A meta-analytic investigation into the moderating effects of situational strength on the conscientiousness–performance relationship

RUSTIN D. MEYER1☆, REESHAD S. DALAL2 AND SILVIA BONACCIO3

1Purdue University, West Lafayette, IN, U.S.A.
2George Mason University, Fairfax, VA, U.S.A.
3Telfer School of Management, University of Ottawa, Ottawa, Ontario, Canada

Summary

Debates about the utility of conscientiousness as a predictor of job performance have focused primarily on mean effect size estimates, despite theoretical and empirical reasons to expect variability across situations. The present study meta-analytically demonstrates that occupation-level situational strength is one important source of this variability. Consistent with theory, predicted unc()orrected conscientiousness–performance correlations ranged from $r = .09$ to .23 (overall performance) and $r = .06$ to .18 (task performance), with stronger correlations observed in weak occupations. These results highlight the need for continued inquiry into the nature of situational strength, its impact on other predictor–outcome relationships, and the implications of these issues vis-à-vis theory and practice. Copyright © 2009 John Wiley & Sons, Ltd.

Introduction

In his 1956 Presidential address to the American Psychological Association, Lee Cronbach proclaimed that psychology has been “limited by the dedication of its investigators to [experimental psychology or correlational psychology] rather than to scientific psychology as a whole” (Cronbach, 1957; p. 671). Cronbach maintained that human behavior is influenced by both the person and the situation, and urged researchers to value and emphasize both—a perspective known as “interactionism.” Although many organizational scientists endorse interactionism (e.g., Barrick, Mount, & Judge, 2001; Griffin, 1997; Hattrup & Jackson, 1996; Johns, 2001; Tett & Burnett, 2003), one of the primary issues preventing its full implementation is a lack of agreement about the structure of situations (Funder, 2006; Johns, 2006).
Despite this dissensus, however, many have argued that one of the most important characteristics to consider is situational strength (Kanfer, Chen, & Pritchard, 2008a; Snyder & Ickes, 1985), especially as a moderator of personality–outcome relationships (Hough & Oswald, 2008; Murphy & Dziewczynski, 2005). The present study contributes to this perspective by meta-analytically examining the moderating effects of occupation-level situational strength on the conscientiousness–performance relationship, thereby improving our theoretical and practical understanding of this predictor, situational strength, and their combined effects on personnel selection.

**Situational Strength**

Situational strength refers to the idea that various characteristics of situations have the capability to restrict the expression and, therefore, criterion-related validity of non-ability individual differences (Mullins & Cummings, 1999; Snyder & Ickes, 1985; Weiss & Adler, 1984). When these characteristics are absent (i.e., when the situation is “weak”), individuals are more likely to be influenced by their behavioral proclivities because these tendencies are the most accessible sources of information regarding potential responses (Mischel & Peake, 1982). When these characteristics are present, however (i.e., when the situation is “strong”), subsequent behaviors are more homogenous than would be expected based on individuals’ trait profiles because strong situational cues replace individual discretion as the most salient source of behavioral information (Chatman, 1989; Forehand & von Haller Gilmer, 1964; Hattrup & Jackson, 1996; Mischel, 1977). Thus, the psychological outcome of strong situations is behavioral homogeneity, the statistical result is restricted variance in the criterion, and the practical effect is attenuated correlations.

The roots of situational strength appear to date back to the work of Rogers (1954), who argued that individual differences in creativity are most likely when situations provide for psychological freedom, psychological safety, and freedom from critical evaluation. Later, Hauge (1964) argued that job specialization reduces the number and variety of behavioral alternatives, while Forehand and von Haller Gilmer (1964) argued that formal and informal characteristics of situations have the ability to place “bounds upon the set of behaviors that might be selected” (p. 371). More recent discussions, however, appear to be based primarily on ideas outlined by Walter Mischel.

Mischel (1977) argued that strong situations restrict the expression of individual differences to the extent that: (1) behavioral cues are salient and uniformly interpreted, (2) appropriate behaviors are clearly defined, (3) the person possesses the necessary skills to engage in the prescribed behaviors, and (4) adequate incentives exist. The relative presence or absence of these conditions can be communicated through factors such as social information, formal policies and procedures, norms, or job characteristics (Mullins & Cummings, 1999). Situational strength is, therefore, best conceptualized as a multifaceted force that homogenizes behavior by providing information about the most appropriate course(s) of action. As a consequence, it is important to understand both the cumulative strength of a given situation (i.e., the total effects of this force), which can be represented by “a configuration or bundle of stimuli” (Johns, 2006, p. 388), as well as the specific effects of individual facets.

This perspective echoes previous conclusions pertaining to the importance of understanding the effects of composite and multiple criteria (Schmidt & Kaplan, 1971). As with criteria, composite moderators help to inform practice by providing information about overall effects, whereas facet-level moderators emphasize psychological understanding by elucidating the specific behavioral mechanisms in play. Thus, a global situational strength construct is used here to (a) provide the first meta-analytic test of the general idea that an important predictor–criterion relationship varies as a function of
situational strength and (b) examine the practical effects of this variability. But, because the cumulative strength of a given situation is ultimately a function of numerous and diverse pieces of information, it is also important to conduct more fine-grained, facet-based analyses that examine the specific mechanisms through which situational strength likely acts. The foundation of these analyses is developed here by first examining some of the myriad operationalizations of situational strength that have been published in the extant literature.

**Operationalizations of situational strength**

Examples of the diverse operationalizations of situational strength that have been employed in previous research include: situational ambiguity and/or uncertainty (Bowles, Babcock, & McGinn, 2005; Gill, Boies, Finegan, & McNally, 2005; Lissek, Pine, & Grillon, 2006), task structure (Meier, 1970), choice of responses to a problem situation (Withey, Gellatly, & Annette, 2005), situational constraints (LaFrance, Hecht, & Paluck, 2003; Wallace, Paulson, Lord, & Bond, 2005), metafeatures of the human resource management system such as consistently enforced rules (Bowen & Ostroff, 2004; see also Davis-Blake & Pfeffer, 1989), industry norms and market uncertainty (Mullins & Cummings, 1999), perceived control (Lee, Ashford, & Bobko, 1990), feedback regarding specific forms of error prevention (LePine, Hollenbeck, Ilgen, Colquitt, & Ellis, 2002), climate strength (Liao & Chuang, 2004; Schneider, Salvaggio, & Subirats, 2002), and transformational leadership (Masood, Dani, Burns, & Backhouse, 2006). Because simultaneously testing the effects of these individual operationalizations would not only be impossible, but would also provide little in the way of information regarding the psychological mechanisms that underlie these effects, published examples were grouped into two broad categories: “constraints” and “consequences” (see the “Conceptual development” subsection of the “Method” section for procedural details).

