specific factor error.

diagonal and in parentheses), these correlations are substantially larger (in absolute value). These results highlight the problem created by ignoring or failing to take into account the effects of all the measurement artifacts in self-report measures in research. Of special interest is the estimated construct-level correlation between JS and OC, which is .91. This correlation is much higher than those previously estimated (cf. Brooke et al., 1988; Cooper-Hakim & Viswesvaran, 2005; Mathieu & Farr, 1991; Nystedt et al., 1999) and thus casts doubt on the empirical distinction between the constructs underlying these measures.

Model 2, which constrains that JS and OC have the same patterns of correlations with PA and NA, also fits the data well: *CFI* = .947, *RMSEA* = .051, *SRMR* = .058, χ^2 = 1744.27, *df* = 1044 ($p < .01$). The Comparisons between Model 2 and Model 1 are shown in Table 3. As can be seen, compared to the fit indices of Model 1, the χ^2 difference is statistically significant ($\Delta\chi^2 = 17.77$, $df = 2$, $p < .01$) but the absolute difference in *CFI* ($|.947 - .948| = .001$) is smaller than the preset cut-off value of .002. As such, it seems that Model 2, which is more parsimonious, may reflect the actual relationships among the constructs. This would mean that JS and OC are not only highly correlated (.91) with each other but also have the same pattern of relationships with PA and NA. Taken together, this finding suggests that these constructs may be empirically redundant.

Table 1
Descriptive statistics and observed correlations.

		M	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	OCQ1	33.93	7.13	*.92*															
2	A&M AOC1	26.85	6.76	.77	*.88*														
3	MSQ1	75.53	13.57	.74	.72	*.92*													
4	Hoppock1	20.36	4.12	.69	.70	.76	*.87*												
5	PANAS.P1	36.31	6.44	.40	.41	.43	.47	*.90*											
6	MPI.P1	38.72	8.28	.26	.28	.37	.39	.70	*.91*										
7	PANAS.N1	17.12	6.03	−.27	−.26	−.35	−.42	−.29	−.31	*.90*									
8	MPI.N1	26.29	8.93	−.18	−.18	−.35	−.34	−.39	−.43	.65	*.91*								
9	OCQ2	31.31	7.61	.82	.74	.69	.70	.45	.33	−.28	−.20	*.93*							
10	A&M AOC2	26.53	6.63	.67	.80	.63	.64	.45	.33	−.28	−.19	.78	*.88*						
11	MSQ2	74.71	13.40	.64	.63	.86	.71	.45	.39	−.32	−.32	.73	.68	*.92*					
12	Hoppock2	20.28	3.89	.63	.66	.71	.87	.47	.34	−.41	−.32	.69	.69	.74	*.87*				
13	PANAS.P2	35.89	7.00	.36	.36	.43	.46	.72	.71	−.33	−.40	.43	.45	.50	.47	*.91*			
14	MPI.P2	38.53	8.02	.28	.29	.37	.42	.67	.90	−.35	−.43	.36	.36	.44	.40	.76	*.92*		
15	PANAS.N2	15.87	5.99	−.16	−.16	−.30	−.32	−.36	−.40	.73	.70	−.22	−.21	−.34	−.35	−.35	−.39	*.91*	
16	MPI.N2	26.60	8.51	−.14	−.13	−.27	−.26	−.37	−.40	.60	.88	−.17	−.16	−.28	−.28	−.37	−.40	.73	*.90*

Notes. N = 255–270. Coefficients alpha are on the diagonal. OCQ1 = Organizational Commitment Questionnaire at Time 1. A&M AOC1 = Allen and Meyer's Affective Commitment at Time 1. MSQ1 = Minnesota Satisfaction Questionnaire at Time 1. Hoppock1 = Hoppock job satisfaction at Time 1. PANAS.P1 = PANAS Positive Affectivity at Time 1. MPI.P1 = Multidimensional Personality Index Positive Affectivity at Time 1. PANAS.N1 = PANAS Negative Affectivity at Time 1. MPI.N1 = Multidimensional Personality Index Negative Affectivity at Time 1. OCQ2 = Organizational Commitment Questionnaire at Time 2. A&M AOC2 = Allen and Meyer's Affective Commitment at Time 2. MSQ2 = Minnesota Satisfaction Questionnaire at Time 2. Hoppock2 = Hoppock job satisfaction at Time 2. PANAS.P2 = PANAS Positive Affectivity at Time 2. MPI.P2 = Multidimensional Personality Index Positive Affectivity at Time 2. PANAS.N2 = PANAS Negative Affectivity at Time 2. MPI.N2 = Multidimensional Personality Index Negative Affectivity at Time 2.

Table 2
Estimated construct-level relationships among variables.

	JS	OC	PA	NA
Job satisfaction (JS)	*.90*	.91 (.72)	.59 (.47)	−.44 (−.36)
Organizational commitment (OC)	.65	*.90*	.53 (.40)	−.27 (−.21)
Positive affectivity (PA)	.42	.36	*.91*	−.56 (−.42)
Negative affectivity (NA)	−.32	−.19	−.38	*.91*

Notes. N = 255–270. Mean coefficient alpha for all measures of the same construct are shown in the diagonals. Values in the lower diagonals are mean observed correlations of the measures for the same constructs obtained *across different times*. Values in the upper diagonals are the estimated construct-level correlations: estimates based on the CFA Model 1 are presented outside the parentheses; corrected correlations based on coefficients alpha are within the parentheses.

