

Singapore Management University Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of
Business

Lee Kong Chian School of Business

8-2018

Prosocial implicit trait policies underlie performance on different situational judgment tests with interpersonal content

Stephan J. MOTOWIDLO
Rice University

Filip LIEVENS
Singapore Management University, filiplievens@smu.edu.sg

Kamalika GHOSH
Rice University
DOI: <https://doi.org/10.1080/08959285.2018.1523909>

Follow this and additional works at: https://ink.library.smu.edu.sg/lkcsb_research

Part of the [Industrial and Organizational Psychology Commons](#), [Organizational Behavior and Theory Commons](#), and the [Personality and Social Contexts Commons](#)

Citation

MOTOWIDLO, Stephan J.; LIEVENS, Filip; and GHOSH, Kamalika. Prosocial implicit trait policies underlie performance on different situational judgment tests with interpersonal content. (2018). *Human Performance*. 31, (4), 238-254. Research Collection Lee Kong Chian School Of Business.

Available at: https://ink.library.smu.edu.sg/lkcsb_research/5990

This Journal Article is brought to you for free and open access by the Lee Kong Chian School of Business at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection Lee Kong Chian School Of Business by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

Prosocial implicit trait policies underlie performance on different situational judgment tests with interpersonal content

Stephan J. Motowidlo^a, Filip Lievens^b, and Kamalika Ghosh^a

^aRice University; ^bSingapore Management University

ABSTRACT

This study tests the hypothesis that situational judgment tests (SJTs) with interpersonal content reflect implicit beliefs about the utility of prosocial action for job effectiveness and that agreeable people are more likely to believe that prosocial action is effective. Two hundred ninety-four undergraduates completed four different SJTs with interpersonal content and a measure of Agreeableness. Results show that the effectiveness of response options in these SJTs is positively correlated with the level of prosociality they express and that because of their prosocial elements, scores on different SJTs are correlated with one another and with Agreeableness. These results shed light on the construct-related validity of SJTs with interpersonal content and point to the possibility that they can assess prosociality in job settings different from those described in their items.

Although situational judgment tests (SJTs) have become popular selection instruments over the last 20 years, with an impressive record of criterion-related validity (McDaniel, Hartman, Whetzel, & Grubb, 2007; McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001), little is known about *why* they are able to predict job performance as consistently as their record shows. The main focus in this literature has been on developing SJTs to maximize their ability to predict performance criteria with little attention to questions about what they measure and how whatever they measure fits into some nomological network. As a result, the construct-related validity of SJTs “is currently a ‘hot mess’ without much theoretical or empirical guidance” (McDaniel, List, & Kepes, 2016, p. 50).

Recent theoretical and empirical work on knowledge determinants of SJT performance and their relation to implicit trait policies (ITPs) begins to address these pressing issues. ITPs are implicit beliefs about the utility of expressing certain personality traits (Lievens, 2017; Lievens & De Soete, 2015; Lievens & Motowidlo, 2016; Motowidlo, 2003, 2017; Motowidlo & Beier, 2010; Motowidlo, Hooper, & Jackson, 2006a, 2006b; Motowidlo, Kell, Martin-Raugh, & Ghosh, 2016). This work leads to the idea that SJTs predict job performance partly because the procedural knowledge they measure includes information about ITPs. A key conceptual advantage of introducing the notion of ITP is that it provides a way to scale individual differences in SJT scores according to a construct that has rich theoretical connections with other constructs, including ability, personality, values, and interests. This permits predicting, testing, and establishing theory-driven linkages between SJT scores, ITPs, and these other constructs.

The rationale for our study proceeds from the contention that SJTs measure procedural knowledge (Clevenger, Pereira, Wiechmann, Schmitt, & Harvey, 2001; McDaniel & Nguyen, 2001; Motowidlo, Borman, & Schmit, 1997; Motowidlo, Hanson, & Crafts, 1997; Weekley & Jones, 1999). There are different kinds of procedural knowledge. Because many, if not most, SJTs in use today include items that ask about interpersonal situations (Christian, Edwards, & Bradley, 2010), procedural knowledge about effective interpersonal behavior is an especially common element of

these SJTs. In these interpersonally oriented SJTs, we hypothesize that response options that describe prosocial actions (actions that express agreeableness, benevolence, compassion, caring, kindness, generosity, courtesy, civility, cooperativeness, and tact) are more effective than response options that describe antisocial actions (actions that express disagreeableness, malevolence, coldness, indifference, cruelty, selfishness, rudeness, incivility, uncooperativeness, and insensitivity). If this hypothesis is correct, people who endorse more prosocial response options will obtain higher procedural knowledge scores. We also hypothesize that agreeable people are more likely to endorse prosocial response options in interpersonally oriented SJTs. This would explain why studies have shown that agreeable people obtain higher procedural knowledge scores (McDaniel et al., 2007; Whetzel & McDaniel, 2009). Studies have also shown that agreeable people are more likely to harbor ITPs that value the importance of prosocial action (Kell, Motowidlo, Martin, Stotts, & Moreno, 2014; Motowidlo et al., 2016) and that people who harbor such ITPs obtain higher procedural knowledge scores (Motowidlo & Beier, 2010). This leads to the inference that interpersonally oriented SJTs have common antecedents in agreeableness and prosocial ITP, and so we hypothesize that the procedural knowledge scores they produce are intercorrelated.

This study develops and tests the hypothesis that one underlying construct that binds SJTs with interpersonal content is prosocial ITP. Although many SJT scholars would probably acknowledge that SJTs often contain interpersonal content (see meta-analysis of Christian et al., 2010), “interpersonal skills” is then typically used as a catch-all term. This lack of precision runs counter to taxonomic work that differentiates among more specific interpersonal skill dimensions (i.e., Hogan & Lock, 1995; Klein, DeRouin, & Salas, 2006). Our study takes a more fine-grained approach because it is based on the idea that an important feature of this interpersonal content is that it contains prosocial and/or antisocial elements.¹ Thus, the novel element in our contribution is that we advance and define a specific construct that we hypothesize underlies interpersonally oriented SJTs. In this way, we hope to lend conceptual precision to the assumption that many SJTs involve some vague and undefined notion of “interpersonal skill.” In turn, this enhanced precision allows formulating better theory-driven connections with other well-defined constructs.

Knowledge determinants of SJT performance

SJTs ask people to judge the effectiveness of response options in SJT items. SJTs often ask for effectiveness judgments directly by instructing people to indicate which options are best and which are worst or by rating the level of effectiveness represented by each option or ask less directly by instructing respondents to pick options they would be most or least likely to carry out or to rate all response options for the likelihood that they would carry them out. Then SJTs are typically scored by comparing someone’s judgments about the effectiveness of SJT response options to judgments about their effectiveness made by subject matter experts. People get higher SJT scores when their judgments resemble experts’ judgments.