This is not to say that other broad operationalizations of situational strength do not exist; rather, these two categories merely represent logically consistent dimensions that can be used to better understand the ways in which situational strength moderates relevant predictor–outcome relationships. But before specific hypotheses about these mechanisms can be derived, it is first important to better understand the behaviors through which the trait in question likely influences the criterion of interest. Thus, the following section explores the conscientiousness–performance relationship in greater detail because the facets of situational strength are posited to moderate this relationship to the extent that they encourage conscientious behaviors among all employees, thereby neutralizing the competitive advantage that conscientious employees typically enjoy (Fleeson, 2007; Withey et al., 2005).

**The Conscientiousness–Performance Relationship**

Conscientious individuals are naturally meticulous, thorough, neat, well-organized, able to hold their impulses in check, dedicated to their goals, persistent, dependable, trustworthy, industrious, and achievement-striving (Digman, 1990; McCrae & John, 1992; Roberts, Chernyshenko, Stark, & Goldberg, 2005). Researchers have found that conscientious employees not only tend to outperform their less conscientious coworkers (e.g., Barrick & Mount, 1991; Hurtz & Donovan, 2000), but are also more punctual (Fallon, Avis, Kudisch, Gornet, & Frost, 2000), have more positive attitudes toward...
goals (Conner & Abraham, 2001), and work more diligently than their non-conscientious counterparts (Gellatly, 1996; Kelly, Johnson, & Miller, 2003).

Given these tendencies, it is not surprising that trait conscientiousness is frequently considered a distal cause of performance-relevant behaviors (Kanfer, Chen, & Pritchard, 2008b). Although the specific nature of this relationship is not yet fully understood, some suggest that trait conscientiousness increases performance at least partially through its impact on motivational states (Parker & Ohly, 2008) such as goal setting (Barrick, Mount, & Strauss, 1993; Judge & Ilies, 2002), self-efficacy (Chen, Casper, & Cortina, 2001; Locke & Latham, 2002), and performance expectancies (Gellatly, 1996). These psychological tendencies are then posited to increase the amount of effort that conscientious employees expend (Mount & Barrick, 1995), the amount of time spent on-task (Biderman, Nguyen, & Sebren, 2007), the quantity and quality of one’s work, and subsequent job knowledge, thereby positively affecting ultimate job performance (see Barrick & Mount, 2000 for more thorough treatment of this process).

Consistent with the primary postulates of situational strength, this study suggests that the conscientiousness–performance relationship will be attenuated in occupations containing cues that encourage conscientious behaviors among all employees. The following section builds a specific theoretical case for the ways in which these behaviors are likely to be influenced by the two broad facets of occupation-level situational strength examined here (i.e., constraints and consequences).

Hypotheses

The first broad operationalization of occupation-level situational strength examined here is “constraints,” defined as “the amount of behavioral/decisional restriction placed on an employee or, conversely, as the amount of autonomy or latitude an employee experiences.” In occupations that are characterized by high constraints, behavioral variability is restricted by external forces (e.g., policies and procedures, government regulations, close supervision) that guide behavior and reduce opportunities for individual discretion. Said differently, high constraints reduce conscientious employees’ natural performance advantage by encouraging goal-setting, high expectancies, and persistence among all employees. Thus, to the extent that constraints exist in a given occupation, the domain of acceptable responses is minimized, state conscientiousness (i.e., conscientious behavior that is induced by situational forces) is increased (Fleeson, 2007), variance in the criterion is restricted, (Snyder & Ickes, 1985) and relevant correlations are attenuated (Johns, 2006).

Conversely, occupations that allow employees to make performance-relevant decisions and to willingly engage in performance-relevant behaviors (i.e., occupations that generally lack constraints) are weaker than those wherein procedures are strictly regimented by an outside source because, in the former, “outcomes depend increasingly on the individual’s own efforts, initiative, and decisions rather than on the adequacy of instructions from the boss or on a manual of job procedures” (Hackman & Oldham, 1976, pp. 257–258, italics in original). The “constraints” operationalization, therefore, taps at least two of the characteristics of situational strength outlined by Mischel (1977): it homogenizes behavior by (1) directing employees’ attention toward those aspects of the task that some external authority has deemed important (i.e., making behavioral cues more salient and uniformly interpreted) and (2) providing external information about what behaviors are most and least appropriate in a given context.

In terms of empirical tests of concepts that are conceptually similar to constraints, the criterion-related validity of conscientiousness has been shown to vary as a function of the perceived job
autonomy of managers (Barrick & Mount, 1993). Consistent with the theoretical perspective presented here, these authors argued that conscientious managers outperform their non-conscientious counterparts in environments that “allow for a great deal of discretion in selecting the appropriate work behaviors to be performed” (p. 117), but that this advantage is minimized when employees are constrained. Additional examples of published operationalizations that fit this category include: task structure (Fleeson, 2007; Meier, 1970); salience of behavioral cues from supervisors (Beaty, Cleveland, & Murphy, 2001); and behavioral rule enforcement (Davis-Blake & Pfeffer, 1989).

Two recent empirical tests have also supported the specific mechanisms through which this facet of situational strength is posited to act. Specifically, conscientiousness was shown to predict provision of effort significantly better when cues indicate that more behavioral options are available, compared to scenarios wherein fewer options are available ($r = .34–.18$, respectively—Withey et al., 2005). Further, situational characteristics that encourage a task-orientation (e.g., tasks that are imposed on individuals and/or create a sense of obligation) were shown to minimize the benefits of trait conscientiousness by encouraging conscientious behaviors among all participants (Fleeson, 2007). Thus, these results provide support for the general mechanism through which the proposed moderation is posited to occur, thereby supporting the underlying premise of the present study.

