Journal of Management
http://jom.sagepub.com/

# Measuring Job-Related Situational Strength and Assessing Its Interactive Effects With Personality on Voluntary Work Behavior

Rustin D. Meyer, Reeshad S. Dalal, Irwin J. José, Richard Hermida, Tiffani R. Chen, Ronald P. Vega, Charlie K. Brooks and Vivek P. Khare

*Journal of Management* 2014 40: 1010 originally published online 8 November 2011
DOI: 10.1177/0149206311425613

The online version of this article can be found at:
http://jom.sagepub.com/content/40/4/1010

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
S•M•A
SOUTHERN MANAGEMENT ASSOCIATION
Southern Management Association

Journal of Management
Vol. 40 No. 4, May 2014 1010-1041
DOI: 10.1177/0149206311425613


# Measuring Job-Related Situational Strength and Assessing Its Interactive Effects With Personality on Voluntary Work Behavior

Rustin D. Meyer
*Georgia Institute of Technology*
Reeshad S. Dalal
Irwin J. José
Richard Hermida
Tiffani R. Chen
Ronald P. Vega
*George Mason University*
Charlie K. Brooks
*Georgia Institute of Technology*
Vivek P. Khare
*George Mason University*

*Situational strength has long been viewed as a useful way of conceptualizing and predicting person–situation interactions. Some have recently argued, however, that more rigorous empirical tests of its behavioral influence are sorely needed. The current article begins addressing this literature gap by (a) developing the Situational Strength at Work (SSW) scale, (b) examining the ways in which individual differences influence perceptions of situational strength, and (c) testing situational strength's moderating effects on two types of voluntary work behavior (i.e., organizational citizenship behavior and counterproductive work behavior). Results indicate*

*Acknowledgments: Portions of this research were funded by the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI), W91WAW-09-C-0096-BAA. The views expressed in this article, however, are solely those of the authors and should not necessarily be regarded as reflecting the positions or policies of ARI.*

*Corresponding author: Rustin D. Meyer, School of Psychology, Georgia Institute of Technology, 654 Cherry St. NW, Atlanta, GA 30332-0170, USA.*

*E-mail: rustin.meyer@psych.gatech.edu*

*strong psychometric properties for the SSW (thereby facilitating future organizational research on situational strength), support for theoretically based predictions regarding the role of individual differences in perceptions of situational strength, support for theoretically based moderator effects on organizational citizenship behavior, and the presence of countertheoretical (yet strong and consistent) moderator effects on counterproductive work behavior. Thus, this study makes several contributions to the situational strength literature but also reveals important areas for future theoretical development and empirical research.*

***Keywords:*** *situational strength; interactionism; personality; moderator; scale development*

---

Social scientists have long argued that human behavior is a joint function of individual differences (e.g., intelligence, personality, interests) and situational characteristics (Cronbach, 1957; Lewin, 1936). Although no true consensus exists about the nature and structure of situations (Funder, 2006; Johns, 2006), many have highlighted the importance of situational strength (e.g., Johns, 2006; Mischel, 1977; Weiss & Adler, 1984). Indeed, Snyder and Ickes (1985: 904) went so far as to argue that situational strength should be viewed as "the most important situational moderating variable." Defined as implicit or explicit cues, provided by entities external to the individual, regarding the desirability of various forms of behavior (Meyer, Dalal, & Hermida, 2010), situational strength has been shown to moderate several relationships that are of particular interest to organizational scientists (cf. Barrick & Mount, 1993; Beaty, Cleveland, & Murphy, 2001; Meyer, Dalal, & Bonaccio, 2009).

The central tenet of situational strength is that it encourages individuals to engage in behaviors that they are unlikely to demonstrate when left to their own devices (Adler & Weiss, 1988; Stagner, 1977). The classic example of a strong situation is a red traffic light (Cooper & Withey, 2009). Here, relevant traits are unlikely to influence behaviors because the most appropriate course of action is so well defined that it overrides most people's natural tendencies, thereby muting the predictive validity of the individual difference(s) in question. Conversely, a yellow traffic light is a relatively weak situation because its behavioral message is open for interpretation. Thus, daring individuals are likely to speed through yellow traffic lights, whereas cautious individuals are likely to stop. As a result, relevant relationships are predicted to be stronger in those situations where the most appropriate course of action is in doubt (i.e., "weak" situations) and weaker in those situations where the most appropriate course of action is apparent (i.e., "strong" situations). In an example that is more relevant to the organizational sciences, conscientiousness is known to be a valid predictor of numerous important outcomes (e.g., task performance—Barrick & Mount, 1991; motivation to learn—Colquitt & Simmering, 1998; turnover—Salgado, 2002). According to the theory underlying situational strength, however, trait conscientiousness is not the only potential source of conscientious behavior (e.g., goal setting, increased time on task). To the extent, then, that situational strength is increased in a manner that encourages conscientious behavior among those who are unlikely to do so when left to their own devices, this should reduce the criterion-oriented validity of trait conscientiousness (Meyer et al., 2009).

## The Present Article

Despite (or perhaps because of) its intuitive appeal, situational strength and its purported moderating effects are often viewed as a behavioral truism, instead of a topic of inquiry in need of legitimate scientific inquiry (Cooper & Withey, 2009). In other words, researchers have prematurely accepted the role that situational strength plays in the organizational sciences, despite a relative dearth of rigorous empirical tests. Indeed, in their review of the strong situation hypothesis, Cooper and Withey (2009: 68) concluded that "despite its 30-year history, it remains only a hypothesis" and that future studies should include

> (a) a range of situation strengths that are clearly measured or manipulated, (b) measures of all relevant personality factors, and (c) statistical analyses capable of confirming whether the ability of those personality factors to predict theoretically relevant behaviors is moderated by the strength of the situation. (Cooper & Withey, 2009: 70)

The current article is intended not only to address this gap in the literature by meeting each of these criteria but also to address Meyer et al.'s (2010) call for the development of a standardized measure of situational strength. Such a measure would obviate the need to rely on ad hoc measures that do not adequately cover situational strength's construct space, thereby aiding researchers' attempts to examine the outcomes (and antecedents) of situational strength.

These goals are achieved through four interrelated studies. In Study 1 we use Meyer et al.'s (2010) four-facet conceptualization of situational strength (discussed subsequently) to describe how an initial bank of job-relevant situational strength items was written, refined, and initially reduced in number. In Study 2 we select the final set of items to form a standardized measure of this construct (the Situational Strength at Work [SSW] scale), assess its psychometric properties, and examine its convergent and discriminant validities with extant job characteristics. In Study 3 we examine the role that individual differences play in *perceptions* of situational strength—a need that was also highlighted in Cooper and Withey's (2009) review. In Study 4 (the data for which were actually collected before Study 3) we examine situational strength's moderating effects on relationships between two important individual differences (the personality traits of conscientiousness and agreeableness) and two important forms of voluntary work behavior (organizational citizenship behavior and counterproductive work behavior; Dalal, 2005; Spector & Fox, 2002). Said differently, the purpose of this article is to begin giving situational strength the serious scientific attention it deserves.

## Study 1—Item Development and Initial Screening

The purpose of Study 1 is to develop an initial bank of situational strength items and retain those with the highest content validity (DeVellis, 1991; Hinkin, 1998; Netemeyer, Bearden, & Sharma, 2003). To ensure that situational strength's construct space is fully covered, a

deductive scale-development process was used here wherein items were "designed to tap a previously defined theoretical universe" (Hinkin, 1995: 969). Specifically, the measure in question is based on the structure proposed by Meyer et al. (2010), who demonstrated that operationalizations of situational strength (or other situational characteristics that are compatible with the definition of situational strength) can be categorized into four interrelated facets.

The first facet of situational strength, "clarity," is defined as "the extent to which cues regarding work-related responsibilities or requirements are available and easy to understand" (Meyer et al., 2010: 125). This facet influences behavior by providing straightforward and easily comprehensible information about work-related responsibilities and/or requirements. For example, when employees receive instructions regarding how to complete a clearly defined task with a finite beginning, process, and end, they are less likely to act in accordance with their individual differences than they are in those situations where they are left to their own devices to define the steps necessary to succeed.

The second facet of situational strength, "consistency," is defined as "the extent to which cues regarding work-related responsibilities or requirements are compatible with each other" (Meyer et al., 2010: 126). This facet influences behavior by uniformly communicating a particular course of action across a variety of channels. Consistency is theoretically distinct from clarity in the sense that, even when presented clearly, multiple cues do not necessarily convey identical messages. For example, an employee may receive specific instructions from multiple managers (i.e., high clarity), but each of the managers may emphasize different tasks, goals, or priorities (i.e., low consistency).

The third facet of situational strength, "constraints," is defined as "the extent to which an individual's freedom of decision and action is limited by forces outside his or her control" (Meyer et al., 2010: 126). This facet influences behavior by preventing employees from exercising discretion pertaining to decisions about which tasks to perform and how or when to perform them. For example, jobs that are characterized by strictly prescribed courses of action are more likely to restrict the expression of one's individual differences than jobs wherein employees are free to engage in the behaviors they believe are most appropriate.

The fourth facet of situational strength, "consequences," is defined as "the extent to which decisions or actions have important positive or negative implications for any relevant person or entity" (Meyer et al., 2010: 127). This operationalization influences employee behavior by encouraging courses of action that increase the probability of positive outcomes and/or decrease the probability of negative outcomes. For example, even a dispositionally reckless employee is likely to *behave* cautiously when doing so will minimize negative outcomes or maximize positive outcomes.

Using this structure as a guide, 77 initial items were written to cleanly tap a single facet. All items were intentionally written to begin with the phrase "On this job . . ." so that researchers interested in using the SSW at different levels of analysis (e.g., "On this team . . .") can easily do so without fundamentally altering the content or meaning of the items. After each item was carefully refined, an independent sample of participants sorted them on the basis of their content (Hardesty & Bearden, 2004; Hinkin, 1998). A combined Method and Results section is used to describe this process.

## *Method and Results*

### *Participants*

Study 1 participants were 101 undergraduates who worked at least part-time and were enrolled in an introductory management course at a large university located in the mid-Atlantic United States. Here, Hinkin (1998: 109) has argued that "it may be appropriate to use a small sample of students as this is a cognitive task not requiring an understanding of the phenomena under examination." Participants were, on average, 24.1 years old, worked an average of 19.7 hours per week, and had held an average of 1.6 full-time jobs in their lifetime. Participants were also diverse with respect to job title, gender (54.7% female), ethnicity (39.6% Asian or Pacific Islander; 31.7% White, non-Hispanic; 8.9% Black, non-Hispanic; 6.9% Hispanic or Latino/a; and 12.9% Other), and primary language (although all participants reported being relatively fluent in English, 46.5% indicated that it was not their first language).