These scores represent their *procedural knowledge* about how to behave effectively in situations described in the SJT items (Clevenger et al., 2001; McDaniel & Nguyen, 2001; Motowidlo, Borman, & Schmit, 1997; Motowidlo, Hanson, & Crafts, 1997; Weekley & Jones, 1999). Motowidlo et al. (2006a) distinguished between two ways this knowledge could be learned. First, basic socialization processes, as represented, for instance, in how people are raised and educated, teach general

¹Note that we do not assume that all kinds of interpersonal interactions feature prosocial and/or antisocial elements. Although we do not know of any, we can imagine SJTs that ask, for instance, about interpersonal interactions such as sales representatives selling cars, demagogues firing up an electorate, platoon leaders ordering soldiers into combat, sociopaths cheating people out of their life savings, and so on. These are all interpersonal situations that require substantial interpersonal skill, but they are not the kinds of interpersonal skill that we believe underlie most SJTs currently in use. Instead, we propose that prosocial action and antisocial action are the important drivers of people’s responses to the kinds of interpersonally oriented SJTs in use today. We expect that people whose most likely or most effective SJT response choices are prosocial ones and whose least likely or least effective choices are antisocial ones obtain higher SJT scores.

principles about costs and benefits of expressing various traits. Second, as people gain experience in a particular job, they may learn more fine-grained information about when it is more or less effective to express certain traits at work.

Motowidlo and Beier (2010) drew upon the extensive literature on knowledge acquisition (e.g., Beier & Ackerman, 2005; Hambrick, 2003; Van Overschelde & Healy, 2001) to frame these knowledge determinants into a more detailed model. The knowledge acquisition literature distinguishes between prior knowledge about a general domain of information and newly acquired knowledge about specific aspects of that domain and emphasizes important differences between these two kinds of knowledge. Motowidlo and Beier (2010) likened prior knowledge about a general domain of information to the knowledge people learned about costs and benefits of expressing various traits, which they called *general domain knowledge*. They also likened newly acquired knowledge about specific aspects of the domain to the knowledge people learned from exposure to a particular job, which they called *job-specific knowledge*.

Two clarifications are in order, though. First, there exist of course different domains of knowledge, and they are not necessarily highly related to one another, and different domains can be associated with different trait constructs. The study described next focuses on knowledge domains that involve costs and benefits of expressing personality traits related to agreeableness and prosociality. But SJTs can be designed to tap knowledge domains associated with other traits as well, such as integrity (Husbands, Rodgeron, Dowell, & Patterson, 2015) and personal initiative (Bledow & Frese, 2009). If the association between an underlying trait and the procedural knowledge score produced by an SJT is strong enough, it may even be reasonable to claim that the SJT “measures” that trait even though the score that reflects it is a procedural knowledge or ITP score. So even though all SJTs measure procedural knowledge, it is still reasonable to view them as methods rather than constructs (Arthur & Villado, 2008) because they can measure different kinds of procedural knowledge with different implications for tapping into underlying trait constructs. Although a procedural knowledge or ITP score can sometimes be interpreted as a measure of an underlying trait (if the association between that trait and ITP or procedural knowledge is strong enough), in our model the underlying traits (agreeableness and prosociality) are treated as separate from, and antecedent to, the ITPs and procedural knowledge measured by the SJTs in our study.

Second, following Kell et al. (2014; and see Campbell, 1990), we make a sharp distinction between interpersonal *knowledge* and interpersonal *skill*. We claim that interpersonally oriented SJTs measure, in part, how much people *know* about how they should act in interpersonal situations. But knowing how they should act does not presume that they can actually *do* what they know should be done. Interpersonal skill can be measured by high-fidelity simulations (assessment center exercises) that have people actually perform in interpersonal situations but not by low fidelity simulations like SJTs.

General domain knowledge in the form of implicit trait policies

General domain knowledge about costs and benefits of expressing certain traits is denoted by the concept of ITP. Motowidlo et al. (2006a, 2006b) introduced this concept to explain how people evaluate the effectiveness of SJT response options. ITPs are implicit policies that describe how people weigh information about the traits expressed by an SJT response option when evaluating the effectiveness of that response option. Such domain knowledge in the form of ITPs can be measured by comparing a person’s judgments about response option effectiveness with estimates of the degree to which the response options express a particular trait. This contrasts to the typically measured procedural knowledge, which is captured by comparing a person’s judgments about SJT response option effectiveness with subject matter experts’ judgments of their effectiveness

Thus, when someone indicates their judgments about the effectiveness of response options in an SJT, those judgments can be used to compute scores for two different constructs. When they are compared to response option effectiveness determined by some other means (e.g., by judgments of

subject matter experts), that comparison yields a score for *procedural knowledge*. This is the score that is commonly computed for SJTs. When they are compared to the degree to which response options express a trait such as prosociality, that comparison yields a score for *prosocial ITP*. Based on theoretical work presented by Motowidlo et al. (2006a, 2006b), Motowidlo and Beier (2010), and Lievens and Motowidlo (2016), we assume that general domain knowledge represented by prosocial ITP is a component of the procedural knowledge measured by an SJT. In studies reported here, we compute scores for procedural knowledge based on the *effectiveness* of response options in three SJTs and scores for prosocial ITP based on the *prosociality* of the same response options.

The concept of dispositional fit (Motowidlo, 2003) explains how ITPs vary across individuals. This concept is predicated on the idea that people are inclined to believe that behavior that expresses their own traits is more effective in interpersonal situations like those described in SJT items. Thus, when such trait-expressive behavior really is more effective, people who possess those traits are more likely to know this. For example, according to this line of reasoning, agreeable people are more likely to believe that agreeable action is effective. If agreeable response options in a SJT are also the more effective options, agreeable people will endorse agreeable response options as more effective and obtain higher SJT scores.

Conceptualizing general domain knowledge as an ITP provides a key benefit for establishing construct-related validity because it provides a way to scale individual differences in general domain knowledge according to a construct that has ready implications for theoretical connections with other known constructs such as ability, personality, values, and interests (Lievens & Motowidlo, 2016; Motowidlo & Beier, 2010; Motowidlo, Ghosh, Mendoza, Buchanan, & Lerma, 2016). Depending on what trait is represented in a particular ITP, that ITP can be readily associated with particular personality traits, values, interests, and so on. For instance, ITP for prosociality is theoretically and empirically associated with individual differences in agreeableness, conscientiousness, benevolent values, emotional intelligence in the form of both understanding and management, social interests, and (negatively) with relativistic ethical ideology (Motowidlo et al., 2016; Motowidlo, Kell, Martin-Raugh, & Ghosh, 2016). As these traits are theoretically presumed to be antecedents of prosocial ITP, they lend themselves to tests of theory-driven connections between procedural knowledge measured by SJTs with interpersonal content, prosocial ITP, and these traits.