A number of theoretical and empirical sources of evidence, therefore, predict that constraints will restrict behavioral variability, thereby serving as a psychologically and empirically meaningful moderator of the conscientiousness–performance relationship.

**Hypothesis 1**: The conscientiousness–performance relationship is stronger in occupations low in constraints than in occupations high in constraints.

The second broad operationalization of situational strength examined here, “consequences,” is defined as “the presence of contingencies between one’s decisions or behaviors and the outcomes accruing to oneself, other employees, the organization as a whole, and/or external stakeholders.” In occupations characterized by high consequences, behavioral variability is restricted due to the salient negative effects of failure, thereby encouraging goal-setting and persistence among all employees regardless of their level of trait conscientiousness. When this occurs, the positive effects of trait conscientiousness are minimized because conscientious behaviors are likely to be demonstrated by all employees.

Past theorizing generally supports this basic contention by positing “dangerous situations tend to constrain behavior” (Hattrup & Jackson, 1996, p. 528) and by indicating that a strong emphasis on quality control and error prevention in occupations wherein outcomes are especially important has led organizations to implement a variety of interventions designed to encourage uniform behaviors (e.g. total quality management—Douglas & Judge, 2001; Pfeffer, 1998). Further, other relevant theorizing such as the “low cost hypothesis” posits that the effects of personality on behavior are likely to be reduced as the personal, financial, or psychological costs of a given behavioral alternative increase. The typical explanation of this phenomenon is that the benefits of avoiding negative outcomes begin to trump the benefits of exercising one’s personal will as the negativity of those outcomes increases, thereby homogenizing observed behaviors toward “safer” alternatives (Bekkers, 2005).

Thus, the “consequences” operationalization taps the third and fourth aspects of situational strength outlined by Mischel (1977) because occupations high in consequences are likely to have (1) performance management systems that provide employees with the necessary skills to engage in prescribed behaviors and (2) implicit and explicit incentives/punishers that homogenize task-related behavior toward those that are least likely to lead to negative outcomes. Thus, although conscientious behaviors are certainly beneficial in highly consequential occupations, this is not the same as saying...
that high trait conscientiousness will increase performance. Specifically, in occupations wherein employees’ decisions are not associated with important outcomes (e.g., tour guides), conscientious behavior is primarily a function of conscientious personality; however, in occupations wherein employees’ decisions are associated with important outcomes (e.g., explosives workers), conscientious behavior is also influenced by salient sanctions, well-established norms, and situationally induced task focus (Funder, 2006).

Consistent with this explanation, a large-scale review of experimental economics studies examined the effects of financial incentives (a form of “consequences”) on performance and found that, although they do not always increase the mean level of performance, incentives do tend to restrict performance variability (Camerer & Hogarth, 1999). This review further indicates that these effects are especially pronounced for effort, task-attention, and persistence, indicating that external incentives can functionally substitute for a number of the behavioral mechanisms through which conscientiousness is posited to influence performance.

In a study that directly assessed the differential effects of dispositional versus situational influences on behavior, Thornton and Knox (2002) found that the personal consequences of a decision affected the variability of likely behavioral responses. Specifically, variability in students’ likelihood to protest an administrative decision was high among those who were not likely to be directly affected, whereas most of the students who were likely to be directly affected reported an increased intent to protest. Thus, the presence of personally relevant negative consequences served as a situational cue that homogenized subsequent behaviors—an outcome that is consistent with a situational strength perspective.

A number of theoretical and empirical sources of evidence, therefore, predict that consequences will restrict behavioral variability, thereby serving as a psychologically and empirically meaningful moderator of the conscientiousness–performance relationship.

Hypothesis 2: The conscientiousness–performance relationship is stronger in occupations low in consequences than in occupations high in consequences.

Method

Study selection

Inclusion criteria
The primary studies considered for this meta-analysis were empirical articles that met four criteria. First, each study had to include a correlation between conscientiousness and individual job performance, or sufficient information to calculate such a correlation. Second, each correlation had to be calculated using actual employees; thus, studies in which participants were not employed (e.g., those using undergraduate students) were excluded. Third, all participants had to have the same occupation, which was necessary to model the unique effects of each dimension of situational strength. For example, if a given study reported a correlation between conscientiousness and task performance but included both clerical workers and factory workers, it would be difficult (if not impossible) to separate the effects of the clerical environment from those of the factory environment. Fourth, each study had to report the sample size used to calculate each conscientiousness–performance correlation. This was necessary because the meta-analytic moderator estimation techniques used here weighted each study...
by a function of its sample size (see “analytic procedure”). No studies were ultimately eliminated based on this criterion because unreported information was obtained by contacting primary study authors. Lastly, a small number of studies that met each of these inclusion criteria were later excluded due to data limitations described in the “coding procedure” subsection.

**Literature search**

Usable primary studies were obtained by examining previous personality–performance meta-analyses (e.g., Barrick & Mount, 1991; Hurtz & Donovan, 2000), conducting cited-reference searches to locate studies that cited previous meta-analyses and relevant primary studies, and conducting literature searches using combinations of pertinent terms in PsycINFO and other relevant databases (e.g., Web of Science, Google Scholar). Further, an e-mail was sent to the Academy of Management’s Research Methods Network (RMNET) to solicit unpublished manuscripts. In addition, recent proceedings from the annual meetings of the Society for Industrial and Organizational Psychology and the Academy of Management were manually perused. Finally, of the several consulting firms that were contacted in an effort to obtain additional data, Hogan Assessment Systems Inc. provided data from 39 studies. Thus, the total number of primary studies used in this meta-analysis was 114 (representing 13 of the U.S. Department of Labor’s 23 Standard Occupational Classification codes), the total number of independent correlations was 162, and the total number of participants was 34,659.