Table 3
Comparing hierarchically nested models.

Model	Description	Parameter estimates	Fit indexes	Compared to less constrained model(s)
1	Most general model No constraint	$\hat{\rho}_{JS-OC} = .91$; $\hat{\rho}_{JS-PA} = .59$; $\hat{\rho}_{OC-PA} = .53$; $\hat{\rho}_{JS-NA} = -.44$; $\hat{\rho}_{OC-NA} = -.27$; $\hat{\rho}_{PA-NA} = -.56$	χ^2 = 1726.49, df = 1042 ($p < .01$); CFI = .948; $RMSEA$ = .051; $SRMR$ = .052	
2	Nested within Model 1 Constraints $\rho_{JS-NA} = \rho_{OC-NA}$ $\rho_{JS-PA} = \rho_{OC-PA}$	$\hat{\rho}_{JS-OC} = .91$; $\hat{\rho}_{JS-PA} = \hat{\rho}_{OC-PA} = .56$; $\hat{\rho}_{JS-NA} = \hat{\rho}_{OC-NA} = -.36$; $\hat{\rho}_{PA-NA} = -.56$	χ^2 = 1744.27, df = 1044 ($p < .01$); CFI = .947; $RMSEA$ = .051; $SRMR$ = .058	$\Delta\chi^2$ = 17.77, Δdf = 2 ($p < .01$); $\Delta CFI = -.001$
3	Nested within Model 2 Constraints $\rho_{JS-NA} = \rho_{OC-NA}$ $\rho_{JS-PA} = \rho_{OC-PA}$ $\rho_{JS-OC} = 1.00$	$\hat{\rho}_{JS-OC} = 1.00$; $\hat{\rho}_{JS-PA} = \hat{\rho}_{OC-PA} = .58$; $\hat{\rho}_{JS-NA} = \hat{\rho}_{OC-NA} = -.38$; $\hat{\rho}_{PA-NA} = -.56$	χ^2 = 1784.62, df = 1045 ($p < .01$); CFI = .944; $RMSEA$ = .053; $SRMR$ = .058	Compared to Model 1 $\Delta\chi^2$ = 58.13, Δdf = 3 ($p < .01$); $\Delta CFI = -.004$ Compared to Model 2 $\Delta\chi^2$ = 17.73, Δdf = 1 ($p < .01$); $\Delta CFI = -.003$

Notes. $\hat{\rho}_{JS-OC}$ = estimated correlation between the constructs of job satisfaction and organizational commitment. $\hat{\rho}_{JS-PA}$ = estimated correlation between the constructs of job satisfaction and positive affectivity. $\hat{\rho}_{JS-NA}$ = estimated correlation between the constructs of job satisfaction and negative affectivity. $\hat{\rho}_{OC-PA}$ = estimated correlation between the constructs of organizational commitment and positive affectivity. $\hat{\rho}_{OC-NA}$ = estimated correlation between the constructs of organizational commitment and negative affectivity. $\hat{\rho}_{PA-NA}$ = estimated correlation between the constructs of positive affectivity and negative affectivity. CFI = comparative fit index. $RMSEA$ = Root Mean Square Error of Approximation. $SRMR$ = Standardized Root Mean Square Residual. $\Delta\chi^2$ = difference in χ^2 between two models. ΔCFI = difference in CFI between two models.

Model 3, which specifies that JS and OC are perfectly indistinguishable by the data, also yields reasonable fit: CFI = .944, $RMSEA$ = .053, $SRMR$ = .058, χ^2 = 1784.62, df = 1045 ($p < .01$). However, as shown in Table 3, compared to Model 2, the absolute difference in CFI ($|.944 - .947| = .003$) is slightly larger than the preset cut-off value of .002. Thus, it is possible that the construct-level correlation between JS and OC, although very high, is not perfect (i.e., equal to 1.00). This suggests that Model 2 above may better reflect the data of this study.

Although the correlation between JS and OC estimated in Model 2 does not reach 1.00, indicating there remain unique variances in these constructs, the patterns of relationships of JS and OC with other constructs (PA and NA) are the same. This finding suggests a possibility that there is a general, higher-order construct underlying JS and OC which largely determines their relationships with other external constructs. To further investigate this possibility, we conducted an additional analysis directly examining the variance components contributing to the observed variation of the items included in the JS and OC measures. Specifically, we estimated the proportions of different sources of variance shown in Eq. (1). In addition, we attempted to "disentangle" the variance attributed to the constructs, separating it into: (a) the variance shared by both the JS and OC and (b) the variance unique to either the JS or OC construct. This analysis allows us to directly examine the extent to which the construct variance in the items of the JS and OC measures is due to the general factor. If the proportion attributable to the general factor is much larger than that due to the unique factor, such a finding would provide certain support for the relative importance of the general job attitude factor underlying the JS and OC measures.