Although ITPs are thought to be shaped in part by traits such as these, they are also presumed to be shaped by experience and learning (Motowidlo et al., 2006a, 2006b). Motowidlo and Beier (2010) proposed that the kind of knowledge represented by ITPs was acquired through basic socialization processes, parenting, and formal education before people entered a job represented in SJT items, whereas job-specific knowledge was acquired after they had some experience in that job. The idea that ITPs are shaped by the joint operation of personality traits and experience evokes the distinction between basic tendencies and characteristic adaptations central to the five-factor model of personality. According to McRae and Costa (1996), basic tendencies are

capacities and dispositions that are generally inferred rather than observed. Basic tendencies may be inherited, imprinted by early experience, or modified by disease or psychological intervention, but at any given period in the individual's life, they define the individual's potential and direction. (p. 68)

Characteristic adaptations, on the other hand, “are acquired skills, habits, attitudes, and relationships that result from the interaction of individual and environment; they are the concrete manifestations of basic tendencies” (p. 69). Thus, prosocial ITP measured by comparing judgments about the effectiveness of SJT response options with response option prosociality fits into the notion of a characteristic adaptation.

This idea has been ventured before (Motowidlo & Beier, 2010; Motowidlo et al., 2006b). But here we also propose that as a characteristic adaptation, the prosocial ITP measured by any particular SJT is determined jointly by a basic tendency in the form of a *generic* ITP and learning experiences that cause people to modify their prosocial ITP for different occupations represented by different SJTs. The generic prosocial ITP we propose here is an inferred disposition that “may be inherited,

imprinted by early experience, or modified by disease or psychological intervention” (McRae & Costa, 1996, p. 68). Accordingly, people may have different prosocial ITPs for occupational contexts represented in different SJTs, but they are all partly determined by their generic ITP, which is relatively invariant across different occupations and SJTs.

Do ITPs generalize across different occupations and SJTs?

We expand Motowidlo and Beier’s (2010) suggestion that ITPs are more general than specific job knowledge by proposing that although they may be more general, prosocial ITPs may still be somewhat specific to the occupation referenced in the SJT from which they were derived. But because generic prosocial ITP affects prosocial ITPs for specific occupations and SJTs, we expect that these SJT-specific prosocial ITPs will share common variance attributable to the effects of generic prosocial ITP on all of them. This implies that people who believe that prosocial action, for instance, is important for effectiveness in one job are likely also to believe it is important for effectiveness in other jobs that share some similar features. For example, professional jobs usually involve interpersonal interaction with clients while providing some professional service. People who believe that prosocial action is effective in interpersonal settings are likely to believe this is true for all professional jobs that involve interpersonal interaction.

Martin (2011) tested the hypothesis that prosocial ITPs measured by SJTs generalize across professional domains. She administered three SJTs that measured prosocial ITP in different professional occupations to 152 undergraduates. One consisted of items about physicians’ actions, another about human factors engineers’ actions, and a third about community service volunteers’ actions. Correlations between the three SJTs ranged from .28 to .52 with a mean of .38, thereby showing that although the three SJTs were developed for very different occupational settings, the prosocial ITP scores they produce still share considerable variance.

Motowidlo et al. (2016) reasoned that if prosocial ITP generalizes across professional occupational settings, it should be possible to construct an SJT that would be context independent with respect to professional occupations by combining prosocial and antisocial items from several different professions. They developed an SJT that consisted of five prosocial and five antisocial items about each of four professions—physicians, lawyers, community service volunteers, and human factors engineers. They computed a separate ITP score based on 10 items for each professional job and conducted a confirmatory factor analysis that showed they all load on a single factor. Correlations between the four scores ranged from .53 to .72 with a mean of .63 in support of the hypothesis that prosocial ITP scores based on SJT items that describe actions in one professional setting correlate with prosocial ITP scores derived similarly for other professional settings. They also computed a total prosocial ITP score by summing across all 40 SJT items. The internal consistency reliability estimate for that total score was .90, and it was correlated as expected with agreeableness ($r = .22$) and both the understanding component ($r = .20$) and the management component ($r = .42$) of emotional intelligence. It was also correlated with conscientiousness ($r = .22$) and with prosocial performance in role-play simulations that involve others needing help ($r = .24$). These results support the reliability and construct-related validity of this measure of prosocial ITP based on items describing actions drawn from four professional jobs and offer further support for the idea that general domain knowledge in the form of prosocial ITP generalizes across occupations, at least across professional occupations that involve interpersonal interaction.

Although these studies found that prosocial ITP generalizes to some extent across occupational contexts, it is important to note that they used SJTs that were developed to be deliberately saturated with prosocial and antisocial items. This is a departure from conventional ways of developing SJTs that do not target any particular underlying trait (e.g., Motowidlo, Dunnette, & Carter, 1990). But if our reasoning holds, even if an SJT was not designed specifically to measure prosocial ITP, if it has interpersonal content, the procedural knowledge scores it produces are nevertheless laced with

prosocial ITP, and so they should also generalize across occupational contexts. At the backdrop of the preceding, the study reported here explores that possibility via several hypotheses that we propose in the next section.

Hypotheses

The overall premise driving this study is that interpersonally oriented SJTs share a common source of variance in prosocial ITP, which partially explains their relations with agreeableness and with one another. Our model of relations between agreeableness, prosocial ITP, and procedural knowledge appears graphically in [Figure 1](#). It shows that procedural knowledge is directly caused by ITPs (Lievens & Motowidlo, 2016; Motowidlo & Beier, 2010; Motowidlo et al., 2006a). In particular, as the procedural knowledge involves elements of interpersonal interaction, the ITPs are prosocial ITPs measured by the same SJT items that produced procedural knowledge scores. So the most immediate cause of SJT-specific procedural knowledge is SJT-specific prosocial ITP. The model also shows that each SJT-specific prosocial ITP is caused by generic prosocial ITP. This leads to the prediction that SJT-specific procedural knowledge scores will also be correlated with one another. Because generic prosocial ITP affects procedural knowledge only through its effects on SJT-specific prosocial ITP, partialling the relevant SJT-specific prosocial ITP scores out of SJT-specific procedural knowledge scores should remove all effects of generic prosocial ITP on procedural knowledge. Finally, the model shows that agreeableness has indirect effects on procedural knowledge through generic and SJT-specific prosocial ITP. This model leads to the following hypotheses that we test in this study.

Hypothesis 1

The first hypothesis is that SJTs with interpersonal content contain prosocial elements that partially account for the effectiveness of response options in interpersonal SJT items. This hypothesis echoes the assumption mentioned earlier that prosocial action in interpersonal situations is more effective

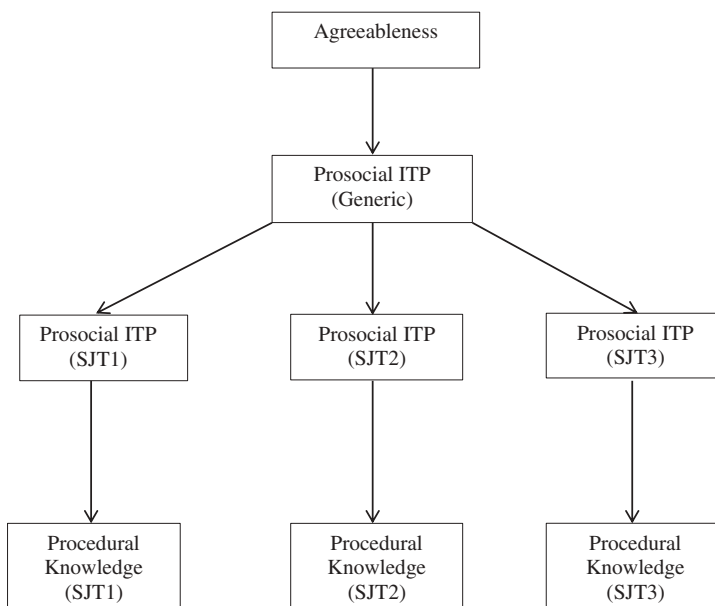


Figure 1. Model of relations between Agreeableness, generic prosocial ITP, SJT-specific prosocial ITP, and SJT-specific procedural knowledge. *Note.* SJT = situational judgment tests; ITP = implicit trait policy.

than antisocial action. If this is true, the most effective response options in a SJT with interpersonal content should express prosocial elements and the least effective response options should express antisocial elements. Thus, H1 states that interpersonal SJT scores are related to prosociality.