**The occupational information network (O*NET)**

**General information**

This study utilized situational data from version 11.0 of the Occupational Information Network (O*NET), a database of occupational characteristics developed to replace the Dictionary of Occupational Titles (DOT). Among other things, O*NET supplies numeric estimates of occupational requirements (i.e., knowledge, skills, and abilities), work contexts, and required activities for over 1100 occupations (referred to as “occupational units” or OUs; Converse, Oswald, Gillespie, Field, & Bizot, 2004; Peterson et al., 2001). These estimates were derived from ratings provided by job incumbents, organizational representatives, and trained occupational analysts (Hubbard et al., 2000; Peterson et al., 2001). Due to the breadth of these data, it is difficult to obtain estimates of their overall reliability and validity. However, estimates calculated using specific portions of O*NET generally demonstrate strong psychometric properties (Campion, Morgeson, & Mayfield, 1999; Converse et al., 2004; Jeanneret & Strong, 2003; McCloy, Waugh, Medsker, Wall, Rivkin, & Lewis, 1999; Peterson et al., 2001). That being said, it is important to note here that all situational information provided by O*NET occurs at the occupation level of analysis. Thus, although meaningful situational variability can be found at multiple levels of analysis, this study will only examine the level afforded by O*NET (we return to this point in the “Strengths and limitations” subsection of the “Discussion”).

**Uses in organizational research**

Several organizational scientists have used O*NET to address substantive questions involving the impact of organizational context. For example, Glomb, Kammeyer-Muller, and Rotundo (2004) used O*NET to conclude that higher levels of emotional labor demands are associated with lower wages for occupations with low cognitive demands, but higher wages for occupations with high cognitive demands. Shaw and Gupta (2004) used O*NET to determine that job complexity misfit has a negative effect on the well-being of low performers, but little effect on high performers. Zimmerman, Christakis, and Vander Stoep (2004) used O*NET to conclude that the higher rates of depression among lower socio-economic status employees may partially be driven by their work environments. Lastly,
Diefendorff, Richard, and Gosserand (2006) used O’NET to determine that “task routineness” moderates the relationship between hesitation and self-management performance, whereas “autonomy” does not. Thus, these studies indicate that O’NET data can be used to help answer myriad questions relevant to the organizational sciences.

Conceptual development

In order to derive broad categories of occupation-level situational strength, the first and second authors located published papers that utilized this construct (e.g., Dierdorff & Surface, 2007; Masood et al., 2006) or closely related concepts such as situational and structural ambiguity (Bowles et al., 2005) or climate strength (e.g., Schneider et al., 2002). They then independently attempted to create the smallest number of broad categories that would reasonably account for this set of operationalizations. Both authors converged on similar two-facet solutions (represented by the aforementioned constraints and consequences categories), and minor disagreements regarding the best placement of a small number of specific operationalizations were resolved through discussion.

The “Occupational Requirements” section of the O’NET content model (http://www.onetcenter.org/content.html; see also Peterson et al., 2001) was then used to locate those O’NET variables that best reflected constraints and consequences. Consistent with related studies (i.e., Glomb et al., 2004; Shaw & Gupta, 2004), we used this category of O’NET variables exclusively when operationalizing situational strength, because it describes the work itself rather than characteristics of typical employees (Peterson et al., 2001). Quantitative estimates of occupational requirements are currently available for the following sub-categories: (1) generalized work activities, defined as “aggregations of similar work activities or behaviors for the accomplishment of major work functions” (Peterson et al., 2001; p. 467), and (2) work context, defined as “the social-psychological and physical conditions under which work is performed” (p. 468).

In order to quantify the two operationalizations of occupation-level situational strength, the first two authors consensually rated whether each item in the aforementioned O’NET sub-categories was highly relevant to situational strength, and, if so, whether it should be assigned to constraints or consequences. (See the Appendix for specific details of the 14 items that were ultimately adjudged relevant. In contrast, 86 items were determined to be non-optimal representations of situational strength, including items such as: “interacting with computers” and “spend time sitting”). The first two authors also discussed whether a high score on each of the final 14 items represented an increase or decrease in occupation-level situational strength; items with high scores representing a decrease in strength were reverse-scored, so that high scores always represented increased strength. This is particularly important because some variable names are misleading; thus, in order to fully understand the nature of each item, it is important to read its full description, as well as its high and low anchors. Cronbach’s $\alpha$s were .72 for constraints and .79 for consequences (these values were calculated using the sample of occupations used in the present meta-analysis, $K = 114$). Scores for constraints and consequences were then used in the weighted least squares (WLS) moderator estimation analyses described in the “analytic procedure” subsection below.

Global situational strength

It should be noted that constraints and consequences were effectively orthogonal ($r = -.03$). Thus, under the common factor model, which is closely related to classical test theory (MacDonald, 1999), it is not possible to posit a traditional higher order factor because there is no shared variance. MacKenzie, Podsakoff, and Jarvis (2005), however, argued that “formative” constructs (constructs caused by, rather than being the cause of, their indicators) can be created with indicators that: (a) represent
characteristics rather than manifestations of the construct, (b) capture unique aspects of the construct space and are therefore not interchangeable, (c) need not be inter-correlated, and (d) need not have common antecedents and consequences. Because each of these conditions applies to the two dimensions of occupation-level situational strength utilized here, a composite representing “global situational strength” was also created to help maintain the construct validity of situational strength and allow for a practical examination of the cumulative effects of situational strength on the conscientiousness–performance relationship.

Meta-analysis procedure

Coding procedure

A three-step coding procedure was used to extract the information necessary for meta-analysis from each primary study. Steps 1 and 2 were conducted independently by the first and third authors (the latter of whom was, at the time, blind to the specific hypotheses). Step 1 consisted of coding the primary studies for various study characteristics (e.g., occupational title, sample size, performance type) and recording (or calculating) the conscientiousness–performance correlation. This step was similar to the procedures used in previous personality–performance meta-analyses (e.g., Barrick & Mount, 1991; Hurtz & Donovan, 2000). Agreement for this step was 93.83%; all discrepancies were resolved through discussion.

