Testing the General Domain Knowledge Component underlying Situational Judgement
Test Performance: A Training Study on Implicit Trait Policies

By

Ying Zhang

A Thesis Submitted to
Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfilment of the Requirements for
the Degree of Master's of Science in Applied Psychology.

December, 2018, Halifax, Nova Scotia

Approved: Dr. Victor Catano
Ph.D., Professor

Approved: Dr. Nicolas Roulin
Ph.D., Professor

Approved: Dr. James O'Brien
Ph.D., Professor

Approved: Dr. Wendy Darr
Ph.D., Defence Scientist

Date: December 19, 2018

Testing the General Domain Knowledge Component underlying Situational Judgement
Test Performance: A Training Study on Implicit Trait Policies

By Ying Zhang

Abstract

December 19, 2018

Situational Judgement Test (SJT) was viewed as a measure of context-dependent knowledge (Krumm et al., 2015). However, Motowidlo and colleagues (2006) proposed a new framework that introduced an additional knowledge component, ITPs (i.e., individuals' understanding of the effectiveness of behaviour based on the personality trait that this behaviour expresses). The current study tested the effect of ITP-based training on individuals' SJT performance. A hundred and thirty-one college students ($M_{age} = 27$, $SD = 7.70$) were tested in a controlled experiment. An independent-samples t-test revealed a small training effect ($M_D = 2.68$; $CI_{95} = 0.84, 4.52$; $d = .32$) thus supporting the existence of ITPs. Other factors, such as personality, cognitive ability, and sex also contributed to the variance of SIT performance. The current findings may have practical implications for inexperienced job seekers. However, such training's impact on trainees' actual job performance remained unknown and would require future investigations.

The Effect of Training on Situational Judgement Test through General Domain

Knowledge: A Training Study on Implicit Trait Policies

The Situational Judgement Test (SJT) is a type of assessment typically used to predict individuals' ability to perform at particular jobs. The original SJT was developed during the 1920s to select soldiers who possessed certain levels of procedural knowledge in responding to various situations (Catano, Brochu, & Lamerson, 2012). In this type of test, respondents are typically presented with various job-related situations; then, they need to make a judgement and decide on behavioural responses to the given situations (McDaniel & Nguyen, 2001). For example, respondents are asked to select the most- and least-appropriate behavioural responses to the situations from a few given options. Since these situations often involve social interactions at work places (e.g., how to effectively deal with interpersonal relations; Lievens & Motowidlo, 2016), many SJTs also have wide applicability in selecting employees for managerial positions (McDaniel, Morgeson, Finnegan, & Campion, 2001).

SJTs have been popular in both research and HR practice for years due to their predictive validity in relation to employees' job performance. This paradigm shares the same idea as other sample-based approaches to selection, such as work samples (Lievens & Motowidlo, 2016). Typically, SJTs are administered in a written format (there are also video-based SJTs, but such a format is used less frequently), and present work situations that employees are likely to encounter in actual jobs; candidates can respond to the situations without actually being present in these situations. Such a low-fidelity simulation saves the cost of work samples during personnel selection. Regardless of the presentation media, SJTs are usually developed with analyses of job-related tasks and with an input from subject matter experts (SMEs; Catano, et al., 2012; Motowidlo &

Beier, 2010). Thus, as other sample-based approaches, SJTs are rich in information that is taken from actual jobs. In turn, they tend to predict job performance (McDaniel, Hartman, Whetzel, & Grubb, 2007). A recent meta-analysis (Christian, Edwards, & Bradley, 2010) reported that the mean criterion-related validity of SJTs predicting job performance ranged from $\rho = 0.25$ to $\rho = 0.47$, depending on presentation formats (i.e., paper-and-pencil or video-based) and the job performance facets that were assessed (i.e., interpersonal skills, leadership skills, or heterogeneous composites).

Implicit Trait Policy

SJTs used to be conceptualized as a measure of context-dependent procedural knowledge (Krumm et al., 2015) since the original idea of this paradigm was to measure individuals' judgment in specific job situations. However, a reconceptualization of SJT (Motowidlo, Hooper, & Jackson, 2006; Motowidlo, Crook, Kell, & Naemi, 2009; Motowidlo, & Beier, 2010, Lievens & Motowidlo, 2016) proposed that such a paradigm might also measure context-independent knowledge. Motowidlo and colleagues (2006) referred to this as general domain knowledge. According to the most recent work on this SJT framework, such knowledge represents individuals' understanding of the degree to which a behavioural response is effective in a given domain (i.e., a wide range of general work situations, rather than in the context of a specific job; Lievens & Motowidlo, 2016).

More specifically, Motowidlo et al. (2006) introduced the concept of implicit trait policy (ITP) and suggested that the information of ITPs describes the general domain knowledge that is captured by SJTs. According to the authors, knowledge of ITPs refers to individuals' understanding or belief of the effectiveness of people's behavioural responses to situations in terms of the underlying personality traits that it expresses (i.e., based on the underlying personality traits that it expresses, one's behavioural response to

a certain situation could be determined as more or less effective). Such knowledge is important for individuals that take SJTs because they may select the most effective behaviours based on their belief of which type of personality traits that will be desired in the given situations. For example, in an SJT, when candidates are responding to a certain situation, there are a few behavioural responses listed as options. These behavioural responses, in fact, may express various levels of Agreeableness. Then, according to the concept of ITP, it is the candidates' knowledge or belief of the level of Agreeableness that is optimal in this given situation that guides their judgement.

Individuals' knowledge of ITPs captured in SJTs is context-independent because it refers to individuals' beliefs about the effectiveness of particular personality trait expressions in general (Motowidlo et al., 2006). According to the notion of ITP, such knowledge is acquired through the fundamental process of socialization. In other words, the general domain knowledge of ITPs is not solely obtained through work; rather, it could also be acquired through general life experience. Thus, individuals without certain work experience might also be able to perform on SJTs to a certain level due to the knowledge of ITPs that they have already acquired through other sources. Recent studies have demonstrated some evidence to support this idea. In Krumm et al. (2015), specific situational descriptions were removed from SJTs before participants' performances were measured. In their results, up to 71% of the SJT items revealed no statistical significance in differentiating between performances of the experimental (i.e., who answered the SJT with specific situation descriptions removed) and control groups. Thus, this finding supported the idea that SJTs might include not only the specific job knowledge but also the general domain knowledge that is relatively context independent.