Hypothesis 2

McDaniel et al. (2007; Whetzel & McDaniel, 2009) reported mean meta-analytic correlations between agreeableness and SJT scores of .19 for SJTs with knowledge instructions (i.e., “identify the best response option” or “rate the effectiveness of all response options”) and .37 for SJTs with behavioral tendency instructions (i.e., “Which response option would you most likely and least likely carry out” or “Rate all response options for the likelihood that you would carry them out”). We contribute to this research by scrutinizing why agreeableness is correlated with SJT scores. According to a review by Christian et al. (2010), many SJTs are interpersonally oriented. We go beyond this general statement by more specifically proposing that the reason that agreeableness is correlated with SJT scores (Whetzel & McDaniel, 2009) might be that because many SJTs have interpersonal content, the procedural knowledge scores they produce are tinged with prosocial ITP, which is correlated with agreeableness. This leads to H2: Agreeableness is related to interpersonal SJTs because these interpersonal SJTs are saturated with prosociality.

Hypotheses 3a and 3b

These hypotheses test the novel idea that scores from different SJTs generalize across occupations and work settings because they tap common elements of prosocial ITP. So we posit that interpersonal SJTs are related to one another because these interpersonal SJTs are saturated with SJT-specific prosociality (H3a) and/or generic prosociality (H3b).

Method

Participants

The students in this study—331 undergraduates—participated for credit applied to their psychology courses, but some cases had missing data. Complete data on all variables are available for 294 undergraduates, including 190 women and 104 men ranging in age from 18 to 27 ($M = 19.23$, $SD = 1.27$). They include 14 African Americans, 133 Asians, 109 Caucasians, 30 Hispanics, and 8 others. All analyses were done with this sample of participants for whom complete data are available.

Measures and procedures

Students took part in this study in group sessions. First, we administered the consent form and a form to collect demographic information. Then we administered several questionnaires in the order shown next.

Management SJT

This is the 30-item SJT reported by Motowidlo et al. (1990), which was designed to predict interpersonal performance in the areas of leadership, assertiveness, flexibility, sensitivity, and communication and predict problem-solving performance in the areas of organization, thoroughness, drive, and resourcefulness. Thus, much of this SJT is interpersonally oriented, even though it includes some content that may not be interpersonally oriented, so we expect it to contain prosocial elements related to the effectiveness of some of the response options.

Each SJT item includes five response options. In the original development of this SJT, 33 or 34 very experienced senior managers rated response options for the level of effectiveness they described. Their mean ratings were used as effectiveness scores for response options.

Participants in our study completed the management SJT by selecting one response option they would most likely carry out and one they would least likely carry out for each SJT item. We computed procedural knowledge scores (denoted as MGT-PK) by summing effectiveness scores for their most likely options and subtracting effectiveness scores for their least likely options. The internal consistency reliability estimate (alpha) for their procedural knowledge score is .75.

Two doctoral students and three undergraduate research assistants rated this SJT's response options for the level of prosociality they represent on a 7-point scale: 1 (*very antisocial*), 2 = (*somewhat antisocial*), 3 = (*slightly antisocial*), 4 = (*neither prosocial nor antisocial*), 5 = (*slightly prosocial*), 6 = (*somewhat prosocial*), and 7 = (*very prosocial*). We defined prosocial for them as expressions of characteristics such as agreeableness, benevolence, compassion, caring, kindness, generosity, courtesy, civility, cooperativeness, and tact. We defined antisocial for them as expressions of characteristics such as disagreeableness, malevolence, coldness, indifference, cruelty, selfishness, rudeness, incivility, uncooperativeness, and insensitivity. The interrater reliability estimate (two-way mixed-effects intraclass correlation) for the mean of five raters is .87. We used their mean rating as the prosociality score for each response option.

We computed a prosocial ITP score based on the management SJT (denoted as MGT-ITP) for each participant in this study by summing prosociality scores for their most likely response option choices and subtracting prosociality scores for their least likely choices. The internal consistency reliability estimate (alpha) for their prosocial ITP score is .72.

We regressed procedural knowledge scores on prosocial ITP scores and saved the unstandardized residuals. These residuals (denoted as MGT-RES) represent what remains in the management procedural knowledge scores after removing prosocial elements.

Student SJT

This SJT was developed by Bess (2001) for her master's thesis. It consists of 23 items with four response options and one item with five response options. Respondents pick one response option that they would most likely take for each SJT item. The items describe various situations that undergraduates might encounter on a university campus, situations intended to elicit either student task performance or contextual performance. Some of the contextual items describe interpersonal situations, so we expect this SJT contains prosocial elements.

Because we could not find a scoring key for the SJT in the master's thesis, we developed our own. We calculated the proportion of students in our sample who indicated that they would perform each response option. Then we designated every response option that was endorsed by at least 30% of the sample as a "good" response. We calculated a procedural knowledge score (denoted as STU-PK) by allocating 1 point if a respondent's preferred option was one of the "good" ones and no points if it was not. The internal consistency reliability estimate (alpha) for this procedural knowledge score is .44.

The same five students rated response options in the student SJT for prosociality on the same 7-point scale used for the management SJT. The interrater reliability estimate (two-way mixed effects intraclass correlation) for the mean of their ratings is .83. We used their mean rating as the prosociality score for each response option.

We calculated a prosocial ITP score (denoted as STU-ITP) by summing prosociality scores for options that respondents indicated they would most likely carry out. The internal consistency reliability estimate (alpha) of this prosocial ITP score is .55.

We regressed student procedural knowledge scores on student prosocial ITP scores and saved the unstandardized residuals. These residuals (denoted as STU-RES) represent what remains in student procedural knowledge scores after removing prosocial elements.

Team Role Test SJT

The Team Role Test (Mumford, Van Iddekinge, Morgeson, & Campion, 2008) was developed to measure knowledge about effective role performance in teams. It taps knowledge about 10 roles, but one was dropped, leaving nine situational scenarios each with 10 response options. The roles are contractor, creator, contributor, completer, critic, communicator, calibrator, coordinator, and consul. Because they all involve interpersonal interaction in team settings, we expect this SJT to be prosocially saturated. Five options for each role are intended to describe effective performance, and five are intended to describe ineffective performance. Respondents rate the effectiveness of each response alternative. We reversed ratings for the options intended to describe ineffective performance and then summed all ratings for one total procedural knowledge score (denoted as TRT-PK). The internal consistency reliability estimate (alpha) for the total procedural knowledge score based on effectiveness ratings of 90 SJT response options is .86.