Step 2 consisted of searching O*NET to determine which OU best matched the occupation used in each primary study. This step was especially important because the accuracy of the hypothesis tests rested on the assumption that the estimates of occupation-level situational strength derived from O*NET closely resembled those of the occupation used in each primary study. Studies were excluded if the occupations they utilized could not be matched to an O*NET OU (seven cases; e.g., “peace corps volunteer”). Agreement for this step was 92.39%; again, all discrepancies were resolved via discussion.

Step 3 consisted simply of transferring the situational estimates provided by O*NET to the master dataset. Unlike Steps 1 and 2, this required neither human judgment nor manual data entry; hence, it was not independently conducted by the second coder. O*NET estimates of situational conditions for specific occupations can be obtained by: (1) going to http://online.onetcenter.org, (2) entering the desired occupation (e.g., “accountant”) in the “Occupation Quick Search” box located in the upper right-hand corner, (3) selecting the O*NET OU that best matches the occupation of interest, and (4) clicking on the “Details” tab located in the “View Report” toolbar located near the top of the O*NET report.

Three performance criteria were utilized in this study: task performance, contextual performance, and overall performance. Based on Borman and Motowidlo (1993), task performance was defined as “the proficiency with which job incumbents perform activities that contribute to the organization’s technical core either directly by implementing a part of its technological process, or indirectly by providing it with needed materials or services” (p. 73), and contextual performance was defined as activities that “support the organizational, social, and psychological environment in which the technical core must function” (p. 73). Measures of overall performance included those that specifically operationalized this construct or those that included aspects of both task and contextual performance. Whenever available, specific items from the criterion measures used by each primary study were examined for content to assist coding efforts; coder agreement on the type of criterion used (a subset of Step 1 coding) was 92.97%. A meta-analytic correlation matrix describing the relationships among the three criteria, generated from primary study samples that measured more than one criterion, is presented in Table 1. As expected (e.g., Viswesvaran, Schmidt, & Ones, 2005), the three criteria exhibit large positive inter-correlations (based on Cohen’s, 1988, rules-of-thumb).
Analytic procedure

In order to model their moderating effects, constraints, consequences, and global situational strength were used as predictors of conscientiousness–performance correlation coefficients in modified WLS regression equations (Hedges & Olkin, 1985). WLS methods are preferable to other available moderator estimation techniques such as ordinary least squares (OLS) regression and “sub-group analysis” for several reasons (Lipsey & Wilson, 2001). First, WLS weights each study by a function of its sample size (referred to as the inverse variance weight), which allows studies with more participants to have a greater impact on analyses. OLS, on the other hand, treats each study as equally meaningful, regardless of its sample size. Second, WLS and OLS are preferable to sub-group analysis because both allow meta-analysts to simultaneously test multiple, continuously distributed moderators, whereas sub-group analysis requires that continuously distributed moderators be artificially divided into two or more sub-groups, thereby reducing statistical power (Chaplin, 1991; Cohen, 1983). Moreover, a recent Monte-Carlo study (Steele & Kammeyer-Mueller, 2002) demonstrated that WLS outperformed all other meta-analytic moderator estimation methods under virtually all conditions, as it was robust against the negative effects of multicollinearity, small N and K (i.e., the sample size across all primary studies and the number of independent primary studies, respectively), and skewed distributions. Because SPSS calculates meta-analytic WLS significance tests incorrectly (Lipsey & Wilson, 2001), an SPSS-add-on package named ZumaStat (Jaccard, 2000), which calculates them correctly, was used.

Results

Preliminary analyses

Mean effect size estimates

Uncorrected (i.e., observed) mean effect size estimates (r) and effect size estimates corrected for unreliability in the predictor and criterion (ρ) were first calculated for the conscientiousness–performance relationship with respect to each of the three criteria (i.e., overall performance, task performance, and contextual performance). Mean uncorrected correlations were .15 for overall performance, .13 for task performance, and .16 for contextual performance. Mean corrected
correlations were .19 for overall performance, .15 for task performance, and .20 for contextual performance. The uncorrected estimates obtained here do not differ substantially from those found in previous meta-analyses (see Table 2). Thus, restricting the primary studies to those that were homogenous in terms of occupation did not appear to bias the average effect size estimates.

**Moderator detection**

A significant $Q$-test of homogeneity (Hedges & Olkin, 1985) rejects the null hypothesis that effect size estimates are drawn from a single population. In the present study, all $Q$-tests were significant. Moreover, all credibility intervals (based on corrected correlations; Whitener, 1990) were relatively wide and at least approached 0 (see Table 3). These results, coupled with the theoretical rationale outlined previously, indicate that a more accurate picture of the conscientiousness–performance relationship may be gained by elucidating the specific sources of this variability (Lipsey & Wilson, 2001; Schneider & Hough, 1995). Tests of moderation, however, were only conducted on overall and task performance because the number of studies that used contextual performance and met our inclusion criteria was too small to yield stable effects. Before examining the results of these analyses, however, it is first necessary to examine examples of strong and weak occupations in order to assess the validity of the operationalizations used here.

**Table 2.** Mean meta-analytic effect size estimates for the conscientiousness–performance relationship, obtained in this and similar studies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K$</td>
<td>$N$</td>
<td>$r^*$</td>
</tr>
<tr>
<td>Overall</td>
<td>82</td>
<td>10,943</td>
<td>.15</td>
</tr>
<tr>
<td>Task</td>
<td>33</td>
<td>4,528</td>
<td>.13</td>
</tr>
<tr>
<td>Contextual</td>
<td>16</td>
<td>2,570</td>
<td>.16</td>
</tr>
</tbody>
</table>

*Mean uncorrected correlation.

†Mean correlation corrected for unreliability in predictor and criterion.

‡Mean correlation corrected for range restriction as well as unreliability in the predictor and the criterion.

§Mean correlation corrected for range restriction and unreliability in the criterion.

||| \*Value estimated by taking the average of the uncorrected conscientiousness-job dedication and conscientiousness-interpersonal facilitation correlations reported by the original authors.