For this analysis, we used a CFA model similar to the models described in "*Analysis*" section. That is, we specified latent factors that represent sources of the variances in a self-report measurement item. There are, however, several important differences. First, for the current model, we attempted to look more closely at the item level instead of the subscale level as in the earlier analysis. Thus, ideally we would include items as indicators in the model. However, due to the requirements of maximum likelihood estimation mentioned earlier (i.e., ratio of sample size to parameter estimates and interval and normally distributed indicators), we had to strike a balance by combining items into pairs and using these pairs of items as indicators. We did this by examining item content and pairing items as close in meaning to each other as possible. There were two exceptions in creating the item pairs. First, when a measure had an odd number of items (e.g., the OCQ), one "pair" of that measure was created by combining three items instead of two. The second exception was for the Hoppock scale which includes items with seven response options (instead of five response options in other measures). Given this larger number of response options (and relatively normal distribution of responses observed in the data), we decided to use the items of the Hoppock scale as indicators instead of combining them into item pairs. As such, results of this additional analysis mostly pertain to pairs of items (except for the Hoppock scale), not the items per se. Nevertheless, we believe that these results still provide us with important information about the properties of the items included in the measures of JS and OC.

The second difference is related to the first one: to maximize the ratio of sample size to parameter estimates, we included only items of the JS and OC measures, not those of the PA and NA measures. Finally, in addition to the latent factors described earlier in "*Analysis*" section (i.e., transient error, scale specific factor error, and construct), we specified a general latent factor representing the general attitude construct to underlie all the indicators. As such, the original JS and OC latent factors (underlying indicators of either measures of JS or OC) now represent the factors "unique" to either the JS or OC construct. Since the shared variance between the JS and OC latent variables is now accounted for by the general factor in the current model, the JS and OC latent factors were constrained to be uncorrelated.

For each indicator, the proportion of variance attributable to the latent factors (transient error, scale specific factor error, the JS or OC unique factor, and the general factor) can be estimated by squaring the standardized loadings of these factors. As noted earlier, residual variance of an indicator includes both item specific factor error and random response error. From this, proportion of variance due to item specific factor error can be estimated by multiplying the residual variance by the correlation between the residual variances of the same item (or item pair) across occasions. Variance due to random response error is then obtained by subtracting the proportion of variance due to item specific factor error variance (just estimated) from the residual variance.

The model has acceptable fit $CFI = .944$, $RMSEA = .050$, $SRMR = .045$, $\chi^2 = 1545.12$, $df = 943$ ($p < .01$). The variance proportions for each indicator (item or item pair) are shown in Table 4. As can be seen, the largest proportion in variance of all the item pairs is due to the general factor, which ranges from .26 to .61. On average, the general factor accounts for about .42 of the variance in JS item pairs (or items) and .46 of variance in OC item pairs. Across all item pairs, the mean proportion due to the general factor is .44. In contrast, the unique factor of the JS construct accounts for only about .04 of the variance in item pairs. The mean proportion due to the unique factor of the OC construct is larger, about .09, but is still much smaller than that of the general factor (.46). Overall, the proportion due to the unique factor of either JS or OC (.06) is only about 12% of that due to the general factor (.44). These results suggest that variation in employees' responses to measures of JS and OC is largely due to the general factor. In other words, it appears that employees tend to respond similarly to the items of these measures; they do not differentiate much between the measures.

Discussion

In this paper we presented and applied a new and more accurate method for estimating construct-level relationships, thereby allowing empirical determination of whether constructs are empirically redundant. The results of the application of the method to an empirical data set suggest the broader possibility that the problem of construct empirical redundancy may be quite widespread in organizational research.

The empirical redundancy of job satisfaction and organizational commitment

We found that the constructs underlying well-established measures of JS and OC are highly correlated. The construct-level correlation, estimated to be .91 in the study, is starkly different from the smaller previously estimated values in the literature (e.g., Brooke et al., 1988; Mathieu & Farr, 1991; Mathieu & Zajac, 1990; Nystedt et al., 1999). Further analysis suggests the possibility that there is a higher-order construct accounting for most of the variation in the measures of the JS and OC constructs. This finding appears to dovetail nicely with Harrison et al.'s (2006) hypothesis that there is a general job attitude construct. Harrison and colleagues examined path models which specify the link between the general job attitude construct underlying JS and OC and a general behavioral criterion. These models were found to fit the data well. Based on their results, Harrison et al. suggested that the general job attitude construct also underlies other job attitudes measures (for example, job involvement) and that this general construct accounts for the relationships between job attitudes and other organizational outcomes. Taken together, current findings seem to be consistent with Harrison et al.'s suggestion that the correlation between JS and OC as well as their correlations with other variables in the nomological network of organizational constructs (both dispositional determinants, PA and NA, as found in the current study, and behavioral outcomes, as shown in Harrison et al.'s study) is probably due to the general job attitude construct. In other words, what really mat-

Table 4
Proportion of variance components in the variance of an item pair in measures of the JS/OC construct.