Knowledge Components Underlying SJT

As Krumm et al. (2015) reported, there may be two knowledge components underlying SJTs – a specific job knowledge component and a general domain knowledge component. Candidates' performance on SJTs would depend on specific job knowledge when an SJT item requires knowledge and skills that can only be learned from particular jobs and when response options contain context-specific rules (i.e., behavioural rules specific to certain types of jobs or organizations). However, those who do not have experiences of doing these particular jobs would still be able to perform well on other SJT items, because they can infer the effective behaviour responses based on general life experience.
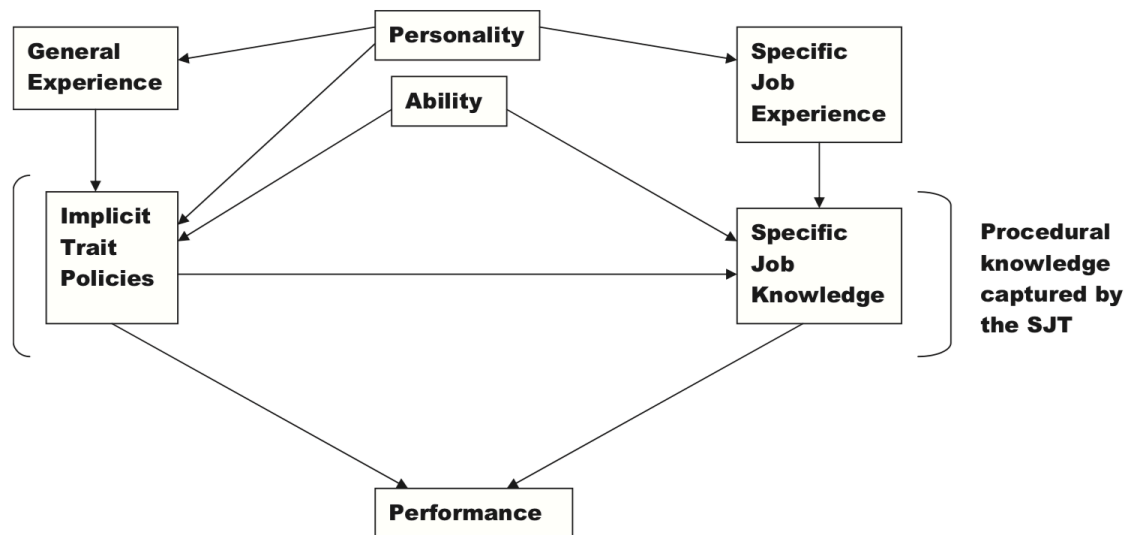

Figure 1. The model illustrating the underlying knowledge components of SJT

performance by Motowidlo and Beier (2010).

In fact, Motowidlo and Beier (2010) re-conceptualized the knowledge constructs underlying SJT scores. In this comprehensive framework, the knowledge components underlying SJT performance are illustrated in Figure 1. According to these authors, there are two constructs that are captured by SJTs. One is the specific job knowledge, and the

other is the knowledge of ITPs (i.e., the general domain knowledge). To support the

existence of these two knowledge components, these authors also presented some

empirical evidence. They compared candidates' scores on the same SJT but marked by

the different the scoring keys that were prepared by university students (i.e., who

possessed the general domain knowledge of ITPs) and SMEs (i.e., who possessed both

the general domain knowledge and specific job knowledge). The results revealed that

both of the scoring keys successfully predicted job performance; thus, Motowidlo and

Beier's (2010) concluded that the general domain knowledge (i.e., ITP) and specific job

knowledge were the two separate components captured by SJT.

Additionally, in Motowidlo and Beier's (2010), the scoring key produced by the

SMEs explained more of the criterion variance than the student key did. As these authors

explained, it was because SMEs had both the knowledge of ITPs and the specific job

knowledge, whereas students only had the knowledge of ITPs; having more knowledge

components underlying the SJT scores would certainly explain more of its variance. In

fact, this finding may also support the authors' further speculation based this SJT model –

these two knowledge components may be used to predict job performance independently.

Specifically, in SJTs, except for those more straightforward situations that can be

explained easily by trait expressions, there are also situations that do not follow the

general trait expression rules. For example, in some situations, the most effective

response might actually express the different ITPs than the general ones. As well, in other

circumstances, responses might simply be irrelevant to the information of trait

expressions. In these cases, the specific job knowledge will play a more important role

because the knowledge of ITPs might no longer be sufficient to produce effective

responses. Thus, Motowidlo and Beier (2010) also argued the incremental validity of the two knowledge components over each other when explaining SJT scores.

The literature of ability to identify criteria (ATIC) may also be relevant to the concept of ITP. According to its definition, ATIC refers to individuals' ability to perceive and identify the tested criteria when performing on an evaluation (Kleinmann et al., 2011). It is relevant to SJT performance because, in addition to having specific job knowledge, both identifying the underlying performance criteria in the SJTs and applying the knowledge of ITPs may help candidates to respond to the situations. However, it is important to distinguish the concept of ITP from that of ATIC. Essentially, ATIC is an ability, but ITP is knowledge. Whereas both of these constructs may provide advantages for candidates to perform on SJTs, they would influence candidates' performance differently. Specifically, the knowledge of ITPs is a resource that candidates can utilize when performing on SJTs, and, because it is knowledge, it can be built through training. In contrast, ATIC is an individual attribute that allows candidates to use any of their suitable resources to perform on SJTs, and, because it is an ability, it would be hard to improve through training.

However, ATIC was not investigated in the current study, because our main purpose (which would be explained in the following sections) had a primary focus on the potentially trainable knowledge component underlying SJTs (i.e., the knowledge of ITPs). Also, because the term "ability" in Motowidlo and Beier's (2010) SJT model implied individuals' general mental ability, we decided to follow their model and examine the overall influence of cognitive ability on SJT performance (i.e., more details would be presented in the following sections). Thus, the specific type of ability, ATIC, was not examined in the current study. Nonetheless, given the important role of ATIC in HR

selection practices and the relevance of its research, findings of the current study would be discussed in comparison with those from ATIC literature.

Training on ITP

The potential existence of the general domain knowledge component underlying SJT performance raises a question on the possibility of improving individuals' performance through ITP-based training. As previously introduced, the general domain knowledge of ITPs and the specific job knowledge were suggested as relatively independent predictors of SJT and job performance (Motowidlo & Beier, 2010). Whereas individuals' specific job knowledge has to be acquired from actual experience in relevant jobs, their knowledge of ITPs may be learned in other situations, including general life experience (Motowidlo, Hooper, & Jackson, 2006). This learning process is similar to the concept of social norms. Like social norms, the effective ways of trait expression might also be identified and communicated through language, including verbally. Such a conceptualization of ITP may indicate that individuals' knowledge of ITPs might be trainable through a relatively short session of verbal communication. Specifically, the effective ways of trait expression can be pointed out for trainees, along with encouragement to apply the ITP when taking SJTs. The trainees who have effectively learned the ITP may then select the responses that are congruent with the most effective trait expressions rules and achieve better performance.

For instance, an SJT's corresponding trait expression rules can be firstly identified as the core training content. Then, the trainees can learn about these rules as well as how to use them in job-related situations. For example, if an SJT favours behaviours that implicate high degrees of Agreeableness, the trait expression rule of "high Agreeableness" can be pointed out to trainees, along with situational examples. During

such training, the trainees may also be told that the employment tests they will be doing to compete for certain types of jobs are likely to favor candidates who express high degrees of Agreeableness in behaviours therefore it might be beneficial to respond to the test items toward such a way of trait expression. Although the knowledge of ITPs, as a general domain knowledge, might have already been acquired through trainee's daily life experience, such a training session would help them to better link the optimal personality trait expressions with particular SJTs.

It is necessary to test the effect of ITP-based training for the three important reasons. First, the possibility of improving candidates' SJT performance through a training session makes it feasible to improve competitiveness of the relatively inexperienced individuals during recruitment and selection processes. As suggested by Motowidlo and Beier (2010), additional to the influence of the knowledge of ITPs on SJT performance, this type of knowledge might even have extended impact on individuals' future job performance because knowing the effective trait expressions may also guide them to behave effectively in real work situations.

Second, studying the effect of ITP-based training on SJT performance may provide HR professionals and employment decision-makers a clearer understanding of the components underlying candidates' SJT scores. Although Motowidlo and colleagues (2006) had proposed the model of SJT based on previous literature and a thorough rationale, direct evidence supporting the general domain knowledge component remains scarce. A test of such a component model through a training study could provide more evidence in support of this model. Solidifying the theoretical basis of SJT paradigm would improve its utility in practice, allowing practitioners to better understand about their selection tool when making decisions.

Third, the results revealed from such a training study may help to quantify the room for improvement that can be achieved through an ITP-based training. That is, the extent to which this type of training would improve candidates' SJT performance. With such insight, decision-makers may then have a clearer idea of how much difference might exist between the candidates who have trained on ITP and those who have not.