<table>
<thead>
<tr>
<th></th>
<th>Uncorrected estimates</th>
<th>Corrected estimates†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$Q$</td>
</tr>
<tr>
<td>Overall</td>
<td>.15</td>
<td>137.7*p</td>
</tr>
<tr>
<td>Task</td>
<td>.13</td>
<td>74.1*p</td>
</tr>
<tr>
<td>Contextual</td>
<td>.16</td>
<td>57.1*p</td>
</tr>
</tbody>
</table>

*Note: $CV = 90\%$ credibility interval.

*p < .01.

†Corrected for unreliability in predictor and criterion.

‡Credibility intervals are only calculated using corrected values (Whitener, 1990).
Table 4. Descriptive data for proposed moderators

<table>
<thead>
<tr>
<th>Operationalization of situational strength</th>
<th>Mean</th>
<th>Range</th>
<th>LQ</th>
<th>UQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constraints</td>
<td>40.07</td>
<td>20.21–59.04</td>
<td>32.75</td>
<td>45.18</td>
</tr>
<tr>
<td>Consequences</td>
<td>58.91</td>
<td>43.04–85.21</td>
<td>46.80</td>
<td>69.57</td>
</tr>
<tr>
<td>Global situational strength</td>
<td>49.49</td>
<td>38.48–61.70</td>
<td>43.73</td>
<td>54.89</td>
</tr>
</tbody>
</table>

Note: The values provided in this table were calculated using the sample of occupations used in this study. The theoretical range for each moderator is 0–100. LQ = Lower quartile of the observed range of moderator scores; UQ = upper quartile of the observed range of moderator scores.

Occupational examples

Some of the strongest O’NET occupations with respect to constraints are “subway and streetcar operators,” and “medical transcriptionists,” both of which allow for little discretion regarding how or when to complete relevant tasks, whereas some of the weakest occupations regarding constraints are “recreation workers,” and “fitness trainers,” both of which have substantial flexibility regarding tasks and the completion thereof. Some of the strongest occupations with respect to consequences are “surgeons,” and “first-line supervisors/managers of police and detectives,” both of which can lead to negative outcomes if performed incorrectly, whereas some of the weakest occupations regarding consequences are “telemarketers,” and “tour guides,” neither of which is typically associated with serious repercussions. With respect to global situational strength, some of the strongest occupations are “airline pilots,” “nuclear equipment operation technicians,” and “subway and streetcar operators,” whereas some of the weakest occupations are “curators,” “poets, lyricists, and creative writers,” and “personnel recruiters.” Thus, although it is possible that additional operationalizations could legitimately be developed, the dimensions used here yield face-valid exemplars of strong and weak occupations. Descriptive data (obtained means, ranges, and confidence intervals) for each operationalization of situational strength are provided in Table 4.

Hypothesis tests

For each hypothesis, the operationalization of situational strength in question was entered as the lone predictor variable into a single WLS regression equation, with the conscientiousness–performance relationship serving as the criterion variable. A negative regression coefficient associated with situational strength indicates that, as hypothesized, the magnitude of the conscientiousness–performance relationship increases as occupation-level situational strength decreases. As recommended by Raudenbush (1994), two-tailed t-tests (with df = K – P – 1, where K is the number of studies and P is the number of predictors in the model) were used to determine if the slopes of the regression lines differed significantly from 0. The use of two-tailed t-tests (rather than z-tests or one-tailed t-tests) makes our hypothesis tests especially conservative. Results of these tests are presented in Table 5.

Hypothesis 1 (constraints) stated that the conscientiousness–performance relationship is stronger in occupations wherein employees’ behavioral and decisional discretion is not restricted. In the case of overall performance, constraints significantly moderated the conscientiousness–performance relationship in the expected direction (β = −.16, p < .05). In the case of task performance (which utilized a substantially smaller sample of primary studies), this effect was only marginally significant, but also in the expected direction (β = −.16, p < .10). Thus, Hypothesis 1 was supported in the case of overall performance and marginally supported in the case of task performance.
Hypothesis 2 (consequences) stated that the conscientiousness–performance relationship is stronger in occupations wherein employees’ decisions and actions do not affect important outcomes. The results of this hypothesis test indicated that consequences significantly moderated the conscientiousness–performance relationship in the expected direction for both overall performance ($\beta = -0.23, p < .05$) and task performance ($\beta = -0.21, p < .05$). Thus, Hypothesis 2 was supported with respect to both criteria.

It is also important to note that, consistent with the above tests and the underlying theory of this study, global situational strength had a significant moderating effect in the expected direction for both overall performance ($\beta = -0.26, p < .01$) and task performance ($\beta = -0.22, p < .05$). Further, the “practical significance” of these effects can be examined by calculating expected conscientiousness–performance correlations for occupations representing extreme levels of situational strength (operationalized here as plus and minus two standard deviations of global situational strength). These analyses demonstrate uncorrected effects for overall performance ranging from $r = 0.09$ in prototypically strong occupations such as “nuclear equipment operation technician,” wherein behaviors are regulated by external sources and have the potential to impact others’ health and safety, to $r = 0.23$ in prototypically weak occupations such as “barber,” wherein employees typically have substantial behavioral/decisional latitude and errors generally do not have serious effects. Commensurate effects for task performance ranged from $r = 0.06$ to $0.18$ for the same occupations.

**Discussion**

**Summary of findings**

Despite a long and continuing history of psychological thought stressing the importance of individual differences and situational characteristics on behavior, personnel selection focuses primarily on the former. This study highlights the benefits of also incorporating the latter, by meta-analytically demonstrating that the criterion-related validity of conscientiousness varies as a function of three operationalizations of occupation-level situational strength. Specifically, meta-analytic WLS moderator analyses indicate that the conscientiousness–overall performance and conscientiousness–
task performance relationships are moderated by constraints (albeit with marginal significance in the case of task performance), consequences, and global situational strength. In other words, the present results meta-analytically demonstrate that the criterion-related validity of conscientiousness is higher in characteristically weak occupations than in characteristically strong occupations. Given that these effects were predicted a priori and confirmed via a conservative meta-analytic test, we believe they have several important implications.