Scale	Proportion of variance to the total observed variance					
Item pair	Item	Random	Transient	Scale	JS/OC	General
Hoppock						
HOP1	.016	.283	.052	.055	.004	.586
HOP2	.022	.217	.020	.127	.000	.611
HOP3	.108	.350	.006	.015	.003	.519
HOP4	.037	.341	.034	.154	.011	.420
Mean Hoppock	.046	.298	.028	.088	.005	.534
MSQ						
MSQ1	.188	.373	.069	.016	.001	.336
MSQ2	.086	.319	.107	.009	.027	.436
MSQ3	.097	.432	.102	.030	.058	.265
MSQ4	.089	.348	.031	.055	.012	.463
MSQ5	.089	.333	.094	.076	.104	.296
MSQ6	.172	.371	.102	.044	.046	.256
MSQ7	.109	.400	.145	.012	.002	.305
MSQ8	.000	.162	.066	.228	.100	.440
MSQ9	.024	.231	.053	.128	.097	.467
MSQ10	.000	.202	.028	.216	.081	.471
Mean MSQ	.086	.317	.080	.081	.053	.374
Mean all JS scales	.074	.312	.065	.083	.039	.419
OCQ						
OCQ1	.020	.242	.059	.036	.040	.613
OCQ2	.031	.351	.089	.001	.147	.380
OCQ3	.013	.173	.138	.047	.088	.547
OCQ4	.008	.198	.085	.149	.099	.474
Mean OCQ	.018	.241	.093	.058	.094	.503
Allen and Meyer						
AOC1	.029	.233	.069	.004	.063	.606
AOC2	.022	.282	.045	.052	.191	.405
AOC3	.030	.387	.075	.180	.016	.307
AOC4	.000	.254	.086	.226	.062	.364
Mean Allen and Meyer	.020	.289	.069	.115	.083	.420
Mean all OC scales	.019	.265	.081	.087	.088	.462
Mean all scales	.054	.295	.071	.084	.057	.435

Notes. Item = item specific factor error. Random = random response error. Transient = transient error. Scale = scale specific factor error. JS/OC = unique factor due to either JS or OC construct. General = general factor shared by both the JS and OC construct.

ters is perhaps the shared variance between JS and OC, not the unshared variances uniquely attributable to each of these constructs.

An alternative explanation for the high construct-level correlation between JS and OC is that the two constructs are strongly and reciprocally causally related (Farkas & Tetrick, 1989). That is, JS and OC are distinct but because of their reciprocally causal relationship, it is not possible to empirically distinguish the constructs in cross-sectional data. Unfortunately, current data do not allow us to resolve the question whether the high construct-level correlation between JS and OC is due to: (a) the existence of the general attitude construct, or (b) the reciprocally causal relationships between the constructs, or (c) the fact that they are indeed the same construct. It should be noted that among the alternative explanations listed here, the second one about the reciprocally causal relationships may be the least plausible because in order to have such a high correlation (.91) one construct should be almost the sole cause of the other. Answering the question about the underlying cause for the high construct-level relationship between JS and OC would require carefully designed longitudinal data. As noted earlier, however, determining the cause is not the focus of the paper. Instead, we are interested in examining the empirical redundancy of the constructs, which, as discussed earlier, is manifested by: (a) a very high construct-level correlation and (b) similar patterns of correlations with other variables. Current results seem to show that these conditions hold with the JS and OC constructs, suggesting they may be empirically redundant. As discussed earlier, empirical redundancy renders it impossible to disentangle the constructs, their effects and relationships with other variables, and consequently their roles in the nomological network of other organizational constructs by *empirical data*.

It is worth reemphasizing here that we did not challenge the *theoretical* distinction between the constructs of job satisfaction and organizational commitment, which has been well-established in the literature. However, the very high correlation among the constructs underlying measures of job satisfaction and organizational commitment found in our study suggests that the constructs cannot be *empirically* distinguished in any practical sense in real research data. As such, while the constructs may be theoretically different, respondents probably cannot reliably distinguish among them, consciously or unconsciously, based upon current measures. Most organizational researchers probably believe that it is very clear that job satisfaction and organizational commitment are distinct conceptually – and indeed they are conceptually different. However, this logical distinctiveness is of no import for research if it is not borne out in employees' responses which are the raw material researchers use to empirically build, test, and refine organizational theories.

Measurement artifacts in measures of organizational constructs

The current study provides the first estimates of the extent to which measurement artifacts, specifically transient error and scale specific factor error, contribute to the observed variances of measures of job attitudes. As seen in Table 4, the proportion of transient error variance relative to total observed variances for the item pairs averaged .062 for the JS measures and .081 for the OC measures. These values suggest that the effect of transient error in measures of job attitudes is not negligible. The proportions of scale specific factor error are even larger, averaging .078 for the item pairs of the JS measures and .097 for those of the OC measures. When compared to the two measurement artifacts accounted for by the coefficient alpha, these values are smaller than the averaged proportion of random response error (.312 for JS and .265 for OC measures) but larger than those of item specific factor error (.069 for JS and .019 for OC measures). It should be noted that these proportions are estimated for item pairs. For a full scale which is created by summing or averaging across the item pairs, the proportions of random response error and item specific factor error in the observed variance will be smaller because different item pairs do not share the same measurement artifacts. Proportions of transient error and scale specific factor error, which are the same for all the item pairs of a measure, will therefore be relatively increased. As such, the findings not only provide evidence confirming the existence of transient error and scale specific factor error, but also indicate the relative importance of these measurement artifacts in measures of organizational constructs, thereby challenging the traditional practice of ignoring them in empirical research.