### *Procedures*

*Sorting task*. Participants were presented with the names and construct definitions of each of the four facets of situational strength (Meyer et al., 2010). They were then presented with a randomized list of all of the items included in the original bank (14 items for clarity, 20 for consistency, 24 for constraints, and 19 for consequences). The practice of developing unevenly sized item pools is consistent with contemporary instrument development standards, in the sense that "it is impossible to specify the number of items that should be included in an initial pool" (DeVellis, 1991: 57) because "at least twice as many items as will be needed in the final scales should be generated" (Hinkin, 1998: 109). Participants were asked to place each item into one of five categories (Hardesty & Bearden, 2004). The first four categories were represented by one of the four facet names and definitions (from Meyer et al., 2010), and the fifth (i.e., "miscellaneous") category was used for items adjudged to fit into none or more than one of the four facets—that is, for "awkward or confusing items" (DeVellis, 1991: 76). The results of this activity were then used to create the following item sorting quality index.

*Sorting quality index*. First, the percentage of participants who sorted each item into the "correct" facet was calculated to favor items that performed as intended. Second, the standard deviation of the percentage of participants who sorted each item into each alternative (i.e., "incorrect") facet was calculated to favor items for which sorting errors were distributed homogenously (Hinkin, 1995). Third, the proportion of participants who sorted each item correctly was compared across demographic subgroups (i.e., men vs. women, those who work fewer than 10 hours per week vs. those who work 10 or more hours per week, those with vs. without full-time work experience, and native vs. nonnative speakers of English) to favor items that were sorted similarly by participants from diverse backgrounds. The final sorting index was created by summing the $z$ scores of each of the aforementioned

three indicators. The 13 items within each facet that had the highest score on the sorting index were retained for use in Study 2. None of the reverse-scored items survived this process, which is not surprising because such items are often confusing to participants and tend to exhibit poor psychometric properties (Hinkin, 1998; Netemeyer et al., 2003).

It is important to note that the comparison of native versus nonnative speakers of English identified six otherwise well-performing items that showed substantial differences across these two groups. Specifically, all six of these items were categorized "correctly" to a greater extent by native than nonnative English speakers, suggesting the possible existence of difficult and/or colloquial language. Consequently, three general changes were made to these items: (a) the word *expectations* was replaced with clearer alternatives (e.g., *informal guidelines*), (b) passive-voice items were rephrased in the active voice, and (c) variations of the word *conflict* (when used as a *verb*) were replaced with an appropriate variation of the word *interfere*. This type of analysis is important because questionnaires that are hard to read are likely to exhibit psychometric problems such as lower reliability and more missing data (Stone, Stone, & Gueutal, 1990). Finally, in an effort to prevent range restriction in the consequences facet (results from a separate pilot sample of 63 employed undergraduate students suggested that these participants disproportionately tended to view their jobs as highly consequential), four additional items that were deliberately written to be more extreme were added to the measure of consequences after the sorting task.

### *Discussion*

Study 1 was designed to reduce the initial bank of items to a more manageable subset via a formalized assessment of content validity (Hinkin, 1998). A total of 13 items within each facet were retained, and 4 relatively extreme items were added to the consequences subscale to ensure an adequate range in scores on that facet. This process resulted in 56 items, which were further refined in the next study.

## Study 2—Secondary Screening and Convergent and Discriminant Validities

The purpose of Study 2 is to (a) determine the final SSW scale (the statistical and conceptual considerations used to select the strongest items are described in detail in the Item Retention portion of the Study 2 Results section), (b) assess its psychometric properties, (c) test its facet structure, and (d) examine its convergent and discriminant validities (DeVellis, 1991). These goals are an important part of the instrument development process, in that they collectively work to ensure that the final set of items is not only psychometrically sound but also behaves in a manner that is consistent with underlying theory, thereby providing construct validity evidence by elucidating its nomological network (DeVellis, 1991).

## *Hypotheses*

*Facet structure*. Given that the SSW was developed to reflect Meyer et al.'s (2010) four-facet structure (outlined previously), this study's first hypothesis tests the extent to which this goal is met.

*Hypothesis 1:* A confirmatory factor analysis will reveal that the intended four-facet model fits the data better than several plausible alternatives.

We now develop several hypotheses concerning convergent and discriminant validation. In this process, researchers traditionally demonstrate that measures of the same construct are more strongly related than measures of different constructs (Campbell & Fiske, 1955). However, because no other measures of situational strength exist, we take a similar but broader approach by comparing the magnitude of *empirical* relationships between constructs that should be more versus less strongly related to situational strength from a *theoretical* or *conceptual* standpoint.

*Clarity*. Feedback (i.e., information from external sources about work-relevant behaviors; Kluger & DeNisi, 1996) overlaps conceptually with clarity because both are focused on the extent to which job-relevant information is provided to employees by an external source. Task significance (i.e., "the degree to which the job has a substantial impact on the lives or work of other people"; Hackman & Oldham, 1976: 257), on the other hand, is conceptually less related to clarity because the former is focused on the importance of one's work whereas the latter is focused on the understandability of relevant cues. We therefore predict that clarity will show a stronger positive correlation with feedback than with task significance.

*Hypothesis 2:* Clarity will be more strongly positively related to feedback than to task significance.

*Consistency*. Role conflict (i.e., the extent to which various job requirements are incompatible or incongruent with each other; Rizzo, House, & Lirtzman, 1970) overlaps conceptually with (in)consistency because both deal with the extent to which various sources of behaviorally relevant information convey similar versus dissimilar messages (although role conflict focuses specifically on role-relevant behaviors whereas consistency attempts to capture more diverse sources of information, such as informal guidance and supervisors' instructions). Task significance (defined previously), on the other hand, is conceptually less related to consistency because the former is focused on the importance of various aspects of one's work whereas the latter is a function of the agreement between sources of work information. We therefore predict that consistency will show a stronger negative correlation with role conflict than with task significance.

*Hypothesis 3:* Consistency will be more strongly negatively related to role conflict than to task significance.

*Constraints*. Autonomy (i.e., "the degree to which the job provides substantial freedom, independence, and discretion to the individual in scheduling the work and determining the procedures to carry it out"; Hackman & Oldham, 1976: 258) overlaps conceptually with the inverse of constraints because they both assess the extent to which behavioral options are influenced by some outside source (although "constraints" is a broader construct that attempts to capture the extent to which employees' behavioral discretion is minimized by *any* source of outside information whereas "autonomy" focuses on the freedom granted by the job itself). Role ambiguity (i.e., a lack of predictability regarding outcomes and/or the appropriateness of particular behaviors; Rizzo et al., 1970), on the other hand, is conceptually less related to constraints because the former pertains to one's certainty regarding responsibilities whereas the latter refers to the extent to which one's freedom at work is restricted by others. We therefore predict that constraints will show a stronger negative correlation with autonomy than with role ambiguity.

*Hypothesis 4:* Constraints will be more strongly negatively related to autonomy than to role ambiguity.

*Consequences*. Production responsibility (i.e., "the cost of errors in terms of both lost output and damage to expensive equipment"; Jackson, Wall, Martin, & Davids, 1993: 754) overlaps conceptually with consequences because both deal with the impact of one's work on important outcomes. Autonomy (as defined previously), on the other hand, is conceptually less related to consequences because the former deals with the amount of choice one has in one's work whereas the latter deals with the importance of work outcomes. Although one could argue that, as the consequences of a job increase, employees will tend to have less autonomy, the presence and diversity of many high-consequences jobs that also inherently contain a large amount of autonomy (e.g., CEO, police officer) suggest that this overlap will likely be relatively trivial. We therefore predict that consequences will show a stronger positive correlation with production responsibility than with autonomy.

*Hypothesis 5:* Consequences will be more strongly positively related to production responsibility than to autonomy.

## *Method*

### *Participants*

Study 2 participants were 394 adults working full-time in a variety of geographic areas. These participants were, on average, 40.6 years old, worked an average of 40.5 hours per week, and had worked in their current organization for an average of 8.8 years. The jobs represented in this sample were kept intentionally diverse in an attempt to capitalize on the notion that situational strength varies meaningfully across job types (Meyer et al., 2009). Participants were also diverse with respect to gender (49.7% female), ethnicity (11.0% Asian or Pacific Islander; 78.7% White, non-Hispanic; 3.6% Black, non-Hispanic; 4.9% Hispanic

or Latino/a; 0.5% Native American; and 1.3% Other), and educational attainment (12.6% had completed high school or less, 14.6% had completed a 2-year college degree, 49.7% had completed at least some college, and 23.1% had completed a postgraduate degree).

This sample was gathered via Syracuse University's StudyResponse Project (http://www.studyresponse.net), which is a nonprofit service that facilitates online research by electronically recruiting adult participants (Stanton, 2006; Stanton & Weiss, 2002; Wallace, 2004; for examples of recently published empirical studies that have used the StudyResponse Project, see Piccolo & Colquitt, 2006, and Young, Baltes, & Pratt, 2007).

### *Procedures*

Participants were sent an invitation e-mail containing eligibility requirements, instructions, and a link to an online survey. As per contemporary guidelines regarding ways to reduce the potential effects of common method bias (e.g., Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), each of the scales was separated physically, psychologically, and temporally (e.g., via transitions, unique directions, and visually distinct presentation formats); furthermore, respondents were guaranteed anonymity, which also has been shown to reduce response distortion (Tourangeau & Yan, 2007). Participants who completed the survey were given a gift certificate to a popular online retailer valued at $7.00, the administration of which was managed by the StudyResponse Project (thereby maintaining anonymity). The median survey response time was roughly 15 minutes.

### *Materials*

*Initial situational strength items*. Before testing the aforementioned hypotheses, the 56 items that remained after the initial screening process were further refined (through the procedures described subsequently) to form the final instrument. All items used a 7-point Likert-type scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*) and were phrased so that higher scores indicated stronger situations.

*Convergent and discriminant scales*. Three scales from the Job Diagnostic Survey (Hackman & Oldham, 1974) were used to help assess convergent and discriminant validities. Feedback was assessed via six items (i.e., by combining the subscales for "external agents" and "the job itself," as is commonly done) and yielded an internal consistency reliability estimate (i.e., Cronbach's alpha) of .77; task significance and autonomy were assessed via three items each and yielded internal consistency reliability estimates of .58 and .63, respectively. The latter two reliability estimates, though low, are comparable to those obtained in Hackman and Oldham's (1974) original validation study ($\alpha$ = .66 for both) as well as a subsequent meta-analysis (Fried & Ferris, 1987; $\alpha$ = .67 and .69, respectively). Each of these constructs was measured using a 7-point Likert-type scale ranging from 1 (*very inaccurate*) to 7 (*very accurate*).

Three additional measures were also used to help assess convergent and discriminant validities. Production responsibility ($\alpha$ = .86) was assessed using Jackson et al.'s (1993) five-item instrument, which uses a Likert-type scale ranging from 1 (*not at all*) to 5 (*a great deal*). Role conflict and role ambiguity ($\alpha$ = .91 for both scales) were measured using Rizzo et al.'s (1970) eight- and six-item instruments, respectively, both of which use a Likert-type scale ranging from 1 (s*trongly disagree*) to 7 (s*trongly agree*).