Implications for theory and research

This paper adds to existing situational strength theory by positing (and demonstrating) that two dimensions can be used to operationalize this construct at the occupational level of analysis. Moreover, constraints and consequences can be combined to create a formative composite representing global situational strength. Each of these three constructs behaves in a manner consistent with a priori expectations, by influencing the criterion-related validity of a common dispositional predictor (i.e., conscientiousness). Thus, these results not only reinforce the validity and importance of situational strength as a general concept, but they also help to clarify its nature. Although, we do not argue that ours is necessarily the structure of situational strength (more research into this question is certainly needed), we do believe it represents important progress by encapsulating many extant operationalizations in a manner consistent with previous theorizing.

Once the ultimate structure of situational strength is determined, a standardized measure of this construct can be developed, thereby allowing researchers and practitioners to accurately assess its presence in specific occupations. Such an instrument could be used alongside traditional job analytic tools “to analyze the context within which the job is embedded” (Murphy & Dzieweczynski, 2005, p. 349), thereby providing locally collected situational strength scores to be used in conjunction with results such as those presented here to better understand and forecast cross-occupational variability in the validities of relevant predictors. This is especially important given that situational strength is posited to restrict behavior in general, thereby highlighting the importance of studying its effect on a variety of other relationships (e.g., other personality–performance relationships). Thus, the most important implication of this study may be its potential to inform the debate about the utility of personality as a predictor.

Implications for the personality debate

Questions concerning the criterion-related validity of trait-based measures of personality have divided the organizational sciences for decades (Guion, 1998). We believe this manuscript makes a meaningful contribution to the most recent iteration of this debate (see Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007b) by showing that questions such as “does personality predict job performance?” fail to adequately capture this issue’s psychological complexity. Instead, more meaningful answers can be obtained by asking “under what conditions do specific personality traits predict specific outcomes?” Our results demonstrate that the validity of conscientiousness varies in practically and theoretically meaningful ways as a function of occupation-level situational strength. Thus, we agree with those who conclude that conscientiousness can be a meaningful predictor of performance, but we also agree with those who state that “there are still compelling reasons for caution when using personality inventories in personnel selection” (Murphy & Dzieweczynski, 2005, p. 354). Said differently, one way to begin resolving the personality debate is to actively conceptualize the role that situational forces play in the expression of valued work outcomes (Barrick & Mount, 2005; Hattrup
& Jackson, 1996; Hough & Oswald, 2008; Johns, 2006; Tett & Christensen, 2007). Our results begin doing this by demonstrating that situational strength is one such force that deserves continued practical and theoretical attention.

This study also has methodological implications for the personality debate. For example, some have recently pointed out that personality–performance meta-analyses should not focus exclusively on corrected validity estimates because uncorrected estimates better reflect the effects that are expected in practice (Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt, 2007a). These authors go on to state, however, that “corrections for range restriction are appropriate when data are available indicating such restriction has occurred” and that corrections for unreliability “are appropriate only when considering unreliability in the criterion” (p. 1034). Given that situational strength restricts criterion variability, and others have argued that criterion unreliability “is a function of, rather than independent of, the situational variables that moderate validities” (James, Demaree, Mulaik, & Ladd, 1992, p. 3), these corrections may mask the impact of situational strength by assuming its effects are statistical artifacts. Thus, although we agree with the importance of examining uncorrected validity estimates, we also urge further research into the appropriateness of meta-analytic corrections of criterion “deficiencies” that may, in fact, be the logical outcome of situational strength.

Strengths and limitations

This study possesses a number of features that complement and extend recent attempts to identify moderators of the conscientiousness–performance relationship. First, it used an interactional perspective and method to meta-analytically model relevant effects. This included basing hypotheses on a single theoretical framework (situational strength) and testing for moderation using the best available meta-analytic technique (WLS regression; Steele & Kammeyer-Mueller, 2002). Second, it utilized a large \( N \) and \( K \) (especially for overall performance), thereby increasing the likelihood of obtaining stable mean effect sizes and moderator estimates. Third, it estimated mean effect sizes for three forms of job performance (overall, task, and contextual), and moderating effects for two of these (overall and task)—an important step given the multi-dimensional nature of job performance (Borman & Motowidlo, 1997). Finally, this was the first study to meta-analytically test whether situational strength moderates the conscientiousness–performance relationship—or even, more generally, whether any situational characteristic moderates any personality–behavior relationship.

The primary limitation of this study stems from its use of O’NET estimates, which are essentially averages representing the conditions one would expect to find in a “typical” setting within a given occupation (Harvey & Wilson, 2000). The present study was, therefore, not able to assess the moderating effects of variability in situational strength that exists within the same occupation but in different organizations (see O’Reilly, Chatman, & Caldwell, 1991), nor variability that exists within the same organization at alternative levels of analysis (e.g., work group). This is important because situational strength likely varies across organizations, across positions within the same organization, and potentially even across time and circumstances within the same position (e.g., constraints may be reduced as tenure increases, consequences may be greater for those decisions or actions that directly affect the organization’s bottom line). Additionally, research on “tightness–looseness” (e.g., Gelfand, Nishii, & Raver, 2006), a construct similar to situational strength, raises the intriguing possibility that situational strength varies across societies or cultures.

Thus, the multi-level nature of situational strength has a number of important implications for the future development and application of this construct. First, the question of relevant levels of analysis (and specific operationalizations of situational strength therein) should be a focal point of any continued theoretical development of situational strength. Second, any test of situational strength
should be explicit about what level(s) of analysis is/are being assessed. Lastly, researchers and theorists should be open to the possibility that the effects of any single operationalization may not be consistent across levels. For example, it is possible that the effects of “constraints” are more pronounced at micro levels of analysis (e.g., constraints imposed by one’s immediate supervisor) than at macro levels of analysis (e.g., constraints imposed by governmental regulations). By measuring situational strength at the occupational level of analysis, the present study provides an initial test of the impact of this construct but also emphasizes the importance of continued investigations at other levels of analysis.