Table 2 illustrates the importance of accounting for all major sources of measurement artifacts in order to estimate construct-level relationships in organizational research. As can be seen, the correlations among JS, OC, PA, and NA estimated by using coefficient alpha (above the diagonal and in parentheses) are much smaller than those obtained by the CFA procedure which accounts for all four sources of measurement artifacts. Thus, these findings suggest that when the appropriate procedure is used, conclusions about organizational phenomena can be substantially different from those based on estimates of construct-level relationships that are distorted by measurement artifacts.

The method for estimating construct-level relationships – considerations and limitations

From the operational definition of constructs described earlier, it can be seen that accurate estimation of construct variance and construct-level relationships depends on appropriate selection of occasions and measures. Specifically, the length of the interval between two measurement occasions is important because it influences how transient error variance is estimated. If the interval is too long, persons' relative standing on the construct may change, causing an inflated estimate of transient error variance (that is, the real changes in the construct are confounded with transient error). Consequently, estimates of construct-level relationships would be inflated. Le et al. (2009) discussed this question in some detail and noted that determining the appropriate interval should be based on accepted theory about the stability of constructs. In the current study, the interval was relative short (1 week), so it is highly unlikely that constructs examined substantially changed during the period. Also, as described in "*Methods*" section, we took the extra step of excluding responses from employees who reported changes in work settings which may have resulted in changes in the constructs of JS and OC. Thus, we can be confident that results were not affected by the confounding effects of real changes in the constructs.

Selection of measures is also critical because it is the shared variance between the measures that operationally determines the construct variance. This fact may raise concerns that the estimates of construct-level correlations can be heavily influenced by how the measures are selected to be included in the analysis. Specifically if the measures are not highly related, the estimated construct variance would be low, leading to a large correction of the observed correlations when estimating the construct-level relationships. The reverse is true when the correlation between the measures selected is high. However, this is essentially a question of theory, not method. Le et al. (2009) argue that for most established measures of major theoretical constructs in organizational and psychological research the common theoretical foundation is sound enough that a large proportion of their observed variance is due to the construct they are intended to measure. As noted earlier, the measures included in the current study are well-established and have been used in empirical research. However, ideally it would be better to include three or more measures for each construct. That would allow the construct to be over-identified and thus better defined (cf. Anderson & Gerbing, 1988).

Naturally, for the approach to work, measures selected as indicators of a construct should reflect the same underlying construct. In the current paper, it can be seen that the Hoppock's scale and the MSQ operationalize the construct of overall job satisfaction differently. Apart from the fact that the former captures the construct more directly with global items while the latter does that indirectly by combining different specific facets, the MSQ appears to be more cognitive/evaluative in nature as compared to the Hoppock which is more affective. This may raise a question about the difference of the constructs underlying these measures. However, as discussed earlier, that difference reflects the specific factor error of the measures because it is likely that empirical researchers wish to generalize their research findings regarding overall job satisfaction beyond the measures used in their studies. This situation was specifically discussed in Le et al. (2009): "In many research areas, several scales have been developed to measure the same theoretical construct (e.g., self-esteem, emotional stability, job satisfaction). These scales may or may not be based on the same theory or model – for example, Job In General (JIG; Ironson, Smith, Brannick, & Gibson, 1989) and Minnesota Satisfaction Questionnaire (MSQ full scale; Weiss et al., 1967) are two measures that are ostensibly based on different conceptualizations of the construct of overall job satisfaction – but they are nevertheless used by researchers to empirically examine the relationships between that construct and other constructs in its nomological network (e.g., job satisfaction as a determinant of turnover: Hom, Caranikas-Walker, Prussia, & Griffeth, 1992). Results are typically not interpreted as being specific to the measures used in those studies (e.g., overall job satisfaction as measured by the JIG), but rather, conclusions are drawn about the construct in general. As such, the conclusions are assumed to be generalizable to all measures for the construct." (pp. 168–169).