## *Results*

### *Item Retention*

An internal item quality (IIQ) index was developed to select the items that would ultimately serve as the final instrument (Stanton, Sinar, Balzer, & Smith, 2002). Specifically, this index consisted of (a) item-total correlations, (b) each item's sorting quality score (from Study 1), and (c) a "1 versus 0" indicator representing whether or not the item was one of the two "anchor items" per facet (i.e., items that were intentionally written to be virtually identical to the definition of the facet in question, as opposed to merely being broadly consistent with the facet definition). The first two indicators were *z* scored, then all three were averaged to yield each item's IIQ score. The top seven items within each facet were selected for ultimate retention. Cronbach's alpha estimates for the resultant scales were all greater than .85 (clarity = .95, consistency = .90, constraints = .89, consequences = .86).

On examining the content of each facet, however, it was determined that clarity and consistency appeared to be potentially "bloated specific" (Cattell, 1978); that is, the Cronbach's alpha estimates of these facets may have been artificially inflated by the similarity of their content and/or phrasing. In an attempt to minimize this issue, a subject matter expert with more than 20 years of instrument development experience (who was blind to the intent of this activity) independently read each item within these two facets and highlighted those that appeared to contain the most overlap. Two items per facet were ultimately deemed to be overly redundant. Because one of these items (per facet) was intentionally written to be an anchor item, the lead author replaced the nonanchor item with the alternative that appeared to best increase item diversity. The resultant internal consistency reliability estimates were identical to the originals to two decimal points (see Table 1 for each of the final items, plus relevant psychometric information).

### *Hypothesis Tests*

According to Hypothesis 1, the factor structure of the SSW should conform to the four-facet structure proposed by Meyer et al. (2010). A confirmatory factor analysis (CFA) was conducted via Amos version 17.0 to test this prediction. As per contemporary recommendations (Thompson, 2000), fit indices for the proposed four-facet structure were compared to a more parsimonious one-factor baseline model and several plausible alternatives in an attempt to avoid the confirmation bias. Specifically, two primary alternative models were tested here: (a) a two-factor model, in which clarity and consistency loaded on one factor

**Table 1**
**Study 2 Items and Basic Psychometric Information**

| | *M* | *SD* | ITC | Skew | Kurtosis | IIQ |
|---|---|---|---|---|---|---|
| Clarity (α = .95) | | | | | | |
| On this job, specific information about work-related responsibilities is provided. | 5.20 | 1.43 | .83 | –.72 | 0.04 | 0.83 |
| On this job, easy-to-understand information is provided about work requirements. | 4.96 | 1.59 | .85 | –.66 | –0.27 | 0.56 |
| On this job, straightforward information is provided about what an employee needs to do to succeed. | 4.98 | 1.62 | .83 | –.78 | –0.10 | 0.56 |
| On this job, an employee is told exactly what to expect. | 4.77 | 1.62 | .81 | –.60 | –0.54 | 0.35 |
| On this job, precise information is provided about how to properly do one's job. | 4.90 | 1.59 | .81 | –.62 | –0.34 | 0.21 |
| [a]On this job, specific information is provided about which tasks to complete. | 4.99 | 1.54 | .78 | –.69 | –0.28 | –0.50 |
| On this job, an employee is told exactly what is expected from him/her. | 5.07 | 1.58 | .83 | –.78 | –0.11 | 0.08 |
| Consistency (α = .90) | | | | | | |
| On this job, different sources of work information are always consistent with each other. | 4.50 | 1.67 | .69 | –.30 | –0.83 | 1.01 |
| On this job, responsibilities are compatible with each other. | 5.08 | 1.45 | .70 | –.76 | 0.12 | 0.93 |
| On this job, all requirements are highly compatible with each other. | 4.92 | 1.47 | .71 | –.62 | –0.35 | 0.48 |
| On this job, procedures remain completely consistent over time. | 4.52 | 1.75 | .70 | –.40 | –0.90 | 0.33 |
| [a]On this job, supervisor instructions match the organization's official policies. | 5.01 | 1.60 | .65 | –.63 | –0.38 | –0.27 |
| On this job, informal guidance typically matches official policies. | 4.85 | 1.55 | .75 | –.60 | –0.18 | 0.04 |
| On this job, information is generally the same, no matter who provides it. | 4.42 | 1.68 | .70 | –.34 | –0.84 | –0.03 |
| Constraints (α = .89) | | | | | | |
| On this job, an employee is prevented from making his/her own decisions. | 3.81 | 1.85 | .72 | .15 | –1.15 | 0.68 |
| On this job, constraints prevent an employee from doing things in his/her own way. | 4.29 | 1.75 | .71 | –.24 | –0.94 | 0.56 |
| On this job, an employee is prevented from choosing how to do things. | 3.91 | 1.70 | .73 | .05 | –0.94 | 0.54 |
| On this job, an employee's freedom to make decisions is limited by other people. | 4.57 | 1.59 | .72 | –.38 | –0.59 | 0.35 |
| On this job, outside forces limit an employee's freedom to make decisions. | 4.48 | 1.67 | .56 | –.40 | –0.66 | 0.26 |
| On this job, procedures prevent an employee from working in his/her own way. | 4.31 | 1.64 | .74 | –.28 | –0.79 | 0.20 |
| On this job, other people limit what an employee can do. | 4.52 | 1.58 | .65 | –.33 | –0.60 | 0.08 |
| Consequences (α = .86) | | | | | | |
| On this job, an employee's decisions have extremely important consequences for other people. | 5.04 | 1.38 | .60 | –.45 | –0.27 | 1.12 |
| On this job, very serious consequences occur when an employee makes an error. | 4.22 | 1.70 | .63 | –.13 | –0.90 | 0.87 |
| On this job, important outcomes are influenced by an employee's actions. | 5.39 | 1.26 | .51 | –.66 | 0.16 | 0.63 |
| On this job, other people are put at risk when an employee performs poorly. | 4.43 | 1.79 | .67 | –.26 | –0.98 | 0.46 |

*(continued)*

**Table 1 (continued)**

| | *M* | *SD* | ITC | Skew | Kurtosis | IIQ |
|---|---|---|---|---|---|---|
| On this job, mistakes are more harmful than they are for almost all other jobs. | 4.23 | 1.79 | .73 | –.18 | –0.92 | 0.43 |
| On this job, tasks are more important than those in almost all other jobs. | 4.54 | 1.69 | .66 | –.42 | –0.65 | 0.29 |
| On this job, there are consequences if an employee deviates from what is expected. | 4.82 | 1.53 | .56 | –.52 | –0.25 | 0.23 |

*Note:* ITC = item total correlation; IIQ = internal item quality score (for details, see the Item Retention section of the Study 2 results). All items used a Likert-type scale, wherein 1 = *strongly disagree* and 7 = *strongly agree*.
[a]Item added to reduce bloated specificity (Cattell, 1978).

**Table 2**
**Study 2 and Study 4 Confirmatory Factor Analysis Results**

| Model | $\chi^2$ | *df* | $\chi^2/df$ | RMSEA | CFI | TLI | SRMR |
|---|---|---|---|---|---|---|---|
| Study 2 | | | | | | | |
| Four-factor | 861.5 | 344 | 2.50 | .07 | .90 | .89 | .06 |
| Three-factor | 1,117.4 | 347 | 3.22 | .08 | .89 | .87 | .07 |
| Two-factor | 1,546.1 | 350 | 4.42 | .11 | .77 | .75 | .14 |
| One-factor | 2,530.9 | 350 | 7.23 | .15 | .58 | .54 | WNE |
| Study 4 | | | | | | | |
| Four-factor | 1,853.1 | 344 | 5.39 | .09 | .88 | .86 | .06 |
| Three-factor | 2,500.8 | 347 | 7.21 | .10 | .83 | .80 | .07 |
| Two-factor | 4,233.2 | 350 | 12.10 | .14 | .70 | .65 | .16 |
| One-factor | 7,317.1 | 350 | 10.62 | .18 | .46 | .37 | .21 |

*Note:* RMSEA = root mean square error of approximation; CFI = confirmatory fit index; TLI = Tucker–Lewis index; SRMR = standardized root mean square residual; WNE = analysis "would not estimate." Four-factor models represent the predicted structure, wherein each item loads on the intended factor and all factors are allowed to correlate. The one-factor model is a standard comparison suggested by Thompson (2000). The two-factor model represents a theoretically viable alternative in which clarity and consistency load on one factor and constraints and consequences load on a second (Deci & Ryan, 1987). The three-factor model represents a theoretically viable alternative in which clarity and consistency load on one factor and constraints and consequences load on their own factors.

and constraints and consequences loaded on the other, because some theorizing (e.g., Deci & Ryan, 1987) suggests that situational influences on behavior can be defined as either "autonomous" (clarity and consistency) or "controlled" (constraints and consequences), and (b) a three-factor model in which clarity and consistency loaded on a single factor (given the potential for conceptual and empirical overlap between these facets) and constraints and consequences each loaded on their own factor.

The results of these tests support Hypothesis 1 in that the primary model demonstrates an acceptable level of congruence with the intended structure and outperforms each of the comparison models (see Table 2). It is important to stress here that, despite an intercorrelation of .81 between clarity and consistency (see Table 3), the proposed four-factor model

**Table 3**
**Study 2 Correlations**

| | *M* | *SD* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Clarity | 5.00 | 1.37 | (.95) | | | | | | | | | |
| 2. Consistency | 4.75 | 1.26 | .81*** | (.90) | | | | | | | | |
| 3. Constraints | 4.29 | 1.32 | .05 | .00 | (.89) | | | | | | | |
| 4. Consequences | 4.67 | 1.17 | .37*** | .34*** | .47*** | (.86) | | | | | | |
| 5. Feedback | 4.83 | 1.15 | .49*** | .36*** | –.21*** | .13** | (.77) | | | | | |
| 6. Task sig. | 5.05 | 1.21 | .08 | –.01 | .00 | .36*** | .30*** | (.58) | | | | |
| 7. Role conflict | 3.99 | 1.50 | –.30*** | –.22*** | .59*** | .30*** | –.40*** | –.04 | (.91) | | | |
| 8. Role ambiguity | 2.51 | 1.17 | –.73*** | –.65*** | .08 | –.28*** | –.57*** | –.19*** | .34*** | (.91) | | |
| 9. Autonomy | 4.94 | 1.27 | .16** | .17** | –.45*** | –.05 | .39*** | .27*** | –.25*** | –.34*** | (.63) | |
| 10. Prod. resp. | 2.91 | 1.09 | .21*** | .22*** | .27*** | .49*** | .06 | .09 | .33*** | –.20*** | .02 | (.86) |

*Note:* Task sig. = task significance; prod. resp. = production responsibility. The situational strength facets, feedback, task significance, autonomy, role conflict, and role ambiguity were measured on a 1–7 Likert-type scale; production responsibility was measured on a 1–5 Likert-type scale. All significance tests are two-tailed.
*p <.05. **$p < .01$. ***$p < .001$.

performed better on every fit index (root mean square error of approximation [RMSEA], confirmatory fit index [CFI], Tucker–Lewis index [TLI], and standardized root mean square residual [SRMR]) than an alternative three-factor model wherein clarity and consistency loaded on the same factor. Moreover, the chi-square difference test revealed that the four-factor model fit the data significantly better than the three-factor model ($\Delta\chi^2 = 255.9, p < .001$). Thus, given its empirical superiority and theoretical foundation, the aforementioned four-facet structure with seven items per facet was retained as the ultimate SSW scale.