Would SJT scores be influenced by the effect of training? Although research directly answering this question is scarce, recent studies along these lines are suggestive. For example, a few studies examined the effect of coaching on situational interviews. The results revealed a statistically significant effect of coaching on interview performance, although the effect size was smaller than that of cognitive oriented tests (Maurer, Solamon, & Troxtel, 1998; Maurer, Solamon, Andrews, & Troxtel, 2001). With such a premise, Cullen, Sackett, and Lievens (2006) tried to investigate training on SJTs directly; they compared the SJT scores of participants who had received training to those who had not. In the sample of one hundred and eleven undergraduate students, the authors administered two SJTs. One of the SJTs was the College Student Questionnaire (CSQ), which assessed the tacit knowledge of the college students. The other SJT was the Situational Judgement Inventory (SJI), which assessed students' knowledge of how to achieve good college performance. According to the authors, the SJI is more situation-specific that requires more complex processing of participants' knowledge and experience.

In their study, Cullen et al. (2006) found that their training intervention improved the participants' performance in the CSQ but not the SJI. The authors explained that the finding was due to the different complexity level of the two SJTs. According to the authors' interpretation, the CSQ was more straight forward and can be easily answered by

the participants who had been advised to answer in a responsible and conscientious manner. In contrast, the SJI required more complex (higher level) thinking and more nuanced understanding of how to respond to the situation-specific items. As a result, only the CSQ but not the SJI had been influenced by training.

Note that the CSQ is similar to the general domain knowledge component proposed by Motowidlo and colleagues' (2006). Meanwhile, the more-complex and situation-specific knowledge assessed in the SJI was similar to the specific job knowledge component in the same framework. Such a framework might help to explain Cullen et al.'s (2006) finding. Participants' performances on SJT were improvable through training, but such improvement might be limited to the component of more straight-forward (i.e., general domain) knowledge. However, this explanation requires further research because Cullen et al.'s (2006) study did not actually incorporate Motowidlo et al.'s (2006) new SJT model. Specifically, although the training for CSQ advised participants to respond in a responsible and conscientious manner, this training only focused on the trait expression rule of Conscientiousness but not the others. The other personality factors might also be important but were not included in Cullen et al.'s (2006) CSQ training. Without incorporating Motowidlo et al.'s (2006) new SJT model and mapping the training contents on all personality factors, we are unable to conclude whether it was the general domain knowledge of ITP or the single personality factor, Conscientiousness, that improved SJT performance. Thus, to better explain Cullen et al.'s (2006) finding and testify its connection with Motowidlo et al.'s (2006) SJT model, a specific ITP-based training study would be necessary to conduct.

As well, the current study aimed to test the general domain knowledge component underlying SJT performance through a controlled experiment. Previously, Motowidlo and

colleagues (2006) had established some evidence to support the existence of ITPs as the general domain knowledge captured by SJTs. However, they have not tested this component through direct training and testing. Thus, it might be necessary to look for more solid evidence that can support the existence of such knowledge component by examining whether or not SJT performance can be improved through training on the knowledge of ITPs.

Relevant to this concern, a more recent study by Lievens et al. (2012) found that individuals who had received coaching programs had better SJT performances in high-stakes selection. However, such a study was only a quasi-experiment, in which participants were self-selected to coaching programs. More importantly, the coaching intervention in this study was through commercial programs. Their contents were not designed specifically for improving ITP knowledge nor had integrated Motowidlo and Beier's (2010) SJT model; it only provided candidates tips and test-taking strategies without clearly identified scientific foundations (i.e., it was unclear what these test-taking tips and strategies were and how they were related to the current scientific knowledge of SJT). To better understand if and how training can influence SJT performance, relevant studies should examine such intervention that is developed based on scientific foundations. Thus, the current study decided to focus the training intervention on the knowledge of ITP, because it is one of the knowledge components in the most recent SJT theoretical framework that was proposed based on relevant scientific research. Thus, the primary purpose of this research was to examine the effect of ITP-based training on SJT performance.

Hypothesis 1: Individuals who have been trained on the knowledge of ITPs will perform better in an SJT than those who have not.

The current study examined this hypothesis by training and testing participants based on Becker's SJT of employee integrity (2005). This particular SJT was chosen for a few reasons. First, this SJT has moderate difficulty so the respondents would tend to have a room for improvement. Second, this SJT is designed to be generic; there is no jargon that is associated with particular types of jobs or organizations. Individuals with various educational background can easily understand the situations. Third, this SJT has relatively simple ITPs that many of the items do not involve complicated interactions of different personality traits. This allowed us to easily test the effect of training in isolation. The current study was a preliminary test of the ITP-based training effect; a simpler version of SJT and ITP training would allow us to clearly identify this effect without further complication.

As well, Becker's SJT (2005) assesses employee integrity. Such a construct was chosen in the current study because of its popularity in I/O psychology research and its potential impact on organizations. Integrity is one of the frequently studied variables in I/O psychology research (Lee, Ashton, & de Vries, 2005), and empirical evidence supports a robust relationship between integrity and counterproductive work behaviours (CWBs; Fine, Horowitz, Weigler, & Basis, 2010). Such behaviours are a broad set of deviant, immoral, and illegal behaviours in employees that could be detrimental to organizations (Sackett & Devore, 2001). For example, employees' theft and fraud behaviours were estimated to cost up to $400 billion in U.S. companies in total in one year (Wells, 1999).

Fine et al., (2010) found that high integrity persistently predicted a low occurrence of CWBs regardless of situational factors, such as security control norms and employee engagement. As well, even in an adverse situation, such as experiencing rejection during

career promotion, higher levels of integrity mitigated its negative influence; employees

with high integrity had fewer CWBs than those with low integrity (Fine, Goldenberg, &

Noam, 2016). Knowing that integrity is related to behaviours that could lead to

tremendous costs and damages in organizations, candidates' quality of integrity may be

an important element in many organizations' hiring considerations. Thus, focusing on

such an assessment tool may potentially help these organizations to make thorough

decisions and avoid negative consequences. Additionally, since CBWs were found to be

highly common (Harper, 1990; Murphy, 1993), our choice of studying an integrity

measure may have implications for a wide range of organizations across various

industries.

   To examine the *Hypothesis 1* using Becker's integrity SJT (2005)*,* the first step of

the current study was to identify the ITPs underlying the SJT. Then, a brief training based

on the identified ITPs were developed and delivered to the participants. Finally, the

participants' subsequent performance on the SJT were measured and compared with a

control group (i.e., the participants who would not receive ITP training). Based on our

hypothesis, we predicted that the participants who received the ITP training would have

higher SJT scores than the control group.

   Note that the current study focused on participants' performance improvement at a

test level, instead of an item level, due to the following considerations. First, such a focus

would avoid potential misinterpretation of the meaning underlying participants'

performance on the chosen SJT. Becker's SJT (2005) was to assess the construct of

integrity; however, there were more than one underlying factor revealed by the study's

factor analysis. Namely, positive relationships, career potential, leadership, and in-role

performance. A focus on any specific item would only reveal individuals' performance in

terms of a single factor, which could not fully represent the holistic construct of integrity. Second, because there are multiple underlying factors, it may introduce confounders if participants' performance is compared at an item level, because this process may also involve comparisons between the different underlying factors. Third, in practice, SJTs were usually developed to assess candidates' performance at a test level. In other words, decision makers are unlikely to interpret candidates' performance based on single items. In fact, single-item assessments are discouraged in psychological test development due to its low reliability (Reichers & Hudy, 1997). Therefore, to examine the effect of training, the current study focused on individuals' total SJT scores.