Conclusions

The present study meta-analytically demonstrates that occupation-level situational strength (operationalized here as constraints, consequences, and a formative global situational strength composite) moderates the conscientiousness–performance relationship, such that conscientiousness better predicts performance in characteristically weak occupations than in characteristically strong occupations. These findings demonstrate that situational strength can be used to improve the theory, research, and practice surrounding personnel selection by helping to calculate more accurate estimates of the expected criterion-related validity of non-cognitive predictors. Thus, we caution against relying exclusively on corrected, bivariate, mean effect size estimates without additionally accounting for variability induced by situational strength. We also encourage future theoretical and empirical investigations into the nature of situational strength as well as its influences on other personality–outcome relationships, at additional levels of analysis. It is only when such inquiry becomes commonplace that we will truly achieve the ideals outlined by Lee Cronbach more than 50 years ago.

Acknowledgements

The authors thank Michael Campion, Jessica Foster, Laurent Lapierre, James LeBreton, and Charlie Reeve for their constructive feedback, as well as Fabian Elizondo for his help with data retrieval from Hogan Assessment Systems, Inc., and Rod McCloy at HumRRO for his O’NET expertise.

Author biographies

Rustin Meyer is a PhD candidate in Industrial/Organizational Psychology at Purdue University and will start as an Assistant Professor of I/O Psychology at the Georgia Institute of Technology in the fall of 2009. His research interests focus primarily on better understanding person-situation interactions as they affect workplace behaviors.

Reeshad Dalal is an Assistant Professor of Industrial/Organizational Psychology at George Mason University. His research interests include person-situation interactions, deviant/counterproductive work behavior, organizational citizenship behavior, advice-giving and advice-taking, and job attitudes.

Silvia Bonaccio is an Assistant Professor of Human Resource Management and Organizational Behaviour Management at the University of Ottawa’s Telfer School of Management. Her research
interests include advice-giving and advice-taking, decision making, and the study of individual differences as they relate to workplace behaviors.

References

Note: References marked with an asterisk indicate studies included in the meta-analysis.


## Appendix. List of O’NET items used to construct situational strength composites

<table>
<thead>
<tr>
<th>Item name</th>
<th>Item definition</th>
<th>Item anchors</th>
<th>O’NET source</th>
<th>Relevant dimension</th>
<th>ITC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making decisions and solving problems (reverse scored)</td>
<td>Analyzing information and evaluating results to choose the best solution and solve problems</td>
<td>H = Reaching conclusions after considering a larger number of choices that are often ambiguous or abstract, where there are competing viewpoints and alternatives that must be considered before reaching final decisions and the solutions decided upon will have a very significant impact. L = Reaching conclusions after considering a few choices that are usually well defined, and the decisions or solutions will have minor impact.</td>
<td>Generalized work activities</td>
<td>Constraints</td>
<td>.49</td>
</tr>
<tr>
<td>Thinking creatively (reverse scored)</td>
<td>Developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions</td>
<td>H = Creating or inventing new and yet-to-be proven practices, technologies, materials, products, or strategies, where the creative effort will have widespread impact and will result in substantial improvements for both an organization and its customers. L = Offering suggestions for some change or improvement to immediate work function or procedures.</td>
<td>Generalized work activities</td>
<td>Constraints</td>
<td>.56</td>
</tr>
<tr>
<td>Freedom to make decisions (reverse scored)</td>
<td>How much decision-making freedom, without supervision, does the job offer?</td>
<td>H = Extensive freedom L = Very little freedom H = Very unstructured</td>
<td>Work context</td>
<td>Constraints</td>
<td>.57</td>
</tr>
<tr>
<td>Structured versus unstructured work (reverse scored)</td>
<td>To what extent is this job structured for the worker, rather than allowing the worker to determine tasks, priorities, and goals?</td>
<td>L = Very structured</td>
<td>Work context</td>
<td>Constraints</td>
<td>.59</td>
</tr>
</tbody>
</table>
| Importance of repeating same tasks | How important is repeating the same physical activities (e.g., key entry) or mental activities (e.g., checking entries in a ledger) over and over, without stopping, to performing this job? | H = Extremely important  
L = Does not apply  
H = Extremely important | Work Context  
Constraints  
.34 |
| Pace Determined by Speed of Equipment | How important is it to this job that the pace is determined by the speed of equipment or machinery? (This does not refer to keeping busy at all times on this job.) | L = Does not apply  
H = Hourly or more often | Work Context  
Constraints  
.53 |
| Time Pressure | How often does this job require the worker to meet strict deadlines? | L = Never  
H = Monitoring and controlling a large number of resources, including managing a large budget | Work Context  
Constraints  
.09 |
| Monitoring and controlling resources | Monitoring and controlling resources and overseeing the spending of money | L = Needing to do little monitoring or controlling of resources or money | Generalized  
Work Activities  
Consequences  
.30 |
| Responsible for Others’ Health and Safety | How much responsibility is there for the health and safety of others in this job? | H = Very substantial  
L = None, no responsibility | Work Context  
Consequences  
.48 |
| Responsibility for Outcomes and Results | How responsible is the worker for work outcomes and results of other workers? | H = Very substantial | Work Context  
Consequences  
.62 |
<table>
<thead>
<tr>
<th>Item name</th>
<th>Item definition</th>
<th>Item anchors*</th>
<th>O'NET source</th>
<th>Relevant dimension</th>
<th>ITC</th>
</tr>
</thead>
</table>
| Consequence of Error            | How serious would the result usually be if the worker made a mistake that was not readily correctable? | H = Extremely serious  
L = Mildly serious  
H = Extreme results | Work Context | Consequences       | .50 |
| Impact of Decisions on Coworkers| How do the decisions an employee makes impact the results of co-workers, clients or the company? | L = Very minor results  
H = Hourly or more often | Work Context | Consequences       | .63 |
| Frequency of Decision-Making    | How frequently is the worker required to make decisions that affect other people, the financial resources, and/or the image and reputation of the organization? | L = Never             | Work Context | Consequences       | .55 |
| Importance of Being Exact or Accurate | How important is being very exact or highly accurate in performing this job? | H = Extremely important  
L = Does not apply | Work Context | Consequences       | .22 |

Note: L = Low, H = high. ITC = item-total correlation.

*Anchors reflect those provided by Jeanneret, Borman, Kubisiak, and Hanson (1999) and Strong, Jeanneret, McPhail, Blakley, and D’Egidio (1999), and are written in their pre-reverse scored form.