Related to the issue, a reviewer pointed out that measures for *different* constructs might include similar items because a new measure might be developed based on knowledge and consideration of existing measures. If that happened, scale specific factor error will be correlated, leading to an inflation of correlations between measures of *different* constructs. As a result, estimated construct-level correlations might be inflated. In the current study, this can potentially be a problem. Specifically, some items of the MSQ concerning different aspects of the job may be interpreted as tapping into the characteristics of the employing organization (vs. only the job itself). For example, the item "The way company policies are put into practice" of the MSQ can invoke impressions about the organization. In particular, this item is most likely to be related to the item "This company really inspires the very best in me in the way of job performance" of the OCQ. In addition, items addressing value congruence in the OCQ and MSQ appear to tap a common antecedent for both JS and OC. As such, the observed correlation between measures of JS and OC may be inflated due to the inclusion of these items in the measures. To examine the problem we re-analyzed the data using the models described in "*Methods*" sections. However, for these analyses, we allowed all the latent factors representing the scale specific factor errors to be correlated. Such correlations account for the potential confounding effects due to similar items included in measures of different constructs discussed above. Results of these additional analyses were essentially the same as those reported in "*Results*" section in terms of estimated construct-level relationships. In particular, the correlation between JS and OC remains at .91. These results suggest that in the current study, similarity of items included in measures of different constructs did not create any notable problem. In general, however, this may pose a more serious problem in other situations. If this is indeed a concern, researchers should examine additional models allowing correlated scale specific factor errors as described here to evaluate the potential effect of the problem.

Another related issue is the heterogeneity and dimensionality of the items within a measure. The CFA approach described here assumes the classic direct reflective models about the relationships between measures/indicators and the latent constructs (cf. Edwards & Bagozzi, 2000). However, as pointed out by Edwards and Bagozzi (2000), several items of the OCQ may not fit well into this model. Instead, other models (e.g., spurious model, indirect reflective model) may be more appropriate. The extent to which this problem may bias the construct-level relationships estimates is not clear. Relevant to the issue, Little, Cunningham, Shahar, and Widaman (2002) argued that when the goal of researchers is to estimate correlations among latent factors, combining items into item parcels (thereby essentially bypassing the issue of item dimensionality and heterogeneity) is justified and does not result in serious bias. That argument suggests that our construct-level correlations estimates obtained here are appropriate. Nevertheless, future research may need to revisit the issue using different measures of OC and JS to triangulate current findings.

It can be argued that efforts should be made to develop measures that better reflect the theoretical constructs of job satisfaction and organizational commitment rather than applying the procedure described here to correct for the effects of measurement artifacts in existing measures. While we agree that more attention is needed for improving measures of organizational constructs, we believe that such effort is complementary, not alternative, to the use of the current procedure for estimating construct-level relationships. First, developing "better" measures may not be easy. After all, the measures included in the current studies are among the best established measures and have been used in organizational research for decades. Improving on these well-researched measures is certainly not a simple task. Even if such improvements can be made, the new measures will not be free from the effects of measurement artifacts (although these effects may not be as large as they are in existing measures). Therefore we still need methods which can take these measurement artifacts into account in order to estimate relationships among constructs from observed correlations among measures. Finally, without the procedure to accurately estimate relationships between constructs, it would not be possible to detect the problem of construct empirical redundancy and the need for improved measures would not be recognized.

A reviewer pointed out that measures of the constructs examined in the current study are self-report, so their observed correlations could be inflated by the effect of a common method factor (cf. Doty & Glick, 1998). Consequently, the construct-level relationships obtained here could be overestimated. However, there is reason to believe that a common method factor due to the use of self-report measures does not pose a problem in the current studies. Le et al. (2009) examined the studies reported in Doty and Glick (1998) and noted that among the three dimensions of common method factor (rater as method, instrument-based method, and temporal method) only the temporal method dimension causes the inflation in observed correlations of self-report measures. This dimension reflects the same effect as that of correlated transient error, which the current estimation procedure took into account, as described earlier. That is, the "method effect" is actually the effect of correlated transient measurement errors, and so controlling for correlated transient measurement errors eliminates this effect.

Conclusion

The current study demonstrated that all four sources of measurement artifacts account for substantial proportions of observed variances of job attitude measures. The effect of measurement artifacts appears to be larger than has been previously assumed by most researchers. Accordingly, it is possible and perhaps likely that current conclusions in the literature about the relationships between constructs have been distorted by overlooked sources of measurement artifacts: scale specific factor error and transient error. It is also possible that in the domain of job and work attitudes, the problem of construct empirical redundancy is more prevalent than realized. If so, there are important implications for alleviating the problem of construct proliferation and for more refined theory development in many areas of research. Conceivably, evidence of construct empirical redundancy will either discourage researchers from creating and/or supporting new constructs which may be conceptually different but empirically indistinguishable from existing ones or enable them to develop better, more creative ways to measure the constructs. Such evidence will also necessitate revisiting theories in the organizational behavior area involving the empirically redundant constructs. Consequently, these theories may be revised and/or simplified. As a first step, more accurate estimates of relationships among organizational constructs must be sought. In the interest of progress in research, it is critical that we resolve questions of construct redundancy and construct proliferation and move in the direction of the ideal of scientific parsimony.

References

Agho, A. O., Price, J. L., & Mueller, C. W. (1992). Discriminant validity of measures of job satisfaction, positive affectivity and negative affectivity. *Journal of Occupational and Organizational Psychology, 65*, 185–196.

Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology, 63*, 1–18.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*, 411–423.

Bargozzi, R. P., & Edwards, J. R. (1998). A general approach to representing constructs in organizational research. *Organizational Research Methods, 1*, 45–87.