The same five students rated response options in the Team Role Test SJT for prosociality on the same 7-point scale used for the management SJT and the student SJT. The interrater reliability estimate (two-way mixed-effects intraclass correlation) for the mean of their ratings is .87. We used their mean rating as the prosociality score for each response option.

To calculate a prosocial ITP score based on items in the team SJT, we identified 49 prosocial items with ratings of 5 or greater on the 7-point scale and 17 antisocial items with ratings of 3 or less. For each participant, we computed the mean effectiveness rating for the prosocial items and the mean effectiveness rating for the antisocial items. The internal consistency reliability estimate (alpha) for the mean of 49 prosocial items is .84, and the reliability estimate for the mean of 17 antisocial items is .73. The correlation between these two means is $-.18$ ($p < .01$). We subtracted the mean antisocial score from the mean prosocial score to derive a prosocial ITP score (denoted as TRT-ITP) for each participant based on team SJT items.

We regressed procedural knowledge scores on prosocial ITP scores and saved the unstandardized residuals. These residuals (denoted as TRT-RES) represent what remains in team SJT scores after removing prosocial elements.

Opinions about professional conduct

The Opinions about Professional Conduct (OPC; Motowidlo et al., 2016) is a single-response SJT deliberately saturated with prosocial/antisocial content to measure prosocial ITP more generally in the broad domain of professional activity. It consists of 40 items that describe prosocial or antisocial behavior of physicians, lawyers, community service volunteers, and engineers. Respondents rate the effectiveness of each SJT item, and a prosocial ITP score (denoted as OPC-ITP) is calculated by subtracting effectiveness ratings for antisocial items from effectiveness ratings for prosocial items. Its internal consistency reliability estimate (alpha) is .88.

Agreeableness

We administered Goldberg's (1992) 50-item International Personality Item Pool. Although the International Personality Item Pool measures all of the Big Five personality factors, only Agreeableness is relevant for our purposes here. The internal consistency reliability estimate for the 10-item Agreeableness scale is .84.

Results

Hypothesis 1

H1 states that interpersonal SJT scores are related to prosociality. To test H1, we examine whether for SJTs with interpersonal content, response option prosociality is positively correlated with response option effectiveness. Recall that each response option in our various SJTs has one score for effectiveness and another score for prosociality. The first hypothesis is about relations

between these two scores in the three sets of SJT items used in this study. It tests the assumption that the more effective response options to interpersonally oriented SJT items are also more prosocial.

To test H1, we computed the correlation between response option effectiveness and response option prosociality separately for each SJT. That correlation is .59 ($p < .01$) for the management SJT ($N = 150$ response options), .29 ($p < .01$) for the student SJT ($N = 97$ response options), and .75 ($p < .01$) for the team SJT ($N = 90$ response options). Thus, the team SJT is the most prosocially saturated and the student SJT is the least prosocially saturated, but all three SJTs contain prosocial content that is associated with their response options' effectiveness. These results support the first hypothesis with respect to the three SJTs used in this study.

Hypothesis 2

Now that H1 is supported, we can move to H2, which states that Agreeableness is related to interpersonal SJTs because these interpersonal SJTs are saturated with prosociality. To test H2, we examined whether (a) Agreeableness was positively correlated with procedural knowledge, (b) Agreeableness was positively correlated with prosocial ITP, and (c) correlations between Agreeableness and procedural knowledge were attributable to the prosocial content of procedural knowledge and so diminished when prosocial ITP was partialled out. As shown in [Table 1](#), Agreeableness correlates $r = .33$ with management procedural knowledge, $r = .21$ with student procedural knowledge, and $r = .20$ with team procedural knowledge. It also correlates $r = .37$ with management prosocial ITP, $r = .40$ with student prosocial ITP, and $r = .26$ with team prosocial ITP. Thus, these correlations support the correlation of Agreeableness with procedural knowledge and prosocial ITP.

If correlations between Agreeableness and procedural knowledge are attributable to prosocial elements of procedural knowledge, correlations between Agreeableness and residual scores that partial prosocial ITP out of procedural knowledge should be lower than correlations between Agreeableness and procedural knowledge. As shown in [Table 1](#), Agreeableness correlates $r = .09$ with the management residual score, $r = .03$ with the student residual score, and $r = -.11$ with the team residual score. Tests for differences between two dependent correlations with one variable in common (Lee & Preacher, 2013a, September) reveal that Agreeableness is correlated significantly less with all three residuals than with their counterpart procedural knowledge scores ($z = 5.29$, $p < .01$, for management correlations; $z = 6.64$, $p < .01$, for student correlations; $z = 4.69$, $p < .01$, for team correlations). In short, all of these results support H2.

Hypotheses 3a and 3b

The basis for the third set of hypotheses is that scores from different SJTs with interpersonal content are related to one another because these SJTs are saturated with prosociality. To test this set of hypotheses, we examined whether (a) procedural knowledge scores derived from such SJTs were correlated with one another and (b) correlations between procedural knowledge scores were attributable to their prosocial content and so diminished when prosocial ITP (measured either with SJT-specific ITP or with general ITP) was partialled out. [Table 1](#) shows that management and student procedural knowledge are correlated $r = .29$ with each other, management and team procedural knowledge are correlated $r = .41$ with each other, and team and student procedural knowledge are correlated $r = .29$ with each other. If these correlations between procedural knowledge measured by different SJTs are attributable to their prosocial content, correlations between residual scores that partial prosocial ITP out of procedural knowledge should be lower than correlations between procedural knowledge scores. [Table 1](#) shows that management and student residuals are correlated $r = .19$ with each other, management and team residuals are correlated $r = .07$ with each other, and team and student residuals are correlated $r = .00$ with each other. Tests for differences between two

Table 1. Means, standard deviations, and correlations between all variables

	1	2	3	4	5	6	7	8	9	10	11
1. MGT-PK	—										
2. STU-PK	.29**	—									
3. TRT-PK	.41**	.29**	—								
4.MGT-ITP	.73**	.18**	.40**	—							
5.STU-ITP	.37**	.47**	.22**	.37**	—						
6.TRT-ITP	.42**	.32**	.91**	.43**	.27**	—					
7.OPC-ITP	.32**	.15*	.46**	.33**	.22**	.46**	—				
8.MGT-RES	.68**	.24**	.18**	.00	.14*	.16**	.11	—			
9.STU-RES	.13*	.89**	.21**	.01	.00	.22**	.05	.19**	—		
10.TRT-RES	.05	-.04	.35**	.00	-.07	-.05	.10	.07	.00	—	
11. Agreeableness	.33**	.21**	.20**	.37**	.40**	.26**	.25**	.09	.03	.25**	—
<i>M</i>	28.95	15.74	36.97	39.26	99.21	7.52	249.64	.00	.00	.00	56.60
<i>SD</i>	12.87	2.70	2.59	15.36	5.15	.60	17.03	8.74	2.39	.89	8.26

Note. *N* = 294. MGT = management SJT; PK = procedural knowledge score; STU = student SJT; TRT = Team Role Test; ITP = prosocial ITP score; OPC = Opinions about Professional Conduct; RES = residual score.