Convergent and discriminant validities for each situational strength facet were tested via Meng, Rosenthal, and Rubin's (1992) two-tailed difference test for dependent correlations. A statistically significant difference between the target facet's correlation with the convergent construct and the discriminant construct indicates support for the hypothesis in question. Because, in some cases, differences between the directions (not just the magnitudes) of the two correlations being compared existed, we opted to take a conservative approach by always testing for the difference between the absolute values of the two correlations.

In support of Hypothesis 2, the difference between clarity's correlation with feedback ($r = .49, p < .001$) and its correlation with task significance ($r = .08, p > .05$) was statistically significant, $t(380) = 7.65$, $p < .001$. In support of Hypothesis 3, the difference between consistency's correlation with role conflict ($r = –.23, p < .001$) and its correlation with task significance ($r = –.01, p > .05$) was statistically significant, $t(385) = 3.23$, $p < .001$. In support of Hypothesis 4, the difference between constraints' correlation with autonomy ($r = –.46, p < .001$) and its correlation with role ambiguity ($r = .08, p > .05$) was statistically significant, $t(381) = 7.07$, $p < .001$. In support of Hypothesis 5, the difference between consequences' correlation with production responsibility ($r = .49, p < .001$) and its correlation with autonomy ($r = –.05, p > .05$) was statistically significant, $t(385) = 6.70$, $p < .001$.

### *Discussion*

Study 2 analyses indicated adequate or better psychometric characteristics and strong evidence of convergent and discriminant validity. Thus, in an effort to contribute to the voluntary work behavior and situational strength literatures (while further contributing to the final instrument's construct validity), we used the SSW to test a series of theoretically based substantive hypotheses. In Study 3, we examine the impact of individual differences on perceptions of situations. Then, in Study 4, we examine the extent to which perceptions of situational strength moderate relationships between individual differences and behavioral outcomes.

## Study 3—Perceptual Influences on Situations

The purpose of Study 3 is to examine the role that individual differences play in perceptions of situational strength (Cooper & Withey, 2009). This is an important issue because strength-relevant cues are unlikely to be objective characteristics of situations; instead, they are likely filtered through one's expectations, experiences, motives, and dispositions (James & McIntyre, 1996; Mischel & Shoda, 1995). This perspective is consistent with what Block and Block (1981) call a "functional" view of situations, which pertains to a given individual's idiosyncratic interpretation of relevant stimuli. The functional view of situations is contrasted with the "physico-biological" view, which pertains to objective characteristics of situations (e.g., temperature, barometric pressure) that are essentially independent of one's perceptions, and the "canonical" view of situations, which pertains to perceivers' generally agreed-upon perceptions.

Consistent with the interactional perspective that drives the current effort, we argue that perceptions of the situation are contingent not only on characteristics of the situation (in this case, its strength) but also on characteristics of the perceiver. That is, we agree with previous theorists who argued that "the most relevant level of situational analysis for theory building and research is the functional situation" (Hattrup & Jackson, 1996: 512) because "contextual factors cannot be disembedded from the psychological meaning given to them by the individual" (Deci & Ryan, 1987: 1033). Thus, it is important to not only quantify interrater agreement within the same situation but also statistically account for the influence that individual differences might have on idiosyncratic perceptions thereof.

This goal is accomplished through two interrelated steps. First, consistent with the idea that "psychological 'situations' (stimuli, treatments) are powerful to the degree that they lead everyone to construe the particular events the same way" (Mischel, 1977: 347), we posit that perceptions of the situation will depend primarily on the situation itself *in strong situations*. Furthermore, consistent with the idea that individual differences "guide us to interpret an ambiguous stimulus in a certain manner; direct our focus of attention on certain aspects of the environment, ignoring others" (Rogers, 1981: 194), we posit that perceptions of the situation will be more greatly influenced by individual differences *in weak situations*. Indeed, when developing the concept of "climate strength," Schneider, Salvaggio, and Subirats (2002: 221) argued that "people in weak situations do not perceive events the same way," thereby

suggesting that an inherent component of weak situations is a relative lack of perceptual consensus. Thus, we posit the following:

*Hypothesis 6:* Interrater agreement regarding the strength of situations will be higher in stronger situations than in weaker situations.

If the previous theorizing is correct, it follows logically that observed agreement within weak situations should increase after accounting for the perceptual influence of relevant individual differences. That is, it should be possible to model the personal factors that lead some types of people to view weak situations differently than other types of people. Unfortunately, however, the theory to make fine-grained predictions about the *specific* individual differences that will influence perceptions of each of the four facets of situational strength does not yet exist. Thus, we examine this question by testing the effects of broad-bandwidth, well-studied individual differences: the "Big Five" personality traits (Goldberg, 1999) and trait positive and negative affect (Watson, Clark, & Tellegen, 1988).

*Hypothesis 7:* Interrater agreement in weak situations will increase after accounting for the influence of participants' personality and trait affect.

## *Method*

### *Participants*

Study 3 participants were 253 adults working full-time in a variety of geographic areas, all of whom were also recruited through the StudyResponse Project (although no overlap existed with participants in any of the other studies). Participants were, on average, 40.4 years old, worked 40.0 hours per week, and had worked in their current organization for 8.5 years. The job titles represented in this sample were, once again, kept intentionally heterogeneous. Participants were diverse with respect to gender (48.2% female), ethnicity (11.6% Asian or Pacific Islander; 74.5% White, non-Hispanic; 4.6% Black, non-Hispanic; 5.3% Hispanic or Latino/a; 1.8% Native American; and 2.2% Other), and educational attainment (9.0% had completed high school or less, 13.1% had completed a 2-year college degree, 51.4% had completed at least some college, and 26.5% had completed a postgraduate degree).

### *Procedures*

Study 3 recruitment procedures were nearly identical to those used in Study 2. The main procedural difference was that, instead of rating their own jobs, Study 3 participants used the SSW to rate the strength of the situational vignettes described subsequently. Participants then responded to a series of individual differences measures, which were used to model the idiosyncratic effects of situational strength perceptions. The median response time in this study was roughly 26 minutes.

*Materials*

*Stimuli*. Eight job vignettes were created to provide a common reference point for participants. All vignettes shared in common an opening paragraph, which described the job of a high school social studies teacher in the Midwestern United States. This paragraph was intentionally written to be moderate vis-à-vis situational strength. Eight statements describing the specific work environment of the job in question were then presented to participants in each vignette. Six of the eight statements were intended to represent moderate situational strength on three of the facets, whereas both of the remaining two statements were intentionally written to represent either a strong situation or a weak situation on the manipulated facet. Thus, there were a total of eight vignettes: a "strong" and a "weak" manipulation for each of the four situational strength facets. We used a within-subjects design, such that each participant responded to two vignettes (i.e., the vignettes involving the "strong" and "weak" levels of one situational strength facet). See the appendix for an example vignette.

*Situational strength*. The SSW (described previously) was used to assess situational strength. Average internal consistency reliability estimates across all vignettes were again greater than .85 for each facet (clarity = .94, consistency = .95, constraints = .95, consequences = .90).

*Job complexity*. Morgeson and Humphrey's (2006) four-item measure of job complexity ($\alpha = .91$) was also included here to show that vignette perceptions not only changed in the ways predicted here, but also did *not* change on a theoretically unrelated situational dimension. In particular, job complexity has not been posited to induce specific behaviors among individuals who are unlikely to engage in them when left to their own devices. Indeed, no known published studies have treated job complexity as an operationalization of situational strength (see Meyer et al., 2010).

*Individual differences*. Individual differences were conceptualized broadly to include trait positive affect (PA), trait negative affect (NA), and each of the Big Five personality traits. PA and NA were measured using Watson et al.'s (1988) Positive and Negative Affect Schedule (PANAS), which consists of 10 PA and 10 NA items (adjectives). Participants rate the extent to which each item describes them on average. Responses were provided on 5-point Likert-type scale ranging from 1 (*very slightly or not at all*) to 5 (*extremely*). Internal consistency reliability estimates were high for both constructs (PA = .91, NA = .92). The Big Five personality constructs were measured using Goldberg's (1999) International Personality Item Pool (IPIP). Ten items were used to assess each of these traits via a Likert-type scale with responses ranging from 1 (*very inaccurate*) to 5 (*very accurate*). Internal consistency reliability estimates were all acceptable (neuroticism = .88, extraversion = .88, openness = .78, agreeableness = .84, and conscientiousness = .86).

## *Results*

### *Manipulation Checks*

Before examining the role of individual differences in perceptions of situational strength, it is necessary to test whether the SSW and the vignettes performed as intended. That is, despite the presumed effects of individual differences on perceptions, a valid measure should be able to detect and model mean changes in situational strength across the stimuli of interest. Consistent with this perspective, paired-samples *t* tests indicate that the strong situation mean for clarity (5.1, *SD* = 1.28) was significantly higher than its weak situation counterpart (3.5, *SD* = 1.86), $t(61) = 6.03$, $p < .001$; the strong situation mean for consistency (5.2, *SD* = 1.04) was significantly higher than its weak situation counterpart (4.1, *SD* = 1.80), $t(63) = 4.51$, $p < .001$; the strong situation mean for constraints (5.4, *SD* = 1.16) was significantly higher than its weak situation counterpart (3.2, *SD* = 1.67), $t(61) = 7.38$, $p < .001$; and the strong situation mean for consequences (5.1, *SD* = 0.95) was significantly higher than its weak situation counterpart (3.8, *SD* = 1.30), $t(63) = 6.45$, $p < .001$. Thus, stimuli that were written to be stronger were perceived as significantly stronger than those written to be weaker. Furthermore, and consistent with the idea that situational strength is not redundant with other characteristics that have been shown to moderate important trait–outcome relationships, mean job complexity scores were unaffected by the manipulation of situational strength across vignettes (the maximum change in mean complexity scores across high versus low manipulations was .01, and all *p* values were greater than .05).