Bateman, T. S., & Strasser, S. (1984). A longitudinal analysis of the antecedents of organizational commitment. *Academy of Management Journal, 27*, 95–112.

Becker, G. (2000). How important is transient error in estimating reliability? Going beyond simulation studies. *Psychological Methods, 5*, 370–379.

Blalock, H. M., Jr. (1968). The measurement problem: A gap between the languages of theory and research. In H. M. Blalock, A. B. Blalock (Eds.), *Methodology in social research* (pp. 5–27). New York, NY: McGraw-Hill.

Bowling, N. A., Beehr, T. A., & Lepisto, L. R. (2006). Beyond job satisfaction: A five year prospective analysis of the dispositional approach to work attitudes. *Journal of Vocational Behavior, 69*, 315–330.

Brannick, M. (1995). Critical comments on applying covariance structure modeling. *Journal of Organizational Behavior, 16*, 201–213.

Brayfield, A. H., & Rothe, H. F. (1951). An index of job satisfaction. *Journal of Applied Psychology, 35*, 307–311.

Brief, A. P., & Roberson, L. (1989). Job attitude organization: An exploratory study. *Journal of Applied Social Psychology, 19*, 717–727.

Brooke, P. P., Russell, D. W., & Price, J. L. (1988). Discriminant validation of measures of job satisfaction, job involvement, and organizational commitment. *Journal of Applied Psychology, 73*, 139–145.

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

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Cohen, P., Cohen, J., Teresi, J., Marchi, M., & Velez, C. N. (1990). Problems in the measurement of latent variables in structural equations causal models. *Applied Psychological Measurement, 14*, 183–196.

Cook, J. D., Hepworth, S. J., Wall, T. D., & Warr, P. B. (1981). *The experience of work: A compendium and review of 249 measures and their use*. London: Academic Press.

Cooper-Hakim, A., & Viswesvaran, C. (2005). The construct of work commitment: Testing an integrative framework. *Psychological Bulletin, 131*, 241–259.

Cramer, D. (1996). Job satisfaction and organizational continuance commitment: A two-wave panel study. *Journal of Organizational Behavior, 17*, 389–400.

Cronbach, L. J. (1947). Test reliability: Its meaning and determination. *Psychometrika, 12*, 1–16.

Cronbach, L. J., Glesser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. John Wiley & Sons.

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

Curry, J. P., Wakefield, D. S., Price, J. L., & Mueller, C. W. (1986). On the causal ordering of job satisfaction and organizational commitment. *Academy of Management Journal, 29*, 847–858.

DeShon, R. P. (1998). A cautionary note on measurement error corrections in structural equation models. *Psychological Methods, 3*, 412–423.

Doty, H. D., & Glick, W. H. (1998). Common methods bias: Does common methods variance really bias results? *Organizational Research Methods, 4*, 374–406.

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

Farkas, A. J., & Tetrick, L. E. (1989). A three-wave longitudinal analysis of the causal ordering of satisfaction and commitment on turnover decisions. *Journal of Applied Psychology, 74*, 855–868.

Griffeth, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management, 26*, 463–488.

Harrison, D. A., & Martocchio, J. J. (1998). Time for absenteeism: A 20-year review of origins, offshoots, and outcomes. *Journal of Management, 24*, 305–350.

Harrison, D. A., Newman, D. A., & Roth, P. L. (2006). How important are job attitudes? Meta-analytic comparisons of integrative behavioral outcomes and time sequences. *Academy of Management Journal, 49*, 305–325.

Harter, J. K., & Schmidt, F. L. (2008). Conceptual versus empirical distinctions among constructs: Implications for discriminant validity. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 36–39.

Hau, K. T., & Marsh, H. W. (2004). The use of item parcels in structural equation modeling: Non-normal data and small sample sizes. *British Journal of Mathematical Statistical Psychology, 57*, 327–351.

Hom, P. W., & Griffeth, R. W. (1991). Structural equations modeling test of a turnover theory: Cross-sectional and longitudinal analyses. *Journal of Applied Psychology, 76*, 350–366.

Hoppock, R. (1935). *Job satisfaction*. New York: Harper Brothers.

Hulin, C. L. (1991). Adaptation, persistence, and commitment in organizations. In M. D. Dunnette & L. M. Hough (Eds.). *Handbook of industrial and organizational psychology* (Vol. 2, 2nd ed., pp. 445–505). Palo Alto, CA: Consulting Psychologists Press.

Hunter, J. E., & Gerbing, D. W. (1982). Unidimensional measurement, second order factor analysis, and causal models. *Research in Organizational Behavior, 4*, 267–320.

Ironson, G. H., Smith, P., Brannick, M. T., & Gibson, W. M. (1989). Constructing a job in general scale: A comparison of global, composite, and specific measures. *Journal of Applied Psychology, 74*, 193–200.

Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling, 10*, 128–141.

Judge, T. A., & Larsen, R. J. (2001). Dispositional affect and job satisfaction: A review and theoretical extension. *Organizational Behavior and Human Decision Processes, 86*, 67–98.