**p* < .05. ** *p* < .01 (two-tailed).

Table 2. Tests of causal models: Fit results and parameter estimates of standardized solution

Model 1 (Complete Mediation Model, Figure 1) ^a					
	MGT-ITP	STU-IT	TRT-ITP	OPC-ITP	Agreeableness
MGT-PK	.73**	---	---	---	---
STU-PK	---	.47**	---	---	---
TRT-PK	---	---	.91**	---	---
MGT-ITP	---	---	---	.33**	---
STU-ITP	---	---	---	.22**	---
TRT-ITP	---	---	---	.46**	---
OPC-ITP	---	---	---	---	.25**
Model 2 (Partial Mediation Model) ^b					
	MGT-ITP	STU-ITP	TRT-ITP	OPC-ITP	Agreeableness
MGT-PK	.70**	---	---	.09*	---
STU-PK	---	.46**	---	.05	---
TRT-PK	---	---	.89**	.05	---
MGT-ITP	---	---	---	.33**	---
STU-ITP	---	---	---	.22**	---
TRT-ITP	---	---	---	.46**	---
OPC-ITP	---	---	---	---	.25**
Model 3 (Fully Recursive Model) ^c					
	MGT-ITP	STU-ITP	TRT-ITP	OPC-ITP	Agreeableness
MGT-PK	.68**	---	---	.08	.06
STU-PK	---	.45**	---	.05	.02
TRT-PK	---	---	.89**	.06*	-.05
MGT-ITP	---	---	---	.25**	.31**
STU-ITP	---	---	---	.13*	.37**
TRT-ITP	---	---	---	.42**	.15*
OPC-ITP	---	---	---	---	.25**

Note. Model 1 is shown in Figure 1. Model 2 includes all paths of Figure 1 and paths from prosocial ITP to the three procedural knowledge scores (MGT-PK, STU-PK, and TRT-PK). Model 3 includes all paths of Model 2 and paths from Agreeableness to the three procedural knowledge scores (MGT-PK, STU-PK, and TRT-PK) as well as paths from Agreeableness to the three ITP scores (STU-ITP, TRT-ITP, and OPC-ITP). MGT = management situational judgment tests; STU = student situational judgment tests; TRT = Team Role Test; OPC = Opinions about Professional Conduct; PK = procedural knowledge score; ITP = prosocial implicit trait policies score; RMSEA = root mean square error of approximation; CFI = comparative fit index.

^a $\chi^2(22) = 178.95^{**}$, RMSEA = .17, CFI = .87. ^b $\chi^2(19) = 169.93^{**}$, RMSEA = .18, CFI = .88. ^c $\chi^2(13) = 83.51^{**}$, RMSEA = .14, CFI = .94. * $p < .05$. ** $p < .01$ (two-tailed).

dependent correlations with no variable in common (Lee & Preacher, 2013b) reveal that correlations between residuals that partial prosocial ITP out of procedural knowledge are significantly lower than correlations between procedural knowledge scores ($z = 2.034$, $p < .05$, for management-student correlations; $z = 5.055$, $p < .01$, for management-team correlations; $z = 4.406$, $p < .01$, for student-team correlations). These results support H3a.

We tested H3b in a similar fashion as H3a. Yet, this time we did not use the SJT-specific prosocial ITP scores but general prosocial ITP as measured by an SJT (i.e., the OPC) that was deliberately designed to measure professional prosocial ITP. As shown in Table 1, professional prosocial ITP correlates .32 with management procedural knowledge, .15 with student procedural knowledge, and .46 with team procedural knowledge. If these correlations between procedural knowledge and professional prosocial ITP are attributable to the prosocial content of procedural knowledge, correlations between professional prosocial ITP and residual scores that partial out ITP scores should be lower than correlations between professional prosocial ITP and procedural knowledge scores. Table 1 shows that professional prosocial ITP correlates .11 with the management residual, .05 with the student residual, and .10 with the team residual. Tests for differences between two dependent correlations with one variable in common (Lee & Preacher, 2013a) reveal that correlations between professional prosocial ITP and residuals that partial prosocial ITP out of procedural knowledge are significantly lower than correlations between professional prosocial ITP and procedural knowledge scores ($z = 4.528$, $p < .01$, for management correlations; $z = 3.404$, $p < .01$, for student correlations; $z = 5.811$, $p < .01$, for team correlations). These results support the third part of H3b.

Tests of full model

As suggested by an anonymous reviewer, we also conducted various tests of our full model. Table 2 presents the results of these model tests. We started by fitting the model in Figure 1 (Model 1: complete mediation model). Results showed significant estimates for all hypothesized path coefficients. This lends support to our Figure 1. As shown in the upper panel of Table 2, this model fit poorly by commonly used goodness-of-fit criteria (e.g., root mean square error of approximation $\leq .06$, comparative fit index $\geq .95$). So, we tested a second model (Model 2: partial mediation model) that added to Figure 1 direct effects from generic/professional prosocial ITP (as measured by the OPC) to the three effectiveness (procedural knowledge) scores. Again, the hypothesized path coefficients (as per Figure 1) were significant (see middle panel of Table 2). Given that the direct effects from generic/professional prosocial ITP to effectiveness (procedural knowledge) scores were low (between .05 and .09; only one direct effect was significant), model fit improved only slightly. Hence, we ran a third model (Model 3: fully recursive model) that added to Model 2 paths from Agreeableness to the three effectiveness (procedural knowledge) scores, as well as paths from Agreeableness to the three ITP scores. As indicated by the bottom panel of Table 2, all hypothesized path coefficients (as per Figure 1) were again significant. In addition, three direct effects from Agreeableness to the ITP scores were also significant. Therefore, this model fitted much better than the partial mediation model, difference in $\chi^2(6) = 86.42, p < .001$.

In sum, these model tests show that Figure 1 was generally supported: There was complete mediation from Agreeableness to the procedural knowledge scores. Yet there were also direct effects from Agreeableness to the ITP scores that were not specified in our Figure 1. Adding these direct effects make sense in line of the theory just laid out. The fact that Model 3 still provided a rather poor fit might indicate some unmeasured sources of model misspecification, which future research might take into account.

Discussion

Main conclusions

We began this study with the premise that little is known about the construct-related validity of SJTs because most of the attention in this literature has been on psychometric properties of SJTs and their criterion-related validity for predicting job performance. Our study takes a different approach. It explores the possibility that SJTs with very different job content might still share common variance in ITPs related to common elements of their job content. According to a review by Christian et al. (2010), many SJTs with items drawn from different work domains are interpersonally oriented. Accordingly, we focused on interpersonal content as a common feature of many SJTs, and we provided more conceptual precision by advancing prosociality as a binding element. That is, we proposed a connection between response option effectiveness and prosociality because in the kinds of interpersonal situations depicted in most SJTs with interpersonal content, prosocial actions are often more effective.