### *Hypothesis Tests*

According to Hypothesis 6, interrater agreement should be higher in strong situations and lower in weak situations. This hypothesis was tested by calculating $r_{wg(j)}$ for each manipulation of each facet. This statistic is a measure of interrater agreement that compares the observed variability in ratings of a single target rated using multiple items (in this case, the seven items per SSW facet) to the variability that would be expected based on chance factors alone (James, Demaree, & Wolf, 1993; LeBreton & Senter, 2008). A value of 1.0 represents perfect agreement, a value of 0.0 represents a complete lack of agreement, and a value of .70 is often used to denote greater-than-chance agreement (LeBreton, Burgess, Kaiser, Atchley, & James, 2003). Although several statistical significance tests have been created to compare observed $r_{wg(j)}$ values (e.g., Cohen, Doveh, & Eick, 2001; Pasisz & Hurtz, 2009), each of these was explicitly created to compare estimates derived from *independent participants* who rate the *same target*. In this study, the *same participants* rated *different targets*. Thus, to circumvent this issue, while also addressing the violation of independence of observations that resulted from these data, tests of dependent variances (Lindell, 2001) were used here. Results generally support this hypothesis, in that observed $r_{wg(j)}$ values were smaller (and, therefore, observed variances significantly larger) in weak situations compared to strong

situations for three of the four comparisons: $r_{wg(j)}$ = .00 versus .84 for clarity, $t(60) = 2.56$, $p < .05$; .33 versus .91 for consistency, $t(62) = 3.43$, $p < .01$; .66 versus .90 for constraints, $t(60) = 2.59$, $p < .05$; and .75 versus .89 for consequences, $t(63) = 1.61$, *ns*. All tests were two-tailed.

According to Hypothesis 7, interrater agreement in weak situations should increase after accounting for the effects of relevant individual differences. This hypothesis was tested by regressing each facet of situational strength on the seven individual differences used here (PA, NA, and the Big Five), then recalculating $r_{wg(j)}$ on the resulting *residuals*, which represent participants' perceptions of situations *after* statistically controlling for the effects of their individual differences profiles. The resultant $r_{wg(j)}$ values were then compared to the original $r_{wg(j)}$ values calculated based on the weak situation vignettes. Because the two variances in question were on different scales (i.e., residualized variances versus those based on the original 1–7 scale), they could not be compared using the Lindell (2001) test. It is important to note, however, that the magnitude of the observed agreement estimates were comparable to those tested in Hypothesis 6 and that all residualized $r_{wg(j)}$ values were greater than or equal to the .70 cutoff sometimes used to denote greater than chance agreement: .00 to .70 for clarity, .33 to .70 for consistency, .66 to .88 for constraints, and .75 to .85 for consequences.

### *Discussion*

Study 3 makes two important contributions. First, it provides additional validity evidence for the SSW, in that changes in manipulated situational strength were reflected in commensurate changes in SSW scores. Second, Study 3 makes a more substantive contribution by showing that, consistent with situational strength theory, (a) interrater agreement was higher in strong situations and lower in weak situations and (b) differences in perceptions of weak situations were in part a function of one's individual differences profile (broadly conceptualized). Furthermore, much of the discrepancy between the relative levels of observed interrater agreement in weak versus strong situations was attributable to participants' individual differences profiles. Given the theoretical basis of the SSW scale, its strong psychometric qualities, and evidence of its ability to account for the role of individual differences-based interpretations of situations, it is now appropriate to examine whether, as predicted by theory, situational strength moderates relevant trait–outcome relationships.

## Study 4—Testing Moderation Hypotheses

The purpose of Study 4 is to examine the moderating effects of situational strength on several personality–outcome relationships. Testing that the effects of situational strength are consistent with extant theory not only is an important part of the scale validation process (DeVellis, 1991) but also tests the claim that relationships between individual differences and behavioral outcomes are stronger in weak situations and weaker in strong situations (e.g., Davis-Blake & Pfeffer, 1989; Hattrup & Jackson, 1996; Meyer & Dalal, 2009;

Mischel, 1977; Snyder & Ickes, 1985; Weiss & Adler, 1984). This general idea forms the basis of the Study 4 hypotheses. The specific theory necessary to make fine-grained predictions about which trait–outcome relationships are most likely to be moderated by the facets of situational strength, however, does not yet exist (Meyer et al., 2010). Consequently, we focus here on trait–outcome relationships that have not yet been examined from an interactional perspective, but for which there are conceptual reasons to expect moderation by situational strength.

With regard to outcomes, we focused on voluntary work behavior (Spector & Fox, 2002)—which consists of both organizational citizenship behavior and counterproductive work behavior—for two reasons. First, because these behaviors are believed to be more discretionary than task performance (i.e., in-role behavior), individual differences are believed to relate more strongly to them than to task performance (Borman & Motowidlo, 1997). This is important because identifying a moderator will be most beneficial if it contributes to our understanding of an already-meaningful trait–outcome relationship. Second, the extent to which situational strength moderates the impact of one of the individual differences studied in the current article—namely, conscientiousness—has already been meta-analytically assessed vis-à-vis task performance (albeit at the occupational level of analysis; see Meyer et al., 2009). These authors discussed the need to study the impact of situational strength on the relationship between individual differences and organizational citizenship behavior (OCB) but were unable to do so themselves. Indeed, the role of situational strength with regard to OCB and/or counterproductive work behavior (CWB) has received little attention in the extant literature (see Beaty et al., 2001, for an exception).

With regard to individual differences, we focused on the personality traits of conscientiousness and agreeableness. Here as well, we chose these constructs for two reasons. First, as mentioned previously, a moderator variable will be most beneficial to the extent that it contributes to our understanding of an already-meaningful trait–outcome relationship. Meta-analytic evidence indicates that conscientiousness is a significant predictor of both OCB and CWB (Dalal, 2005) and that agreeableness is related to CWB (Berry, Ones, & Sackett, 2007; Salgado, 2002) and, potentially, to OCB (Organ & Ryan, 1995). Second, it is also important to test those trait–outcome relationships that are most likely to demonstrate moderation. Here, neurological research suggests that neuroticism (Munafo et al., 2003) and extraversion (Depue & Collins, 1999) are the most biologically engrained of the Big Five personality traits and, therefore, "may prove to be less sensitive to situation factors" (Wood & Beckman, 2006: 459). Because conscientiousness and agreeableness are the only Big Five personality traits to meet *both* of these criteria, we test the following hypotheses:

*Hypothesis 8:* The positive conscientiousness–OCB relationship will be stronger in situations perceived as weak situations than in situations perceived as strong.

*Hypothesis 9:* The positive agreeableness–OCB relationship will be stronger in situations perceived as weak situations than in situations perceived as strong.

*Hypothesis 10:* The negative conscientiousness–CWB relationship will be stronger in situations perceived as weak situations than in situations perceived as strong.

*Hypothesis 11:* The negative agreeableness–CWB relationship will be stronger in situations perceived as weak situations than in situations perceived as strong.

## Method

### Participants

Study 4 participants were 588 adults working full-time in a variety of geographic areas, all of whom were also recruited through the StudyResponse Project (although no overlap existed with participants in the previous studies). Participants were, on average, 39.1 years old, worked 40.9 hours per week, and had worked in their current organization for 6.8 years. The job titles represented in this sample were, again, kept intentionally heterogeneous. Participants were diverse with respect to gender (46.7% female), ethnicity (11.7% Asian or Pacific Islander; 76.3% White, non-Hispanic; 4.7% Black, non-Hispanic; 4.7% Hispanic or Latino/a; 0.9% Native American; and 1.7% Other), and educational attainment (12.5% had completed high school or less, 14.6% had completed a 2-year college degree, 48.9% had completed at least some college, and 24.0% had completed a postgraduate degree).

### Procedures

Study 4 procedures were identical to those used in Study 2. It is important to note here that Study 4 data were collected chronologically before Study 3 data, but this article was intentionally arranged in the present order to facilitate a more logical flow of ideas. The median Study 4 response time was roughly 16 minutes.

### Materials

*Situational strength*. The SSW (described previously) was used to assess situational strength. Internal consistency reliability estimates were again greater than .85 for each facet (clarity = .94, consistency = .91, constraints = .94, consequences = .89). A global situational strength composite ($\alpha$ = .92) was also calculated here. Following the arguments of Wainer (1976) and Dawes (1979) regarding the effectiveness and elegance of linear, unit-weighted combinations, this global composite was based on a simple average (mean) of all 28 situational strength items.

*Predictors*. Conscientiousness ($\alpha$ = .88) and agreeableness ($\alpha$ = .87) were measured via Goldberg's (1999) IPIP. A total of 20 items were used to assess these traits (10 items per trait) via a Likert-type scale with responses ranging from 1 (*very inaccurate*) to 5 (*very accurate*).

*Criteria*. OCB ($\alpha$ = .90) was assessed via Williams and Anderson's (1991) 14-item measure, which uses a Likert-type scale with responses ranging from 1 (*never*) to 5 (*always*). All items were reframed here to be self-reports (as opposed to supervisor reports); examples include "assist supervisor with his/her work (when not asked)" and "adhere to informal rules devised to maintain order." It is important to note here that research supports the use of self-reports when assessing OCB, in that they tend to yield similar scores when compared to

**Table 4**
**Study 4 Correlations**

| | *M* | *SD* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Clarity | 5.13 | 1.36 | (.94) | | | | | | | | |
| 2. Consistency | 4.88 | 1.29 | .74*** | (.91) | | | | | | | |
| 3. Constraints | 4.08 | 1.46 | .05 | .05 | (.84) | | | | | | |
| 4. Consequences | 4.91 | 1.23 | .32*** | .29*** | .33*** | (.89) | | | | | |
| 5. Global SS | 4.75 | 0.91 | .76*** | .75*** | .55*** | .69*** | (.92) | | | | |
| 6. Conscientiousness | 3.97 | 0.71 | .23** | .12** | –.18*** | .12** | .09* | (.88) | | | |
| 7. Agreeableness | 3.83 | 0.70 | .24** | .21*** | –.20*** | .06 | .11** | .52*** | (.87) | | |
| 8. OCB | 3.40 | 0.60 | .26*** | .28*** | .09* | .29*** | .33*** | .18*** | .21*** | (.90) | |
| 9. CWB | 2.03 | 1.26 | –.06 | –.01 | .30*** | .11** | .13** | –.47*** | –.47*** | .17*** | (.96) |

*Note:* SS = situational strength; OCB = organizational citizenship behavior; CWB = counterproductive work behavior. Each of the situational strength facets and CWB are measured on a 1–7 Likert-type scale. OCB, conscientiousness, and agreeableness are measured on a 1–5 Likert scale. All significance tests are two-tailed.
*$p < .05$. **$p < .01$. ***$p < .001$.

supervisory ratings—thereby helping to ameliorate concerns regarding impression management (Allen, 2006).

CWB ($\alpha = .96$) was assessed via Bennett and Robinson's (2000) 19-item measure, which uses a Likert-type scale with responses ranging from 1 (*never*) to 7 (*daily*). Examples include "made an ethnic, religious, or racial remark at work" and "discussed confidential company information with an unauthorized person." Because CWB is frequently performed in a private and unobserved manner, supervisors have little opportunity to detect its occurrence; consequently, supervisor reports of such behavior are likely to be severely contaminated by halo error (Dalal, 2005; Dalal, Baysinger, Brummel, & LeBreton, in press; Sackett, Berry, Wiemann, & Laczo, 2006; Spector & Fox, 2002). Therefore, self-reports may actually be the best way of assessing CWB under conditions of anonymity and when the researchers are unconnected with the organization (e.g., Aquino, Galperin, & Bennett, 2004)—both of which were the case for the present data.