Individual Differences

Motowidlo and Beier's (2010) reconceptualization of the knowledge components underlying SJT performance includes ITPs and specific job knowledge. Additionally, as shown in Figure 1, the authors have also suggested some factors that may lead to individual differences in the acquisition of ITPs and specific job knowledge and eventually the SJT performance. One of the important factors that leads individuals' different SJT scores is personality.

First of all, we need to clarify that individuals' personality traits and their knowledge of ITPs are two distinct but interacting concepts. According to McCrae and Costa (1996), personality traits are individuals' natural and basic tendencies; however, through ongoing interactions with environments, individuals learn skills, preferences, attitudes, habits and so on that guide them to act more effectively in social situations (e.g., through mechanisms of social learning and reinforcement). The authors named this process characteristic adaptation. Motowidlo and Beier (2010) considered the acquisition of ITPs as an example of characteristic adaptation. Taken together, while personality

traits are individuals' natural tendencies that guide them to behave, this process is also under the influence of the knowledge of ITPs that they acquire through experience. In other words, individuals would respond to environments under the influence of their personal disposition, but it is the knowledge of ITP that helps them to adapt and act more appropriately in the future. Together, personality and the knowledge of ITPs help to shape individuals' behaviours in various situations.

The interaction between individuals' personality and knowledge of ITPs may also happen in another way – personality influences the acquisition of ITPs. Motowidlo's (2003) notion of disposition fit have explained this process. According to this notion, judgements of effective behaviours in a situation tends to reflect individuals' own personality traits. In other words, when individuals have the natural behavioural tendencies that are aligned with the optimal behaviours in a work situation, they are more likely to be motivated to acquire and utilize the knowledge of ITPs. For instance, they make more accurate judgements and would behave more properly in that situation by knowing and applying the appropriate rules. In contrast, when individuals' personality traits do not naturally guide them toward the effective responses in a situation, they may have more difficulties acquiring the knowledge of ITP and would not tend to behave accordingly. For example, in a workplace that encourages employees to be humble, kind, and empathic with clients, individuals who have lower Agreeableness may have a harder time embracing this trait expression rule and would not perform as well as others (i.e., who have higher Agreeableness).

Differences among individuals' ability to acquire and utilize relevant knowledge of ITPs might ultimately result in differences in their SJT performance. However, studies

examining the influence of personality on SJT performance through its interaction with ITP remain scarce. Therefore, the current study aimed to test such an influence.

Hypothesis 2a: Individuals who possess personality traits that are aligned with the corresponding ITPs will generally perform better in an SJT than those who do not.

Hypothesis 2b: Individuals who possess personality traits that are aligned with the corresponding ITPs will show more improvement in an SJT after being trained on the knowledge ITPs.

The current study tested these hypotheses with an assessment of participants' personality using the Big Five Personality factors (Goldberg, 1992). All participants' personality profiles (i.e., the combination of their Big Five personality factors) were compared to the ideal personality profile (i.e. the ITPs) favored by Becker's integrity SJT (2005). Before conducting the experiment, the current study identified the ITPs underlying Becker's SJT (2005; the specific procedure would be described in the following section), which was summarized as high Extraversion, moderate Agreeableness (i.e., somewhat Agreeable), high Conscientiousness, high Emotional Stability, and high Openness. Such a combination of Big Five personality traits was used as the reference point (i.e., ideal profile) in the measure of personality (i.e., in the current study, the variable of personality specifically referred to *the degree to which one's personality profile is aligned with the ITPs*). Specifically, after participants' Big Five personality factors were collected, their personality profile was compared with the ideal profile. A distance between each participant's personality profile and the ideal profile was calculated as a quantitative measure of personality (i.e., denoted as the *personality fit score*).

Testing the main effect of such a personality variable in predicting participants' SJT performance will provide evidence for *Hypothesis 2a*, whereas testing the interaction effect between participants' personality fit score and training condition in predicting SJT performance will allow us to test *Hypothesis 2b*. Based on our *Hypothesis* 2a, we predicted a negative effect of personality on SJT performance in all participants (i.e., smaller numerical values in participants' personality fit scores would predict higher SJT scores across both conditions). Also, based on Motowidlo and Beier's (2010) SJT model and our *Hypothesis 2b*, we predicted a negative interaction between personality and training condition when predicting SJT performance (i.e., smaller personality fit scores would predict even higher SJT scores in training condition).

Note that Becker's SJT (2005) measures the construct of integrity, and there was existing literature on the relationship between personality and integrity. For example, the HEXACO personality model (Ashton, Lee, & de Vries, 2014) incorporated Honesty-Humility as a personality factor, which is close to the construct of integrity by definition. However, since the current study chose to focus on the more widely known Big Five personality theory (Goldberg, 1992) as our basis of ITP knowledge, we focused on the literature that explained the relationship between the Big Five personality factors and integrity. The previous literature suggested that there were moderate positive associations between integrity and three personality factors, namely, Agreeableness, Conscientiousness, and Emotional Stability. Specifically, a meta-analysis by Ones et al., (1994) estimated the correlation between Conscientiousness and integrity to be 0.45, and other studies (Neuman & Baydoun, 1998; Sackett & Wanek, 1996) found correlations between Agreeableness and integrity to be .44 and between Emotional Stability and integrity to be .37. The rules of trait expression regarding these three personality factors

developed in the current study were consistent with these findings. However, while

Becker's SJT also favors the trait expression of high Extraversion and Openness, previous

studies had barely established association between integrity and these two personality

factors (Neuman & Baydoun, 1998; Sackett & Wanek, 1996). Nonetheless, since the

current study had a main focus on testing the effect of ITP-based training for a particular

SJT rather than integrity in general and that these two characteristics existed in the SJT in

a consistent manner, we included them as two of the five rules of trait expression.

Individuals' personality fit scores included all these five personality factors.

Another important individual factor in Motowidlo and Beier's (2010) SJT model

is cognitive ability. Based on the empirical finding on the strong and robust relationship

between cognitive ability and job performance (Schmidt & Hunter, 1998), Motowidlo and

Beier (2010) suggested that cognitive ability would also increase individuals'

performance on SJTs because such an ability implies higher capacity for knowledge

acquisition and retention. In the current study, since having the knowledge of ITPs for

answering the SJT would require individuals' ability to learn, we expected that those who

with higher levels of cognitive ability would be prepared with this type of knowledge

better than others.

Specifically, we proposed that cognitive ability would play an important role in

both individuals' learning from everyday experience and the specific training in the

current study. First, in their everyday experience, a higher mental capacity would help

individuals to process information and receive the knowledge of ITPs more effectively

despite the noises in the learning environment. After realizing the ITPs, then, individuals

with higher cognitive ability would also have a higher capability to remember such

knowledge. Thus, before participating in the current study, these individuals would have

better prepared with the mental resource (i.e., the effective ITPs) that would allow them to perform well in the SJT. Additionally, during our training, when they were also provided with the highly organized information essential for the SJT, these individuals might also learn better because they would be more likely to understand and remember such information. In other words, even though the ITP-based training was designed to improve all individuals' SJT performance, those who with higher levels of cognitive ability were expected to master the training contents better than others. Such higher effectiveness of learning could be revealed by these individuals' greater increase of SJT scores after training. Thus, the current study aimed to examine the following hypotheses regarding the influence of cognitive ability.

Hypothesis 3a: Individuals who possess higher cognitive ability will overall perform better in an SJT than others.

Hypothesis 3b: Individuals who possess higher cognitive ability will show more improvement in an SJT after being trained on the knowledge of ITPs.