This study's main conclusion is that prosocial implicit trait policies underlie performance on different SJTs with interpersonal content. In other words, procedural knowledge scores of SJTs with interpersonal content are correlated with one another and with Agreeableness because procedural knowledge measured by these SJTs includes knowledge (i.e., prosocial ITP) about the effectiveness of prosocial expressions in interpersonal situations and because agreeable people are more likely to have this knowledge. In particular, our results show that (a) effectiveness of response options in SJTs with interpersonal content is positively correlated with the prosociality of response options so that people who tend to endorse more prosocial response options get higher procedural knowledge scores, (b) agreeable people harbor more prosocial ITPs and so are more inclined to endorse prosocial response options and get higher procedural knowledge scores, and (c) people who endorse prosocial response options in one SJT with interpersonal content also tend to endorse prosocial

response options in other SJTs with interpersonal content, so that people who score higher in procedural knowledge on one SJT with interpersonal content also tend to score higher in procedural knowledge on other SJTs with interpersonal content.

Caveats

All our results are based on only four SJTs. Thus, although we frame our conclusions as though they apply to all SJTs with interpersonal content, we acknowledge that can never be proven. The best we can hope for is disconfirmation in the form of evidence that two other SJTs with interpersonal content do not show the pattern of correlations we report here between agreeableness, procedural knowledge, and prosocial ITP. Thus, we offer our conclusions as hypotheses awaiting disconfirmation.

It is important to note, however, that although our study included only four SJTs, they are very different from one another in many important respects. One (Motowidlo et al., 1990) is a multiple-response SJT with items describing common management and administrative situations, with response options designed to predict problem-solving and interpersonal performance in these situations, with instructions to pick one most likely and one least likely option for each situation, with scoring for procedural knowledge by subtracting the effectiveness of least likely options from the effectiveness of most likely options, and with scoring for prosocial ITP by subtracting prosociality ratings of least likely options from the prosociality of most likely options. Another (Bess, 2001) is a multiple-response SJT with items describing common situations that university students might encounter on campus, with response options designed to predict students' task and contextual performance, with instructions to pick one most likely response option for each situation, with scoring for procedural knowledge by allocating 1 point for response option choices that were endorsed by at least 30% of the sample, and with scoring for prosocial ITP by summing prosociality scores for most likely response options. A third (Mumford et al., 2008) is a multiple-response SJT with items describing effective and ineffective performance in teams, with response options designed to measure knowledge about effective role performance in teams, with instructions to rate the effectiveness of each response option, with scoring for procedural knowledge by subtracting effectiveness ratings for ineffective options from effectiveness ratings for effective options, and with scoring for prosocial ITP by subtracting the mean of effectiveness ratings for very antisocial response options from the mean of effectiveness ratings for very prosocial response options. The fourth (Motowidlo et al., 2016) is a single-response SJT with items describing prosocial and antisocial actions by doctors, lawyers, community service volunteers, and engineers, with instructions to rate the effectiveness of items, without any scoring for procedural knowledge, and with scoring for prosocial ITP by subtracting effectiveness ratings for antisocial items from effectiveness ratings for prosocial items. With these marked differences, it is less likely that correlations between scores derived from these SJTs can be explained simply as artifactual results of correlated method variance.

Variability in test motivation, however, is another factor that might artifactually explain correlations between test scores. If some students completed our instruments carelessly or indifferently because they were not particularly motivated to score well on these tests, this might artifactually increase correlations between test scores because the unmotivated students will tend to score lower on all tests than students who were conscientiously motivated to do well. However, this issue will likely plague any attempt to study relations between SJTs that are as different from one another as the ones used in this study because it will be difficult to find a sample and a research setting that will guarantee that test participants will feel highly motivated to perform well on all of the varied SJTs they are asked to complete.

Implications for SJT theory and research

This study contributed to ITP theory. In prior research, there was evidence for a link between ITPs and (a) self-reported ratings on similar traits and (b) behavioral expressions of these traits as observed in actual situations (Motowidlo et al., 2016). This study goes beyond these results by demonstrating the key conceptual role of ITPs for binding a diversity of SJTs. We call for future studies that use the same construct-driven approach to test for other ITPs that might unify SJTs with other common elements. For example, SJTs that focus on task-oriented behavior that is not interpersonal might share common variance in ITPs for facets of conscientiousness related to drive, organization, and detail orientation. SJTs that focus on ethical or unethical behavior that is not interpersonal might share common variance in ITPs for other facets of conscientiousness related to integrity and rule compliance. Essentially, we propose that SJTs with items drawn from different job and occupational domains might be organized according to the kinds of ITPs that drive the procedural knowledge they purport to measure. Such future research can be conducted by placing SJTs in a network with available self-report measures and by developing marker tests that tap into the procedural knowledge related to these domains (see the OPC in this study related to prosociality). Currently, such marker tests are scarce, and more of them need to be developed.

We also call for studies that test whether a SJT that is saturated with ITPs relevant for performance in a particular job—such as prosocial ITP for jobs that involve interpersonal interaction in helping or service contexts—can predict performance in jobs even if its items are drawn from different job contexts. This would lead to the possibility of developing more “universal” SJTs to measure ITPs related to job performance in a much broader range of job and occupational contexts than those explicitly represented in their items.

Implications for SJT practice

Recently, Harvey (2016) posited that

SJT development has typically been costly and nonstandardized, being based on the questionable assumptions that (a) a high degree of situational customization of scenarios and responses is mandatory and that (b) scoring must be done only in an item-level, situationally customized fashion based on the effectiveness of each behavioral response in each setting. If targeted trait SJTs that are composed of more generic content and scored using standardized methods can still predict job effectiveness, such SJTs could be used far more widely and cost effectively than current ones. (p. 64)

This study provides some encouraging evidence that this might indeed be possible, at least for interpersonal SJTs. So this study and other similar ones might lead to a shift in focus on more generic interpersonal SJTs and their underlying constructs. We agree with Harvey (2016) that this has various practical benefits in terms of item development, response option development, and scoring. In the end, such a shift in emphasis in the development of interpersonal SJTs highlights that practical efficiencies and a deeper understanding of the constructs that interpersonal SJT measure are not contradictory but go hand in hand.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Arthur, W., & Villado, A. J. (2008). The importance of distinguishing between constructs and methods when comparing predictors in personnel selection research and practices. *Journal of Applied Psychology, 93*, 435–442. doi:10.1037/0021-9010.93.2.435