## *Results*

### *Reevaluation of the SSW's Factor Structure*

Additional evidence of the SSW's factor structure was obtained by conducting a second CFA using procedures that were identical to those utilized previously. This analysis provides additional support for Hypothesis 1. Again, despite an intercorrelation of .74 between clarity and consistency (see Table 4), the proposed four-factor model performed better on every fit index (RMSEA, CFI, TLI, and SRMR) than an alternative three-factor model wherein clarity and consistency loaded on the same factor (see, again, Table 2). Moreover, the chi-square difference test revealed that the four-factor model fit the data significantly better than the three-factor model ($\Delta\chi^2 = 647.7$, $p < .001$).

**Table 5**
**Situational Strength's Moderating Effects on the Conscientiousness–OCB Relationship**

| | *r* | β | $R^2$ | $\Delta R^2$ |
|---|---|---|---|---|
| Conscientiousness | .18*** | .16*** | .032 | .032*** |
| Global situational strength | .33*** | .30*** | .133 | .101*** |
| Conscientiousness × global situational strength | –.16*** | –.10** | .144 | .010** |
| Conscientiousness | .18*** | .13*** | .032 | .032*** |
| Clarity | .26*** | .25*** | .084 | .052*** |
| Conscientiousness × clarity | –.10* | –.14*** | .102 | .019*** |
| Conscientiousness | .18*** | .16*** | .032 | .032*** |
| Consistency | .28*** | .26*** | .098 | .066*** |
| Conscientiousness × consistency | –.10* | –.13*** | .114 | .016*** |
| Conscientiousness | .18*** | .20*** | .032 | .032*** |
| Constraints | .09* | .14*** | .047 | .015** |
| Conscientiousness × constraints | –.13*** | –.12*** | .065 | .017*** |
| Conscientiousness | .18*** | .14*** | .032 | .032*** |
| Consequences | .29*** | .29*** | .114 | .081*** |
| Conscientiousness × consequences | –.10* | –.09* | .123 | .009* |

*Note:* OCB = organizational citizenship behavior. $N = 581$. Beta weights are standardized regression coefficients from the final step in the analysis. All significance tests are two-tailed. The $R^2$ and $\Delta R^2$ columns represent unadjusted values; parsimony adjusted values do not differ appreciably.
*$p < .05$. **$p < .01$. ***$p < .001$.

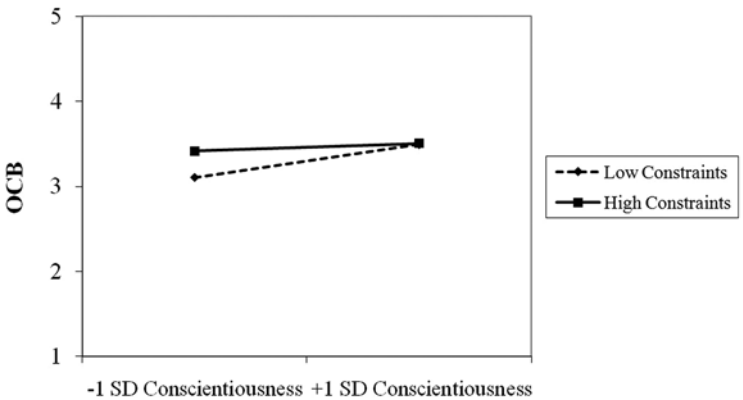
*Hypothesis Tests*

The remaining hypotheses were tested using moderated multiple regression, wherein the standardized Big Five predictor of interest was entered in Step 1, the standardized situational strength facet of interest was entered in Step 2, and the cross-product of these two standardized predictors was entered in Step 3. This three-step procedure was used here (in lieu of the more typical and parsimonious two-step procedure) to isolate situational strength's main and moderator effects, that is, to better estimate the amount of additional variance that was explained by including situational strength in the predictive equations. The decision to use standardized scores (i.e., mean-centered scores divided by their standard deviation) in lieu of mean-centered scores does not affect conclusions associated with tests of either the regression coefficient or the incremental validity associated with the interaction term (Aguinis & Gottfredson, 2010). All significance tests were two-tailed.

According to Hypothesis 8, perceptions of situational strength should moderate the relationship between conscientiousness and OCB, such that observed correlations should be stronger in weak situations and weaker in strong situations. This hypothesis was fully supported, in that all five interaction terms were statistically significant (see Table 5 for a summary). Visual examinations of the resultant interaction plots confirmed that all effects were in the hypothesized direction (see Figure 1 for an example).

According to Hypothesis 9, perceptions of situational strength should moderate the relationship between agreeableness and OCB, such that this relationship would be stronger in weak situations and weaker in strong situations. This hypothesis was generally supported, in

**Figure 1**
**Plot of a Significant Interaction Between the Big Five Trait of Conscientiousness and the Situational Strength Facet of Clarity in the Prediction of Organizational Citizenship Behavior (OCB)**




*Note:* As hypothesized, the strength of relevant relationships was stronger in weak situations and weaker in strong situations.

that four of the five interaction terms were statistically significant and the remaining effect (i.e., for clarity) was marginally significant (see Table 6 for a summary). Visual examinations of the resultant interaction plots confirmed that all effects were in the hypothesized direction.

According to Hypothesis 10, perceptions of situational strength should moderate the relationship between conscientiousness and CWB, such that this relationship should be stronger in weak situations and weaker in strong situations. This hypothesis was not supported. Although all five interaction terms were statistically significant (see Table 7 for a summary), visual examinations of the resultant interaction plots indicated that all effects were in the direction opposite to that predicted on the basis of extant theory (see Figure 2 for an example). In other words, the observed relationships were stronger in strong situations and weaker in weak situations.

According to Hypothesis 11, perceptions of situational strength should moderate the relationship between agreeableness and CWB, such that this relationship should be stronger in weak situations and weaker in strong situations. This hypothesis, too, was not supported. Although all five interaction terms were statistically significant (see Table 8 for a summary), visual examinations of the resultant interaction plots indicate that all effects were once again consistently in the direction opposite to that predicted on the basis of extant theory.

## *Discussion*

The Study 4 CFA provided independent support for the SSW's facet structure, which is important because "it is preferable to validate a factor structure across different samples and

**Table 6**
**Situational Strength's Moderating Effects on the Agreeableness–OCB Relationship**

| | *r* | β | $R^2$ | $\Delta R^2$ |
|---|---|---|---|---|
| Agreeableness | .21*** | .19*** | .045 | .045** |
| Global situational strength | .33*** | .30*** | .142 | .097** |
| Agreeableness × global situational strength | .11** | –.08* | .149 | .006* |
| Agreeableness | .21*** | .17*** | .045 | .045*** |
| Clarity | .26*** | .23*** | .092 | .047*** |
| Agreeableness × clarity | –.04 | –.07 | .097 | .005 |
| Agreeableness | .21*** | .17*** | .045 | .045*** |
| Consistency | .28*** | .24*** | .101 | .056*** |
| Agreeableness × consistency | –.10* | –.10** | .112 | .011** |
| Agreeableness | .21*** | .25*** | .045 | .045*** |
| Constraints | .09* | .14*** | .063 | .018*** |
| Agreeableness × constraints | –.09* | –.10** | .074 | .011** |
| Agreeableness | .21*** | .19*** | .045 | .045* |
| Consequences | .29*** | .29*** | .129 | .084*** |
| Agreeableness × consequences | –.06 | –.07* | .136 | .006* |

*Note:* OCB = organizational citizenship behavior. $N = 581$. Beta weights are standardized regression coefficients from the final step in the analysis. All significance tests are two-tailed. The $R^2$ and $\Delta R^2$ columns represent unadjusted values; parsimony adjusted values do not differ appreciably.
*$p < .05$. **$p < .01$. ***$p < .001$.

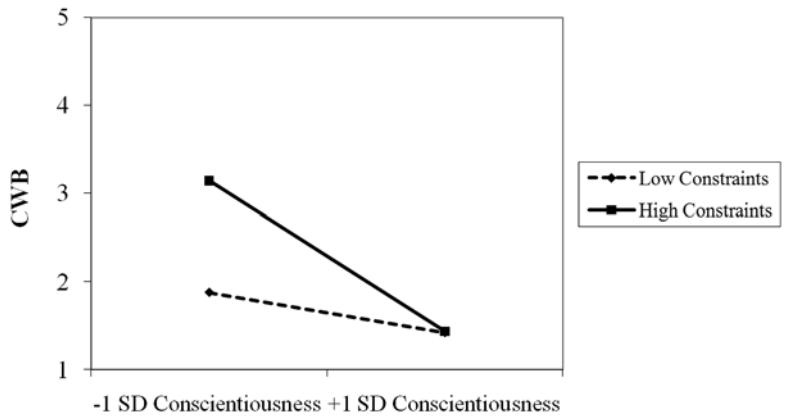
**Table 7**
**Situational Strength's Moderating Effects on the Conscientiousness–CWB Relationship**

| | *r* | β | $R^2$ | $\Delta R^2$ |
|---|---|---|---|---|
| Conscientiousness | –.47*** | –.47*** | .219 | .219*** |
| Global situational strength | .13** | .14*** | .253 | .034*** |
| Conscientiousness × global situational strength | –.28*** | –.22*** | .307 | .054*** |
| Conscientiousness | –.47*** | –.47*** | .219 | .219*** |
| Clarity | –.06 | .07 | .222 | .003 |
| Conscientiousness × clarity | –.20*** | –.18*** | .251 | .030*** |
| Conscientiousness | –.47*** | –.46*** | .219 | .219*** |
| Consistency | –.01 | .07 | .222 | .003 |
| Conscientiousness × consistency | –.19*** | –.15*** | .246 | .024*** |
| Conscientiousness | –.47*** | –.43*** | .219 | .219*** |
| Constraints | .30*** | .26*** | .272 | .053*** |
| Conscientiousness × constraints | –.23*** | –.25*** | .342 | .070*** |
| Conscientiousness | –.47*** | –.49*** | .218 | .218*** |
| Consequences | .11** | .16*** | .242 | .025*** |
| Conscientiousness × consequences | –.18*** | –.15*** | .267 | .026*** |

*Note:* CWB = counterproductive work behavior. $N = 582$. Beta weights are standardized regression coefficients from the final step in the analysis. All significance tests are two-tailed. The $R^2$ and $\Delta R^2$ columns represent unadjusted values; parsimony adjusted values do not differ appreciably.
*$p < .05$. **$p < .01$. ***$p < .001$.

to use the same method, either EFA or CFA, in both samples" (Kline, 2005: 205). The results of the Study 4 hypothesis tests provide further support for the quality of the SSW as a measure. Specifically, 9 of 10 relationships were statistically significant (the remaining one was

**Figure 2**
**Plot of a Significant Interaction Between the Big Five Trait of Conscientiousness and the Situational Strength Facet of Constraints in the Prediction of Counterproductive Work Behavior (CWB)**




*Note:* Contrary to theory, the strength of relevant relationships was stronger in strong situations and weaker in weak situations.

marginally significant) in the predicted direction for OCB, and all 10 were statistically significant in a direction opposite to that predicted for CWB. As opposed to serving as a strike against the construct validity of the SSW, however, the consistency of the latter finding suggests the potential presence of an effect that is worthy of additional theoretical attention—an issue that, in conjunction with the overall implications of all four studies, is explored further in the General Discussion section.