Le, H., Schmidt, F. L., & Putka, D. J. (2009). The multifaceted nature of measurement artifacts and its implications for estimating construct-level relationships. *Organizational Research Methods, 12*, 165–200.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151–173.

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.

Lodahl, T. M., & Kejner, M. (1965). The definition and measurement of job involvement. *Journal of Applied Psychology, 49*, 24–33.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Marsh, H. W., & Hocevar, D. (1988). A new, more powerful approach to multitrait–multimethod analyses: Application of second-order confirmatory factor analysis. *Journal of Applied Psychology, 73*, 107–117.

Mathieu, J. E., & Farr, J. L. (1991). Further evidence for the discriminant validity of measures of organizational commitment, job involvement, and job satisfaction. *Journal of Applied Psychology, 76*, 127–133.

Mathieu, J. E., & Zajac, D. M. (1990). A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment. *Psychological Bulletin, 108*, 171–194.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*, 568–592.

Meyer, J. P., Stanley, D. J., Herscovitch, L., & Topolnytsky, L. (2002). Affective, continuance, and normative commitment to the organization: A meta-analysis of antecedents, correlates, and consequences. *Journal of Vocational Behavior, 61*, 20–52.

Morrow, P. C. (1983). Concept redundancy in organizational research: The case of work commitment. *Academy of Management Review, 8*, 486–500.

Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior, 14*, 224–247.

Nystedt, L., Sjöberg, A., & Hägglund, G. (1999). Discriminant validation of measures of organizational commitment, job involvement, and job satisfaction among Swedish army officers. *Scandinavian Journal of Psychology, 40*, 49–55.

Organ, D. W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology, 48*, 775–802.

Porter, L. W., Steers, R. M., Mowday, R. T., & Boulian, P. V. (1974). Organizational commitment, job satisfaction and turnover among psychiatric technicians. *Journal of Applied Psychology, 59*, 603–609.

Price, J. L. (1997). Handbook of organizational measurement. *International Journal of Manpower, 18*, 303–558.

Ree, M. J., & Carretta, T. R. (2006). The role of measurement error in familiar statistics. *Organizational Research Methods, 9*, 99–112.

Rousseau, D. M. (2007). A sticky, leveraging, and scalable strategy for high-quality connections between organizational practice and science. *Academy of Management Journal, 50*, 1037–1042.

Roznowski, M., & Hulin, M. (1992). The scientific merit of valid measures of general constructs with special reference to job satisfaction and job withdrawal. In C. J. Cranny, P. C. Smith, & E. F. Stone (Eds.), *Job satisfaction*. New York: Lexington.

Rusbult, C. E., & Farrell, D. (1983). A longitudinal test of the investment model: The impact on job satisfaction, commitment, and turnover of variations in rewards, costs, alternatives, and investments. *Journal of Applied Psychology, 68*, 429–438.

Scarpello, V., & Campbell, J. P. (1983). Job satisfaction: Are all the parts there? *Personnel Psychology, 36*, 577–600.

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.

Schmidt, F. L., Le, H., & Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual differences constructs. *Psychological Methods, 8*, 206–224.

Schwab, D. E. (1980). Construct validity in organizational behavior. In B. M. Staw & L. L. Cummings (Eds.). *Research in organizational behavior* (Vol. 2, pp. 3–43). Greenwich, CT: JAI Press.

Singh, J. (1991). Redundancy in constructs: Problem, assessment, and an illustrative example. *Journal of Business Research, 22*, 255–280.

Tesser, A., & Krauss, H. (1976). On validating a relationship between constructs. *Educational and Psychological Measurement, 36*, 111–122.

Thoresen, C. J., Kaplan, S. A., Barsky, A. P., Warren, C. R., & Chermont, K. D. (2003). The affective underpinnings of job perceptions and attitudes: A meta-analytic review and integration. *Psychological Bulletin, 129*, 914–945.

Thorndike, R. L. (1949). *Personnel selection: Test and measurement techniques*. New York, NY: John Wiley.

Thorndike, R. L. (1951). Reliability. In E. F. Lindquist (Ed.), *Educational measurement* (pp. 560–620). Washington: American Council on Education.

Toulmin, S. (1961). *Foresight and understanding: An enquiry into the aims of science*. Bloomington: Indiana University Press.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133*, 1038–1066.

Vandenberg, R. J., & Lance, C. E. (1992). Examining the causal order of job satisfaction and organizational commitment. *Journal of Management, 18*, 153–167.

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.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219–235.

Weiss, D. J., Dawis, R. V., England, G. W., & Lofquist, L. H. (1967). *Manual for the Minnesota Satisfaction Questionnaire*. Industrial Relations Center, University of Minnesota, Minneapolis, MN.

Wiener, Y., & Vardi, Y. (1980). Relationship between job, organizational and career commitments and work outcomes: An integrative approach. *Organizational Behavior and Human Performance, 26*, 81–96.

Williams, L. J., & Hazer, J. T. (1986). Antecedents and consequence of satisfaction and commitment in turnover models: A re-analysis using latent variable structural equation models. *Journal of Applied Psychology, 71*, 219–231.