- Beier, M. E., & Ackerman, P. L. (2005). Age, ability, and the role of prior knowledge on the acquisition of new domain knowledge: Promising results in a real-world learning environment. *Psychology and Aging, 20*, 341–355. doi:10.1037/0882-7974.20.2.341
- Bess, T. (2001). *Exploring the dimensionality of situational judgment: Task and contextual knowledge*. Thesis submitted in partial fulfillment of the requirements for the degree, Master of Science in Psychology. Virginia Polytechnic Institute and State University, Blacksburg, Virginia.
- Bledow, R., & Frese, M. (2009). A situational judgment test of personal initiative and its relationship to performance. *Personnel Psychology, 62*, 229–258. doi:10.1111/peps.2009.62.issue-2
- Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 1, 2nd ed., pp. 687–732). Palo Alto, CA: Consulting Psychologists Press.
- Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology, 63*, 83–117. doi:10.1111/peps.2010.63.issue-1
- Clevenger, J., Pereira, G. M., Wiechmann, D., Schmitt, N., & Harvey, V. S. (2001). Incremental validity of situational judgment tests. *Journal of Applied Psychology, 86*, 410–417.
- Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment, 4*, 26–42. doi:10.1037/1040-3590.4.1.26
- Hambrick, D. Z. (2003). Why are some people more knowledgeable than others? A longitudinal study of knowledge acquisition. *Memory & Cognition, 31*, 902–917. doi:10.3758/BF03196444
- Harvey, R. J. (2016). Scoring SJTs for traits and situational effectiveness. *Industrial and Organizational Psychology, 9*, 63–71. doi:10.1017/iop.2015.119
- Hogan, J., & Lock, J. (1995, May). *A taxonomy of interpersonal skills for business interactions*. Paper presented at the 10th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Husbands, A., Rodgerson, M. J., Dowell, J., & Patterson, F. (2015). Evaluating the validity of an integrity-based situational judgment test for medical school admissions. *BMC Medical Education, 15*, 144. doi:10.1186/s12909-015-0424-0
- Kell, H. J., Motowidlo, S. J., Martin, M. P., Stotts, A. L., & Moreno, C. A. (2014). Testing for independent effects of prosocial knowledge and technical knowledge on skill and performance. *Human Performance, 27*, 311–327. doi:10.1080/08959285.2014.929692
- Klein, C., DeRouin, R. E., & Salas, E. (2006). Uncovering workplace interpersonal skills: A review, framework, and research agenda. In G. P. Hodgkinson & J. K. Ford (Eds.), *International review of industrial and organizational psychology* (Vol. 21, pp. 80–126). New York, NY: Wiley & Sons, Ltd.
- Lee, I. A., & Preacher, K. J. (2013a, September). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Retrieved from <http://quantpsy.org>.
- Lee, I. A., & Preacher, K. J. (2013b, October). Calculation for the test of the difference between two dependent correlations with no variable in common [Computer software]. Retrieved from <http://quantpsy.org>
- Lievens, F. (2017). Assessing personality-situation interplay in personnel selection: Toward more integration into personality research. *European Journal of Personality, 31*, 424–440. doi:10.1002/per.2111
- Lievens, F., & De Soete, B. (2015). Situational judgment test. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (Vol. 22, 2nd ed., pp. 13–19). Oxford, UK: Elsevier.
- Lievens, F., & Motowidlo, S. J. (2016). Situational judgment tests: From measures of situational judgment to measures of general domain knowledge. *Industrial and Organizational Psychology, 9*, 3–22. doi:10.1017/iop.2015.71
- Martin, M. P. (2011). *The generalizability of knowledge as measured by a single-response situational judgment test across domains*. Thesis submitted in partial fulfillment of the requirements for the degree, Master of Arts. Rice University, Houston, Texas.
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology, 60*, 63–91. doi:10.1111/peps.2007.60.issue-1
- McDaniel, M. A., List, S. K., & Kepes, S. (2016). The ‘Hot Mess’ of situational judgment test construct validity and other issues. *Industrial and Organizational Psychology, 9*, 47–51. doi:10.1017/iop.2015.115
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology, 86*, 730–740.
- McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment, 9*, 103–113. doi:10.1111/1468-2389.00167
- McRae, R. R., & Costa, P. T., Jr. (1996). Toward a new generation of personalities: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 51–87). New York, NY: Guilford.
- Motowidlo, S. J. (2003). Job performance. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology* (Vol. 12, pp. 39–53). New York, NY: Wiley.
- Motowidlo, S. J. (2017). Implicit trait policies in personality research. *European Journal of Personality, 31*, 472–473.

- Motowidlo, S. J., & Beier, M. E. (2010). Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. *Journal of Applied Psychology, 95*, 321–333. doi:[10.1037/a0017975](https://doi.org/10.1037/a0017975)
- Motowidlo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance, 10*, 71–83. doi:[10.1207/s15327043hup1002_1](https://doi.org/10.1207/s15327043hup1002_1)
- Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology, 75*, 640–647. doi:[10.1037/0021-9010.75.6.640](https://doi.org/10.1037/0021-9010.75.6.640)
- Motowidlo, S. J., Ghosh, K., Mendoza, A. M., Buchanan, A. E., & Lerma, M. N. (2016). A context-independent situational judgment test to measure prosocial implicit trait policy. *Human Performance*. doi:[10.1080/08959285.2016.1165227](https://doi.org/10.1080/08959285.2016.1165227)
- Motowidlo, S. J., Hanson, M. A., & Crafts, J. L. (1997). Low-fidelity simulations. In D. L. Whetzel & G. R. Wheaton (Eds.), *Applied measurement methods in industrial psychology*. Palo Alto, CA: Consulting Psychologists Press.
- Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006a). A theoretical basis for situational judgment tests. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 57–82). Mahwah, NJ: Lawrence Erlbaum.
- Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006b). Implicit policies about relations between personality traits and behavioral effectiveness in situational judgment items. *Journal of Applied Psychology, 91*, 749–761. doi:[10.1037/0021-9010.91.4.749](https://doi.org/10.1037/0021-9010.91.4.749)
- Motowidlo, S. J., Kell, H. J., Martin, M.P., & Ghosh, K. (2016). Implicit Trait Policies about Prosocial Professionalism. In P. Podsakoff, S. B. Mackenzie, & N. P. Podsakoff (Eds.), *The oxford handbook of organizational citizenship behavior* (pp. 543-558). New York, NY: Oxford University Press. doi:[10.1093/oxfordhb/9780190219000.013.14](https://doi.org/10.1093/oxfordhb/9780190219000.013.14)
- Mumford, T. V., Van Iddekinge, C. H., Morgeson, F. P., & Campion, M. A. (2008). The team role test: Development and validation of a team role knowledge situational judgment test. *Journal of Applied Psychology, 93*, 250–267. doi:[10.1037/0021-9010.93.2.250](https://doi.org/10.1037/0021-9010.93.2.250)
- Van Overschelde, J. P., & Healy, A. F. (2001). Learning of nondomain fact in high- and low-knowledge domains. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1160–1171.
- Weekley, J. A., & Jones, C. (1999). Further studies of situational tests. *Personnel Psychology, 52*, 679–700. doi:[10.1111/peps.1999.52.issue-3](https://doi.org/10.1111/peps.1999.52.issue-3)
- Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review, 19*, 188–202. doi:[10.1016/j.hrmr.2009.03.007](https://doi.org/10.1016/j.hrmr.2009.03.007)