## General Discussion

### *Implications and Future Research*

The current findings suggest that the SSW scale demonstrates adequate psychometric qualities. Specifically, a rigorous item retention strategy ensured content validity, multiple CFAs supported the predicted four-factor solution (Meyer et al., 2010), each subscale demonstrated strong internal consistency reliability across three samples, and all subscales were sensitive to manipulations of situational stimuli. More substantively, results indicate strong evidence of convergent and discriminant validities, patterns of perceptions that are influenced by individual differences in ways that were consistent with underlying theory, moderating effects in the predicted direction for OCB, and moderating effects in the opposite direction for CWB. The following paragraphs outline three general implications of these findings.

**Table 8**
**Situational Strength's Moderating Effects on the Agreeableness–CWB Relationship**

| | *r* | β | $R^2$ | $\Delta R^2$ |
|---|---|---|---|---|
| Agreeableness | –.47*** | –.47*** | .223 | .224*** |
| Global situational strength | .13*** | .16*** | .261 | .037*** |
| Agreeableness × global situational strength | –.25*** | –.16*** | .289 | .029*** |
| Agreeableness | –.47*** | –.48 | .223 | .224*** |
| Clarity | –.06 | .07 | .225 | .004 |
| Agreeableness × clarity | –.14** | –.09* | .232 | .008* |
| Agreeableness | –.47*** | –.49*** | .223 | .224*** |
| Consistency | –.01 | .10** | .234 | .010** |
| Agreeableness × consistency | –.16*** | –.11*** | .248 | .014*** |
| Agreeableness | –.47*** | –.41*** | .224 | .224*** |
| Constraints | .30*** | .24*** | .274 | .050*** |
| Agreeableness × constraints | –.22*** | –.20*** | .318 | .044*** |
| Agreeableness | –.47*** | –.48*** | .224 | .224*** |
| Consequences | .11** | .14*** | .242 | .018*** |
| Agreeableness × consequences | –.16*** | –.11*** | .257 | .015*** |

*Note:* CWB = counterproductive work behavior. $N = 582$. Beta weights are standardized regression coefficients from the final step in the analysis. All significance tests are two-tailed. The $R^2$ and $\Delta R^2$ columns represent unadjusted values; parsimony adjusted values do not differ appreciably.
*$p < .05$. **$p < .01$. ***$p < .001$.

First, the SSW provides a common way for authors of future primary studies to conceptualize and assess situational strength. Although Barrick and Mount (1993), Beaty et al. (2001), Masood, Dani, Burns, and Backhouse (2006), and many others have used situational strength to frame important questions pertaining to person–situation interactions, they were forced to utilize ad hoc and impoverished operationalizations and measures of this concept (Cooper & Withey, 2009; Meyer et al., 2010). The development and validation of an instrument that can be used to assess situational strength in a standardized manner is a meaningful advancement that will permit future researchers to better address a variety of important research questions (some of which are outlined subsequently).

Second, the current efforts show that accounting for situational strength can increase practitioners' ability to predict a variety of organizationally relevant phenomena. For example, adding situational strength to the Study 4 equations (including both its main and interactive effects) accounted for an average of 7.3% and 5.5% additional variance (in OCB and CWB, respectively) beyond the trait of interest. Thus, in addition to being used as a research instrument, the SSW may be able to serve as a companion to traditional job analytic tools, thereby yielding behaviorally relevant information that is traditionally missed by common approaches (Barrick & Mount, 2005; Johns, 2006; Murphy & Dzieweczynski, 2005). It is our hope that this information can then be used to inform selection, motivation, job design, and training decisions (to name a few).

Last, these findings have important implications for future theoretical development. Most notably, the fact that relationships between the individual differences examined here and CWB were, contrary to expectations, stronger in strong situations and weaker in weak situations

suggests the presence of a complex interplay between (a) employees' individual differences, (b) the ways in which employees filter situational strength through their trait profiles to give their work environment psychological meaning, and (c) the ways in which this meaning is manifested via subsequent behavioral reactions. Although space considerations prevent a full discussion of relevant issues, we predict that fundamental processes such as self-determination (Deci & Ryan, 1987), differential framing (James & McIntyre, 1996), and psychological reactance (J. W. Brehm, 1966; S. S. Brehm & Brehm, 1981) will play an important role in our understanding of this unexpected, yet robust, finding.

For example, consistent with the idea that humans have a relatively universal need to control their environments and behaviors (Deci & Ryan, 1987), we posit that key individual differences will influence the extent to which this need is perceived to be thwarted by strong situations. That is, the manner in which a person frames the situation will influence his or her subsequent responses. For example, employees who are characterized by a high need for achievement may view instantiations of extremely high clarity (e.g., very specific instructions from one's boss) as an indicator of a lack of trust, whereas employees who are characterized by a high fear of failure may view the same situation as an indicator of support. Thus, the same situation is framed by one employee as a threat to self-determination (thereby resulting in psychological reactance and CWB) but is framed by another employee as an act of care (thereby resulting in increased perceptions of support and OCB). Thus, providing a universal, theoretically grounded framework for examining the processes that can be used to explain the countertheoretical findings obtained in Study 4 would be a valuable contribution that could be directly tested using the SSW.

### *Limitations and Conclusions*

Some might argue that the effects of this study may be artifacts of common method bias (i.e., that parameter estimates may be inflated by the fact that much of the data used in this study were collected via self-report). In addition to the practical steps that were intentionally taken to minimize this possibility (discussed previously), various lines of research suggest that common method variance (CMV) can only *attenuate* interaction effects such as those consistently observed in Study 4 (e.g., Evans, 1985; Siemsen, Roth, & Oliveira, 2010). Indeed, after conceptually and empirically analyzing the effects of CMV, Siemsen and his colleagues (2010: 470) ultimately concluded that

> empirical researchers should not be criticized for CMV if the main purpose of their study is to establish interaction effects. On the contrary, finding significant interaction effects despite the influence of CMV in the data set should be taken as strong evidence that an interaction effect exists.

Given that person–situation interactions are the main focus of situational strength theory, and that predicting and detecting such interactions was one of the primary foci of the current research, we argue that the nature of the data used here does not substantively alter our conclusions.

A second potential limitation of this article (and, more broadly, the situational strength literature as a whole) is the fact that sufficiently granular theory does not yet exist to inform hypotheses about the ways in which specific individual differences influence perceptions of situational strength and the ways in which each facet of situational strength affects specific behaviors and relevant outcomes. Although the current results represent an important step forward, we hope that the development and validation of the SSW in Studies 1 and 2, the promising results obtained using the broad approach used in Study 3 (i.e., examining the perceptual influences of the Big Five, PA, and NA), and the countertheoretical results obtained in Study 4 will encourage and enable more fine-grained investigations into this important line of inquiry. Said differently, we hope that future theory and research will build on the foundations put in place by the present article to develop a deeper understanding of this important and nuanced construct.

## Appendix

### *Example Vignette (High Clarity Manipulation)*

The job in question is a Social Studies teacher at a public high school located in the Midwestern United States. The person in this position is responsible for teaching 9th- through 12th-grade students about the fundamentals of civics, government, history, and geography. In addition to standard teaching duties, this person serves as the faculty advisor to a student club. This teacher's direct supervisor is the school's principal, but relevant decisions and behaviors can also be influenced by the county's school board, the Parent Teacher Association (PTA), numerous state/federal agencies, and the students themselves. The following characteristics define the specific work environment at this particular school:

- Teachers' non-academic responsibilities and requirements sometimes match their areas of academic expertise. Thus, teachers' non-academic responsibilities are somewhat compatible with their academic duties.
- This school has very easy-to-understand policies in place, which guarantee that teachers know what their responsibilities and requirements are.
- Some aspects of teachers' freedom of choice and action are influenced by the school board and the principal, but teachers can make their own decisions regarding some issues.
- Teachers here who do not improve student learning outcomes may be reprimanded.
- Some of this school's goals are aligned with those of the state and the district, but others are not, so when teachers consult their supervisors, colleagues, or the school's official policy manual, they sometimes receive the same message.
- Teachers here have a moderate amount of freedom. That is, teachers' decisions and actions are jointly influenced by themselves and by relevant others.
- Very specific and straightforward information is provided about what teachers must do to be successful at this school.
- Teacher behavior has a moderate impact on the lives of students because there are mentors in the community who may help students succeed if the teachers do not.

Note: In this sample vignette, the second and seventh statements correspond to the facet being manipulated (in this case, clarity). These statements are worded such that they represent high clarity. The equivalent statements in the counterpart low-clarity vignette were phrased as, "This school has very difficult-to-understand policies in place, which make it impossible to know what teachers' responsibilities are" and "Very vague and confusing information is provided about what teachers must do to be successful at this school." Moreover, in the vignette above, the nonmanipulated facets (all presented at a moderate level) are as follows: the first and fifth statements represent consistency, the third and sixth statements represent constraints, and the fourth and eighth statements represent consequences.

## References


Adler, S., & Weiss, H. M. 1988. Recent developments in the study of personality and organizational behavior. In C. L. Cooper & I. Robertson (Eds.), *International review of industrial and organizational psychology*: 307-330. New York: John Wiley.

Aguinis, H., & Gottfredson, R. K. 2010. Best-practice recommendations for estimating interaction effects using moderated multiple regression. *Organizational Research Methods*, 31: 776-786.

Allen, T. D. 2006. Rewarding good citizens: The relationship between citizenship behavior, gender, and organizational rewards. *Journal of Applied Social Psychology*, 36: 120-143.

Aquino, K., Galperin, B. L., & Bennett, R. J. 2004. Social status and aggressiveness as moderators of the relationship between interactive justice and workplace deviance. *Journal of Applied Social Psychology*, 34: 1001-1029.

Barrick, M. R., & Mount, M. I. 1991. The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44: 1-25.

Barrick, M. R., & Mount, M. K. 1993. Autonomy as a moderator of the relationships between the Big Five personality dimensions and job performance. *Journal of Applied Psychology*, 78: 111-118.

Barrick, M. R., & Mount, M. K. 2005. Yes, personality matters: Moving on to more important matters. *Human Performance*, 18: 359-372.

Beaty, J. C., Cleveland, J. N., & Murphy, K. R. 2001. The relation between personality and contextual performance in "strong" versus "weak" situations. *Human Performance*, 14: 125-148.

Bennett, R. J., & Robinson, S. L. 2000. Development of a measure of workplace deviance. *Journal of Applied Psychology*, 85: 349-360.

Berry, C. M., Ones, D. S., & Sackett, P. R. 2007. Interpersonal deviance, organizational deviance, and their common correlates: A review and meta-analysis. *Journal of Applied Psychology*, 92: 410-424.