The current study tested these hypotheses by examining the relationship between participants' cognitive ability and SJT performance. Information indicating participants' cognitive ability were collected through a brief cognitive ability test. According to Motowidlo and Beier's (2010) SJT model, we predict that there will be a positive association between the cognitive ability test scores and SJT scores among all participants. To test *Hypothesis 3b*, the current study examined the interaction between the participants' cognitive ability test scores and training condition in the prediction of SJT performance. We expected a stronger association between participants' cognitive ability test scores and SJT performance in the training condition. In plain words, it means

that individuals with higher cognitive ability were expected to benefit more from the training intervention.

Previous literature on SJT also suggested that there might be a sex difference. Whetzel et al., (2008) conducted a meta-analysis on SJT performances among different subgroups. Their findings revealed that SJT scores were slightly higher in female than male respondents, especially when SJT were correlated with Conscientiousness and Agreeableness. This was explained by the advantage of generally higher Conscientiousness, Agreeableness, and Emotional Stability in women (McDaniel & Nguyen, 2001; McDaniel et al., 2007; Costa, Terracciano, & McCrae, 2001). Although men tended to have a slight advantage of emotional stability (i.e., Neuroticism), Whetzel et al., (2008) argued that women's advantage of Conscientiousness and Agreeableness outweighed the disadvantage of Emotional Stability in SJTs and therefore women performed better than men overall. This is relevant to the current study because the SJT we used (i.e., Becker's SJT of employee integrity; Becker, 2005) tend to favor high Conscientiousness, Agreeableness, and Emotional Stability. However, existing studies confirming whether or not there is such a sex difference in Becker's SJT of employee integrity remains scarce. Therefore, the current study investigated this specific question. Clarifying such a sex difference may have important implications for HR policy and practices, especially for those organizations who plan to assess job candidates' integrity under the SJT paradigm. For example, knowing that SJT might tend to favor female candidates would help employers to make decisions with less bias.

One might argue that the sex difference in SJT performance might be redundant if such a difference could be explained by personality difference between two men and women. However, previous studies proposed such an explanation only based on the

coincidence that women tend to have higher Conscientiousness, Agreeableness, and

Emotional Stability while performing better in SJTs that also favor such personality

trends. The explanation for the sex difference through personality has not been

established in terms causation. Thus, the current study had measured both sex and

personality effects on SJT performance. By examining whether or not sex would explain

the variance in SJT scores above and beyond the personality factor, the current study

would improve to our understanding of the relationship between sex and SJT

performance. Thus:

Hypothesis 4: Female participants will overall perform better in SJT than male

participants.

To test such a hypothesis, the current study examined if sex (i.e., being female)

would predict higher SJT scores. Additionally, although not being hypothesized as the

predictors of SJT performance, the current study also collected participants' other

relevant demographic information. Questions regarding age, ethnic group, field of study,

and length and type of work experience were asked. This was to obtain an accurate

description of the current sample, for example, to ensure that potential applications of the

current findings would be externally valid.

Further Validation of the Integrity Measures

The SJT used in the current study was Becker's SJT of employee integrity (2005).

Becker developed this SJT because he believed that measuring integrity using SJT

paradigm ensures validity in predicting job performance (McDaniel, Morgeson, Finnegan,

Campion, & Braverman, 2001) and can prevent faking (McDaniel & Nguyen, 2001). By

presenting business scenarios with moral conflicts to respondents, Becker (2005) argued

that these items would be more realistic and less transparent to respondents than more

traditional and direct measures of integrity. Although Becker had validated this scale in the original study, follow-up studies intended to examine its correlation with other integrity measures (i.e., the concurrent validity) are scarce. Although such a measure might have been adopted in HR practice, we did not know the degree to which its score would correlate with other integrity measures. Therefore, the current study addressed this gap by examining the correlation between Becker's SJT and a ten-item Trait Self-Descriptive (TSDI) Integrity Measure (Catano, O'Keefe, Francis, & Owens, 2018), which was recently developed and validated.

Hypothesis 5: Participants' scores on Becker's SJT will be positively correlated with TSDI scores.

The current study tested this hypothesis by examining the correlation between participants' scores on the two measures. Because the two measures focus on the same construct, we expected that the two scores would be positively correlated.

Method

Participants

Participants of the current study were recruited from three sources - Amazon Mechanical Turk (MTurk), social media, and Saint Mary's University's Experimental Sign Up System (SONA). Participants were compensated. However, due to the Ethics constrains and the operations in different participating platforms, participants coming from different sources received different types of compensation. Individuals participated through MTurk received monetary compensation, and those who participated through SONA received course credit for one of the psychology courses they were taking. The amounts of compensation were determined through the calculation on a basis of one US dollar or 0.25 course credit per 15 minutes of participation. Each participant was expected

to spend approximately 45 minutes on average to complete the experiment; thus, participants from MTurk received four dollars whereas those who from SONA received 0.75 course credit (i.e., it was the participants' choice on which of the courses they were taking that would receive the credit, but all the courses were on a 100-point grade scale). Participants from both of the two platforms were anonymized; their identifying information was held by a third party and was inaccessible to the researchers. However, since social media participants participated directly through our anonymous web link, individualized compensation was not feasible unless there was a link between their data and identifying information. Thus, all these participants were entered into a random draw, where they could win a $50 gift card.

Participants were eligible to participate in the current study if they were currently studying at or recently (i.e., within a year) graduated from a college or university program. This was because the current study targets on a population that has relatively less work experience, because they were expected to have general domain knowledge but comparably little job experience. Unlike SONA, participation through MTurk and social media was open to the public, thus individuals who were not students or recent graduates could potentially choose to participate. MTurk participants were screened through a third-party tool, TurkPrime, which made sure that only participants identified themselves as students or recent graduates were able to find the current study. However, through social media, it was hard to control who could find the current study. To avoid ineligible participants, we tried to post advertisement in social groups that were dominated by students, and the eligibility criteria was emphasized within the advertisement. Based on the fact that the demographic compositions of social media participants' age and length of work experience were similar to MTurk, we would conclude that our participants from

social media could represent a student sample. Further, although participants from all

ethnic groups are encouraged to participate, they were required to be fluent in English.

This was because all instructions, training materials, and the SJT used in the current study

were in English; low English proficiency might introduce confounders. Also, all

participants were required to have normal or corrected vision, which was necessary for

them to understand and perform appropriately during our experimental procedure (i.e., all

materials were presented in a written from; the entire experimental procedure required

reading).

      Before collecting data, the number of participants was determined through power

analysis using G*Power 3.1 (2014) based on an estimated effect size, $R^2 = 0.2$, for the

current study. This effect size was estimated based on the previous research on the effect

of training, personality, sex, and cognitive ability on SJT performance (Cullen, et al.,

2006; Lievens & Sackett, 2006; Mumford, 2015). In total, 133 individuals participated in

the current study. However, two of them were excluded due to low validity of their data

(i.e., the specific exclusion procedure would be explained in Results). The demographic

composition of the 131 participants were summarized in Table 1. These participants had

a mean age of 27 ($SD = 7.70$). There was an even distribution of sex; around half of the

participants were females (n = 70), around half were males (n = 60), and one participant

indicated "other." With random assignment, the numbers of participants assigned into the

two conditions were also even; there were 66 participants in the control condition and 65

participants in the training (i.e., experimental) condition. More demographic

characteristics of participants, broken down by sample source, is shown in Table 1.

Table 1.

Summary of demographic information by participant source.

Demographic Characteristic	Mturk n (%)	Social Media n (%)	SONA n (%)	All Sources n (%)
Age				
18-29	68 (68)	17 (81)	9 (90)	94 (71.8)
30-39	25 (25)	4 (19)	1 (10)	30 (22.9)
40-49	5 (5)	0 (0)	0 (0)	5 (3.8)
50-59	0 (0)	0 (0)	0 (0)	0 (0)
60+	2 (2)	0 (0)	0 (0)	2 (1.5)
Sex				
Male	48 (48)	9 (43)	3 (30)	60 (45.8)
Female	51 (51)	12 (57)	7 (70)	70 (53.4)
Other	1 (1)	0 (0)	0 (0)	1 (0.8)
Ethnicity				
Caucasian	70 (70)	3 (14.3)	7 (70)	80 (61.1)
Black	16 (16)	0 (0)	0 (0)	16 (12.2)
Middle Eastern	0 (0)	0 (0)	2 (20)	2 (1.5)
Hispanic	7 (0)	0 (0)	0 (0)	7 (5.3)
Asian	6 (6)	17 (81)	1 (10)	24 (18.3)
Aboriginal	1 (1)	0 (0)	0 (0)	1 (0.8)
Other	0 (0)	1 (4.7)	0 (0)	1 (0.8)
Field of Study				
Arts and Humanities	19 (19)	1 (4.8)	0 (0)	20 (15.3)
Social Science	14 (14)	3 (14.3)	7 (70)	24 (18.3)
Natural Science	17 (17)	1 (4.8)	0 (0)	18 (13.7)
Formal Science	11 (11)	5 (23.8)	0 (0)	16 (12.2)
Professional and Applied Science	38 (38)	10 (47.6)	1 (10)	49 (37.4)

Other	1 (1)	1 (4.8)	2 (20)	4 (3.1)
Grade Point Average				
0-2.0	2 (2)	0 (0)	0 (0)	2 (1.5)
2.0-3.0	24 (24)	5 (23.8)	2 (20)	31 (23.7)
3.0-4.0	73 (73)	16 (76.2)	7 (70)	96 (73.3)
Prefer not to tell	1 (1)	0 (0)	1 (10)	2 (1.5)
Length of Work Experience				
0-12 months	23 (23)	7 (33.3)	1 (10)	31 (23.7)
1-3 years	30 (30)	7 (33.3)	4 (40)	41 (31.3)
3-5 years	13 (13)	3 (14.3)	3 (30)	19 (14.5)
5 years +	32 (32)	4 (19)	2 (20)	38 (29)
Prefer not to tell	2 (2)	0 (0)	0 (0)	2 (1.5)
Type of Work Experience				
N/A (No Work Experience)	19 (19)	6 (28.6)	1 (10)	26 (19.8)
Manager	10 (10)	1 (4.8)	0 (0)	11 (8.4)
Professional	17 (17)	4 (19)	0 (0)	21 (16)
Technician or Associate Professional	17 (17)	4 (19)	0 (0)	21 (16)
Clerical Support Worker	9 (9)	0 (0)	1 (10)	10 (7.6)
Service and Sales Worker	20 (20)	2 (9.5)	6 (60)	28 (21.4)
Skilled Agricultural, Forestry and Fishery Worker	0 (0)	0 (0)	0 (0)	0 (0)
Craft and Related Trade Worker	0 (0)	0 (0)	0 (0)	0 (0)
Plant and Machine Operator, or Assembler	4 (4)	0 (0)	0 (0)	4 (3.1)
Elementary Occupation	0 (0)	0 (0)	0 (0)	0 (0)
Armed Forces Occupation	2 (2)	1 (4.8)	0 (0)	3 (2.3)
Other	1 (1)	3 (14.3)	1 (10)	5 (3.8)

Prefer Not to Tell	1 (1)	0 (0)	1 (10)	2 (1.5)
ESL Student				
Yes	6 (6)	16 (76.2)	3 (30)	25 (19.1)
No	93 (93)	5 (23.8)	7 (70)	105 (80.2)
Prefer Not to Tell	1 (1)	0 (0)	0 (0)	1 (0.8)

Training Materials

ITP Training. All participants were randomly assigned to one of the two conditions: the training condition (i.e., go through the ITP training prior to SJT) and the control condition (i.e., go through a control task prior to SJT). Participants who were assigned to the training condition were presented reading material that was designed to deliver the knowledge of ITPs for the following SJT. The training material was developed carefully by the researcher to ensure that it only focuses on explaining the rules of personality trait expression that appeared in Becker's SJT (2005). During the entire development process, the definition of ITP (Motowidlo, Hooper, & Jackson, 2006) and the Big Five personality theory (Goldberg, 1992) were strictly followed. To identify the trait expression rules, each item of Becker's SJT (2005) was carefully reviewed, and the personality related information was extracted and organized in terms of Extraversion, Agreeableness, Conscientiousness, Emotional Stability (Neuroticism), and Openness (to experience) based on descriptions of each personality trait provided by Big Five personality theory (Goldberg, 1992). For example, if an item favors the behavioural response that expresses high creativity and openness to new solutions, the corresponding rule of trait expression was summarized as "high Openness". After reviewing all items of the SJT, the current study summarized that a person who performs well in Becker's SJT

would express high Extraversion, moderate Agreeableness (i.e., somewhat Agreeable), high Conscientiousness, high Emotional Stability, and high Openness.

The knowledge of ITPs was communicated with the participants through the written training material. In the ITP training page, participants were first introduced to the definitions of each personality factor (Ackerman, 2017; John, Robins, & Pervin, 2008; John & Srivastava, 1999). This was because such an understanding would be a basis for the participants to understand trait expression rules. Following the introductions to personality factors, the trait expression rules for Becker's SJT were presented to the participants. This was then followed by a behavioural example to facilitate learning. The contents of the ITP training are provided in Appendix A.

Control Task. To make sure that the only difference between the training and control condition was the knowledge of ITPs, participants of the control condition were given reading material about personality as well. However, it only introduced the five factors of personality. In other words, participants of both conditions were introduced to the same definition of personality factors, but participants in the control condition did not receive information on the specific trait expression rules identified from Becker's SJT (2005). However, to maintain the similarity of learning experience, some general behavioural examples corresponding to each personality factors were also presented to the participants of the control condition.

Note that while the behavioural examples presented in the training condition were specifically made to illustrate the desired ways of trait expression (i.e., a high-performing employee would express high Conscientiousness and how they would behave in a work situation), the behavioural examples presented in the control condition were only simple demonstrations of each personality factor (i.e., a typical behaviour of a high

Conscientiousness person and a typical behaviour of a low Conscientiousness person).

Such a material design was intended to isolate the independent variable (i.e., whether the

participants learn the rule of trait expression) and to maintain the other features the same

across conditions (i.e., the topic is focused on the Big Five personality theory, all

participants need to involve understanding and memorizing verbal information, etc.). The

contents for the control task are provided in Appendix B.

Before being used on any of the participants, both materials for the ITP training

and the control task were examined by graduate students and faculty members of the

Department of Psychology at Saint Mary's University. Based on the feedbacks from

graduate students and faculty members, modifications were made so that the materials for

each condition delivered the intended level of knowledge on ITP.

Measures

Situational Judgement Test. A revised version of Becker's SJT of employee

integrity (Cronbach's alpha > .90; 2005) was used to measure participants' performance

on an SJT after training on the general domain knowledge (i.e., the ITPs). As on other

SJTs, the answer choices in Becker's SJT had been weighted based on how effective the

behaviours would be in the given situation. The maximum score of this test is 18 and the

minimum is -18. In the original validation study (Becker, 2005), participants scored from

-3 to 17 ($M = 10$, $SD = 4$). In the current study, the original SJT items have been slightly

revised (See Appendix C) to ensure a proper context for student participants. However,

the revisions were minor and only aimed to adjust context; no change had been made to

modify the behavioural tendencies that would be assessed in the original SJT.