Block, J., & Block, J. H. 1981. Studying situational dimensions: A grand perspective and some limited empiricism. In D. Magnusson (Ed.), *Toward a psychology of situations: An interactional perspective*: 85-106. Hillsdale, NJ: Lawrence Erlbaum.

Borman, W. C., & Motowidlo, S. J. 1997. Task performance and contextual performance: The meaning for personnel selection research. *Human Performance*, 10: 99-109.

Brehm, J. W. 1966. *A theory of psychological reactance*. New York: Academic Press.

Brehm, S. S., & Brehm, J. W. 1981. *Psychological reactance: A theory of freedom and control*. New York: Academic Press.

Campbell, D. T., & Fiske, D. W. 1955. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56: 81-105.

Cattell, R. B. 1978. *The scientific use of factor analysis*. New York: Praeger.

Cohen, A., Doveh, E., & Eick, U. 2001. Statistical properties of the rwg(j) index of agreement. *Psychological Methods*, 6: 297-310.

Colquitt, J. A., & Simmering, M. J. 1998. Conscientiousness, goal orientation, and motivation to learn during the learning process: A longitudinal study. *Journal of Applied Psychology*, 83: 654-665.

Cooper, W. H., & Withey, M. J. 2009. The strong situation hypothesis. *Personality and Social Psychology Review*, 13: 62-72.

Cronbach, L. J. 1957. The two disciplines of scientific psychology. *American Psychologist*, 12: 671-684.

Dalal, R. S. 2005. A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology*, 90: 1241-1255.

Dalal, R. S., Baysinger, M., Brummel, B. J., & LeBreton, J. M. in press. The relative importance of employee engagement, other job attitudes, and trait affect as predictors of overall employee job performance. *Journal of Applied Social Psychology*.

Davis-Blake, A., & Pfeffer, J. 1989. Just a mirage: The search for dispositional effects in organizational research. *Academy of Management Review*, 14: 385-400.

Dawes, R. M. 1979. The robust beauty of improper linear models in decision making. *American Psychologist*, 34: 571-582.

Deci, E. L., & Ryan, R. M. 1987. The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology*, 43: 1024-1037.

Depue, R. A., & Collins, P. F. 1999. Neurobiology of the structure of personality: Dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, 22: 491-517.

DeVellis, R. F. 1991. Scale development: Theory and applications. In L. Bickman & D. J. Rog (Eds.), *Applied social research methods series* (Vol. 26). Newbury Park, CA: Sage.

Evans, M. G. 1985. A Monte-Carlo study of the effects of correlated method variance in moderated multiple regression analysis. *Organizational Behavior and Human Decision Processes*, 36: 305-323.

Fried, Y., & Ferris, G. R. 1987. The validity of the job characteristics model: A review and meta-analysis. *Personnel Psychology*, 40: 287-322.

Funder, D. C. 2006. Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality*, 40: 21-34.

Goldberg, L. R. 1999. A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7): 7-28. Tilburg, the Netherlands: Tilburg University Press.

Hackman, J. R., & Oldham, G. R. 1974. *The Job Diagnostic Survey: An instrument for the diagnosis of jobs and the evaluation of job redesign projects* (Tech. Rep. No. 4). New Haven, CT: Yale University Department of Administrative Sciences.

Hackman, J. R., & Oldham, G. R. 1976. Motivation through the design of work: A test of a theory. *Organizational Behavior and Human Performance*, 16: 250-279.

Hardesty, D. M., & Bearden, W. O. 2004. The use of expert judges in scale development: Implications for improving face validity of measures of unobservable constructs. *Journal of Business Research*, 57: 98-107.

Hattrup, K., & Jackson, S. E. 1996. Learning about individual differences by taking situations seriously. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations*: 507-547. San Francisco: Jossey-Bass.

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

Hinkin, T. R. 1998. A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1: 104-121.

Jackson, P. R., Wall, T. D., Martin, R., & Davids, K. 1993. New measures of job control, cognitive demand, and production responsibility. *Journal of Applied Psychology*, 78: 753-762.

James, L. R., Demaree, R. G., & Wolf, G. 1993. rWG: An assessment of within-group interrater agreement. *Journal of Applied Psychology*, 78: 306-309.

James, L. R., & McIntyre, M. D. 1996. Perceptions of organizational climate. In K. R. Murphy (Ed.), *Individual differences and behavior in organizations:* 417-450. San Francisco: Jossey-Bass.

Johns, G. 2006. The essential impact of context on organizational behavior. *Academy of Management Review*, 31: 386-408.

Kline, R. B. 2005. Principles and practice of structural equation modeling. New York: Guilford.

Kluger, A. N., & DeNisi, A. 1996. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119: 254-284.

LeBreton, J. M., Burgess, J. R. D., Kaiser, R. B., Atchley, E. K. P., & James, L. R. 2003. The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar? *Organizational Research Methods*, 6: 80-128.

LeBreton, J. M., & Senter, J. 2008. Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11: 815-852.

Lewin, K. 1936. *Principles of topological psychology*. New York: McGraw-Hill.

Lindell, M. K. 2001. Assessing and testing interrater agreement on a single target using multi-item rating scales. *Applied Psychological Measurement*, 25: 89-99.

Masood, S. A., Dani, S. S., Burns, N. D., & Backhouse, C. J. 2006. Transformational leadership and organizational culture: The situational strength perspective. *Proceedings of the Institution of Mechanical Engineers*, *Part B: Journal of Engineering and Manufacture*, 220: 941-949.

Meng, X. L., Rosenthal, R., & Rubin, D. B. 1992. Comparing correlated correlation coefficients. *Psychological Bulletin*, 1: 172-175.

Meyer, R. D., & Dalal, R. S. 2009. Situational strength as a means of conceptualizing context. *Industrial and Organizational Psychology*, 2: 99-102.

Meyer, R. D., Dalal, R. S., & Bonaccio, S. 2009. A meta-analytic investigation into situational strength as a moderator of the conscientiousness-performance relationship. *Journal of Organizational Behavior*, 30: 1077-1102.

Meyer, R. D., Dalal, R. S., & Hermida, R. 2010. A review and synthesis of situational strength in the organizational sciences. *Journal of Management*, 36: 121-140.

Mischel, W. 1977. The interaction of person and situation. In D. Magnusson & N. S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology*: 333-352. Hillsdale, NJ: Lawrence Erlbaum.

Mischel, W., & Shoda, Y. 1995. A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, 102: 246-286.

Morgeson, F. P., & Humphrey, S. E. 2006. The work design questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology*, 91: 1321-1339.

Munafo, M. R., Clark, T. G., Moore, L. R., Payne, E., Walton, R., & Flint, J. 2003. Genetic polymorphism and personality in healthy adults: A systematic review and meta-analysis. *Molecular Psychiatry*, 8: 471-484.

Murphy, K. R., & Dzieweczynski, J. L. 2005. Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Human Performance*, 18: 343-357.

Netemeyer, R. G., Bearden, W. O., & Sharma, S. 2003. *Scaling procedures: Issues and applications.* Thousand Oaks, CA: Sage.

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

Pasisz, D. J., & Hurtz, G. M. 2009. Testing for between-group differences in within-group interrater agreement. *Organizational Research Methods*, 12: 590-613.

Piccolo, R. F., & Colquitt, J. A. 2006. Transformational leadership and job behaviors: The mediating role of core job characteristics. *Academy of Management Journal*, 49: 327-340.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. 2003. Common method biases in behavioral research. A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88: 879-903.

Rizzo, J. R., House, R. J., & Lirtzman, S. I. 1970. Role conflict and ambiguity in complex organizations. *Administrative Science Quarterly*, 15: 150-163.

Rogers, T. B. 1981. A model of the self as an aspect of the human information processing system. In N. Cantor & J. F. Kihlstrom (Eds.), *Personality*, *cognition*, *and social interaction:* 193-214. Hillsdale, NJ: Lawrence Erlbaum.

Sackett, P. R., Berry, C. M., Wiemann, S. A., & Laczo, R. M. 2006. Citizenship and counterproductive behavior: Clarifying relations between the two domains. *Human Performance*, 19: 441-464.

Salgado, J. F. 2002. The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment*, 10: 117-125.

Schneider, B., Salvaggio, A. N., & Subirats, M. 2002. Climate strength: A new direction for climate research. *Journal of Applied Psychology*, 87: 220-229.

Siemsen, E., Roth, A., & Oliveira, P. 2010. Common method bias in regression models with linear, quadratic, and interaction effects. *Organizational Research Methods*, 13: 456-476.

Snyder, M., & Ickes, W. 1985. Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed.): 883-948. New York: Random House.

Spector, P. E., & Fox, S. 2002. An emotion-centered model of voluntary work behavior: Some parallels between counterproductive work behavior and organizational citizenship behavior. *Human Resource Management Review*, 12: 269-292.

Stagner, R. 1977. On the reality and relevance of traits. *Journal of General Psychology*, 96: 185-207.

Stanton, J. M. 2006. *The StudyResponse open recruitment method* (Tech. Rep. No. 13007). Syracuse, NY: Syracuse University, School of Information Studies.

Stanton, J. M., Sinar, E. F., Balzer, W. K., & Smith, P. C. 2002. Issues and strategies for reducing the length of self-report scales. *Personnel Psychology*, 55: 167-194.

Stanton, J. M., & Weiss, E. M. 2002. *Online panels for social science research: An introduction to the StudyResponse Project* (Tech. Rep. No. 13001). Syracuse, NY: Syracuse University, School of Information Studies.

Stone, E. F., Stone, D. L., & Gueutal, H. G. 1990. Influence of cognitive ability on responses to questionnaire measures: Measurement precision and missing response problems. *Journal of Applied Psychology*, 75: 418-427.

Thompson, B. 2000. Ten commandments of structural equation modeling. In L. Grimm & P. Yarnold (Eds.), *Reading and understanding more multivariate statistics:* 261-284. Washington, DC: American Psychological Association.

Tourangeau, R., & Yan, T. 2007. Sensitive questions in surveys. *Psychological Bulletin*, 133: 859-883.

Wainer, H. 1976. Estimating coefficients in linear models: It don't make no nevermind. *Psychological Bulletin*, 83: 213-217.

Wallace, J. C. 2004. *The StudyResponse Project: A description and evaluation of using standing panels of participants for psychological research* (Tech. Rep. No. 13006). Syracuse, NY: Syracuse University, School of Information Studies.

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.

Weiss, H. M., & Adler, S. 1984. Personality and organizational behavior. *Research in Organizational Behavior*, 6: 1-50.

Williams, L. J., & Anderson, S. E. 1991. Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behavior. *Journal of Management*, 17: 601-617.

Wood, R. E., & Beckman, N. 2006. Personality architecture and the FFM in organisational psychology. *Applied Psychology: An International Review*, 55: 453-469.

Young, L. M., Baltes, B. B., & Pratt, A. K. 2007. Using selection, optimization, and compensation to reduce job/family stressors: Effective when it matters. *Journal of Business and Psychology*, 21: 511-539.