Manipulation Check. To test the training effect, we needed to ensure that the

training manipulation had achieved its desired results. Thus, a manipulation check was

necessary. All participants took a short multiple-choice question test. In this test, participants were asked questions about the desired personal traits in high-performing employees. The questions and answer keys were developed based on the rules of trait expression presented in the ITP training material (See Appendix D). To avoid potential practice effect or priming effect that might interfere participants' performance on the SJT, this test was placed after the SJT. Also, because all participants were taking this test regardless of their condition (i.e., whether or not they had learned the ITPs), instructions were provided so that all of them were prepared to answer the questions.

   Integrity. All participants took a ten-item Trait Self-Descriptive (TSDI) Integrity Measure. The ten items were developed and validated recently (Catano, et al., 2018). According to the authors, the ten-item integrity measure has established both good reliability (Cronbach's alpha = .76) and validity (i.e., construct validity, convergent validity, discriminant validity, etc.). This integrity measure was presented to participants as shown in Appendix E.

   Personality. All participants took a personality measure. The personality measure used in the current study were 50 items from the International Personality Item Pool (IPIP; International Personality Item Pool, 2017a). Reliability of the items measuring each of the Big Five personality factors was beyond .79 (see Table 2), with a mean reliability being Cronbach's alpha = .84 (Goldberg, 1992). IPIP is an empirically valid personality measure; items measuring each factor are highly correlated with other measures, such as NEO-PI-R Facet Scales (r = .73; Johnson, 2014). This personality measure was presented to participants as shown in Appendix F, and the distribution of items for each factor was shown in Table 2. The results of the participants' personality measure were scored in terms of the Big Five personality factors. In other words, each

participant had a personality profile that consisted of five sub-scores. Based on IPIP's

administration guide, participants' score of each Big Five factor were calculated by

summing all corresponding items for this factor for data analyses.

Table 2.

Reliability of the IPIP Personality Measure by Each Big-Five Domain

Big-Five Domain	Number of Items	Coefficient Alpha
Extraversion	10	.87
Agreeableness	10	.82
Conscientiousness	10	.79
Emotional Stability	10	.86
Openness	10	.84
Total/Mean	50	.84

Importantly, one of our hypotheses (*Hypothesis 2a*) predicted that individuals with

personality profile aligned with the ITPs would perform better on the SJT. The identified

ITPs for Becker's SJT were high Extraversion, Conscientiousness, Emotional Stability,

and Openness, and moderate Agreeableness. Thus, the ideal scores for Big Five

personality factors in the IPIP measure were determined as Extraversion = 5,

Agreeableness = 4, Conscientiousness = 5, Emotional Stability = 5, and Openness = 5.

These scores were determined according to IPIP's official guide, which suggested five

equally anchored points corresponding to the different levels of personality (International

Personality Item Pool, 2017b). For example, Extraversion = 1, 2, 3, 4, and 5 can be

interpreted as *introverted*, *somewhat introverted*, *average*, *somewhat extraverted*, and

extraverted. Thus, in the current study, the ideal scores of Extraversion,

Conscientiousness, Emotional Stability, and Openness were determined to be 5, and the

ideal score of Agreeableness corresponding to *moderate Agreeableness* was determined

to be 4, which, according to IPIP's interpretation guide, refers to *somewhat agreeable.*

Then, to measure the degree to which participants' personality profiles were

aligned with the ITP, a *personality fit* score was calculated for each participant through

summarizing the distance between his or her each personality dimensional score and the

ideal dimensional score corresponding to the ITP (i.e., $X_{\text{Personality fit}} = |(X_{\text{Extraversion}} - 5) +$

$(X_{\text{Agreeableness}} - 4) + (X_{\text{Conscientiousness}} - 5) + (X_{\text{Emotional Stability}} - 5) + (X_{\text{Openness}} - 5)|$). In other

words, such a score would directly reflect how far each individual is distant from the ideal

personality profile. For example, participant A with Extraversion = 3, Agreeableness = 3,

Conscientiousness = 2, Emotional Stability = 2, and Openness = 3 would have a

personality fit score of 11, and participant B with Extraversion = 4, Agreeableness = 4,

Conscientiousness = 4, Emotional Stability = 4, and Openness = 4 would have a

personality fit score of 4. In this case, participant A, who had a higher personality fit

score, would be more distant from the ideal personality profile and less aligned with the

ITP.

Cognitive Ability. All participants were measured on cognitive ability. The

current study used the 16-item International Cognitive Ability Resource (ICAR) test

(Condon & Revelle, 2014). This test has been established as a reliable and valid measure

of cognitive ability. Reliability coefficients of this test were beyond Cronbach's alpha of

0.81, and validity coefficients of this test ranged from 0.60 to 0.85. This cognitive ability

measure was presented to participants as shown in Appendix G.

Procedure

The entire experiment was conducted online through the Qualtrics platform. The

participants were provided with a web link to the experiment once after they had signed

up to participate. On the front page of the experiment, the participants were greeted and

informed that they could proceed to participate at a time of their convenience, but they

should prepare sufficient time to complete the experiment (i.e., at least 30 minutes and at

most 70 minutes, depending on their speed of reading and answering questions). All

participants were reminded that they need to focus on the materials and try to avoid

distraction for the entire experiment. They were also notified that all of their progress

would be timed. They can either choose "Yes, I'm ready now." to start the experiment or

choose "No, I will come back later." to defer the experiment. If they choose to defer the

experiment, they could access this page again and proceed to participate at any time.

If the participants click "Yes, I'm ready now.", they were directed into a new

page, where they would be informed about the purpose of the current study. All were told

that this study was aimed to help college students prepare for future job applications; the

materials in the current study might be beneficial to increase their competitiveness and

they would also have an opportunity to practice on some of the popular employment tests.

Confidentiality and anonymity of participation were reiterated. All participants had to

click "Agree" to consent participation. Then, the experiment started with a demographic

questionnaire. Information regarding to their age, sex, ethnic group, the field of study,

grade point average, length and type of work experience, and whether or not they are ESL

students were asked. Then, they were directed to complete the personality measure (i.e.,

IPIP) and integrity measure (i.e., TSDI). All these steps were grouped as a beginning

phase, in which the participants voluntarily started their participation and provided

information about themselves. This phase took maximally 20 minutes.

The next phase was the training phase. In this phase, half of the participants were

presented with the ITP training material, and half were presented with the control task

material. Participants in both conditions were told that the current study had provided

them with a material that aimed to improve their competitiveness during employment

selection tests. The potential benefits of this material were emphasized in the general

introduction to motivate learning. However, separate instructions for each condition were

given following the general introduction. The participants in the control condition were

told that they were going to study personality theories because some employers might be

interested in the relevant aspects. In contrast, the participants in the training condition

were told that they were going to study some test-taking strategies for one of the

commonly used employment tests. However, participants in both conditions were

reminded the importance of understanding and remembering the learning contents to

obtain the potential benefits.

 All participants were given the minimum of five minutes and maximum of 15

minutes to study; a count-down timer was shown on the top of the page. In other words,

the participants needed to spend at least 5 minutes to study the material. They might study

for longer; however, they would be directed to the next page after 15 minutes. Thus, the

total time for the training phase ranged from the minimum of five minutes and the

maximum of 15 minutes. The minimum and maximum time control was to ensure the

quality of data (i.e., participants could complete these materials with a reasonable amount

of time while distractions and memory decay remained relatively controlled).

 The next phase was a test phase, which included two main tests (i.e., the Becker's

SJT and the ICAR) and a brief manipulation check. The SJT was to assess the effect of

training, whereas the ICAR was to assess participants' cognitive ability. All participants

were given the same instruction, which introduced them to complete the following

employment tests. To simulate the psychological status of employment test preparation in

job seekers, in the instruction, all participants were also encouraged to apply what they have learned from the previous training. After the SJT, participants were prompted to complete the manipulation check, which was then followed by the ICAR. Similar to the previous phase, all of these tests had time restrictions and a displayed count-down timer. Similar to the training phase, participants could proceed after the minimal given time and had to proceed after the maximal given time.

The manipulation check was brief (i.e., it only takes one to three minutes). However, participants had 10 to 30 minutes to work on the SJT and five to 10 minutes to work on the ICAR. After completing the ICAR, all participants were debriefed and told the exact purpose of the current study. The ITP training materials were presented to the participants in the control condition if they were interested. All participants were reassured with anonymity and confidentiality and thanked for their time; compensation was then delivered. The entire experiment took approximately 45 minutes; however, some participants finished earlier (i.e., as little as 30 minutes) and a few participants took longer (i.e., up to 80 minutes) to complete the procedure.

All data collected were collected through Qualtrics^{TM} and then were exported to SPSS for statistical analysis. Independent-sample *t*-tests and Multiple linear regression analyses were performed to test hypotheses.

Results

Data scoring. Participants' responses to manipulation check, SJT, TSDI Integrity measure, and ICAR cognitive ability test were scored based on the corresponding marking keys. Mean scores of the IPIP personality measure were calculated in terms of Big Five personality factors, with all the reverse coded items been recoded according to IPIP's administration guide. In other words, each participant had a mean score for

Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness,

respectively. Then, based on the scoring method described in the previous section, each

participants' personality fit score was calculated. This score was a sum of participants'

IPIP sub-dimensional scores comparing to the reference point (i.e., the ideal sub-

dimensional scores in the ITP). Table 3 showed means and standard deviations of all

measured variables, including TSDI Integrity, ICAR cognitive ability, SJT, manipulation

check, and each of the Big Five personality factors of IPIP.

Table 3.

Means and Standard Deviations of Measured Variables among Different

Sample Sources.

	Mturk ($n = 100$)		Social Media ($n = 21$)		SONA ($n = 10$)		All Sources ($N = 131$)	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Extraversion	2.83	0.88	3.01	0.58	3.05	1.02	2.87	0.86
Agreeableness	3.83	0.72	4.01	0.54	4.08	0.66	3.88	0.70
Conscientiousness	3.71	0.79	3.63	0.58	3.66	0.69	3.70	0.75
Emotional Stability	3.20	0.72	2.91	0.49	2.71	0.56	3.12	0.70
Openness	3.84	0.66	3.81	0.43	3.39	0.42	3.80	0.62
TSDI Integrity	32.72	6.17	38.90	5.10	36.2	4.73	33.98	6.33
Situational Judgement Test	7.32	5.75	6.17	4.42	9.90	4.25	7.42	5.48
Cognitive Ability	5.96	3.14	5.95	3.55	4.00	2.79	5.81	3.20
Manipulation Check	3.64	1.22	3.67	1.39	3.60	1.17	3.64	1.24

Note. All personality measures had a scale of 0 to 5; TSDI integrity measure had a scale of 0 to 50; SJT had a scale of -18 to 18; ICAR cognitive ability had a scale of 0 to 16; manipulation had a scale of 0 to 5.

Data Screening and cleaning. The data file was screened and cleaned before

analysis. In terms of missing values, there were one case with missing information on age

and one with cognitive ability score missing. Because the missing values are under 5% of the data, it was assumed that the data were missing completely at random (MCAR), and these cases were used in the following analysis with pairwise deletion.

      Outliers and normality. Univariate outliers were determined if the given variable has a value exceeding $Z = 3.29$. There were two outliers. Case 44 had Agreeableness $Z = 3.85$, and case 63 had Openness $Z = 3.99$; both of them had exceeded 3.29 standard deviations from the mean. Because these outliers stayed on the continuum of the distribution, they were considered as extreme members of the same population. Thus, rather than deleting the outliers, the extreme values of the outliers were converted so that their corresponding Z scores were equal to 3.29. Note that the decision of winsorizing was made because the data did not need to be transformed, based on the result of normality checking. Normality of our data was determined through Skewness and Kurtosis. None of the variables had values comparable to the conventional problematic values (i.e., Skewness = 2, Kurtosis = 7), thus we concluded that the current sample had normal distribution in all measured variables. Mahalanobis distance was then calculated for each participant to detect multivariate outliers. We found two outliers with Mahalanobis distances equal to 34.43 and 32.14, both of which exceeded $\chi^2(9) = 27.88$. Because multivariate outliers have unusual combination of different variables, these cases may threaten validity of the following analyses. Thus, we performed listwise deletion in these two cases. There were in total 130 participants included in the following analysis. All the analyses used in the current study corresponding to our hypotheses were summarized in Table 4.

Table 4.

The corresponding statistical analysis methods in testing each hypothesis.

Hypothesis	Statistical Analysis	Parameter
H1	Independent sample t-test & multiple linear regression	Main effect of training
H2a	Multiple linear regression	Main effect of personality fit
H2b	Multiple linear regression	Interaction between personality fit and training
H3a	Multiple linear regression	Main effect of cognitive ability
H3b	Multiple linear regression	Interaction between cognitive ability and training
H4	Multiple linear regression	Main effect of sex
H5	Pearson's Correlation	N/A

Independent-samples *t*-test. To determine the effect of training, independent-samples *t*-tests were conducted to compare participants' performance on manipulation check and SJT between the experimental and control condition. In order to test the training effect, we needed to make sure that the manipulation of ITP training was effective; in other words, we expected that participants in the training condition had learned from the training material and therefore performed better on manipulation check. As expected, participants in the training group had significantly different scores on manipulation check ($M = 4.28$, $SD = 1.05$) than those in the control group ($M = 3.02$, $SD = 1.07$), $t(129) = 6.79$, $p = .000$. Also, in this test, there was no violation of the assumption

of equal variance, based on Levene's Test, $F(1, 128) = .08$, $p = .78$. Moreover, the effect size is large (Cohen's $d = 1.18$) for the manipulation check score difference of 1.26, 95% CI [.89, 1.63], between the two conditions. Figure 2 demonstrated the comparison of manipulation check between the two groups. Thus, our training manipulation was effective, and the independent-samples t-test were then conducted to determine the effect of training on SJT performance.
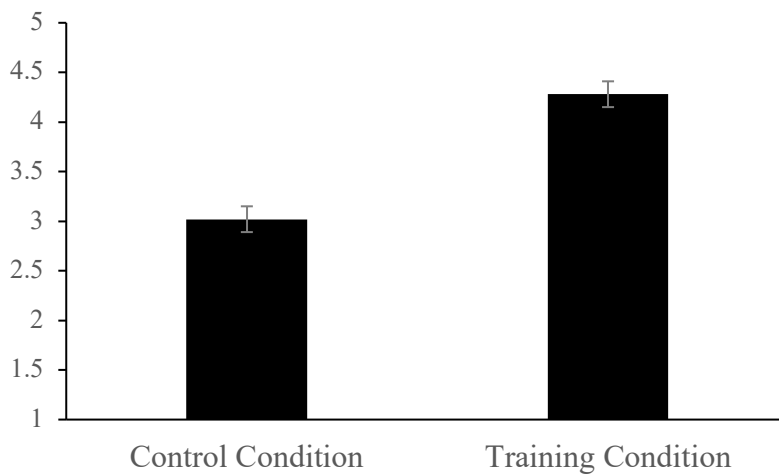

Figure 2. Average manipulation check scores in participants of the control versus training condition. Error bar indicates the standard error of sample mean.

It was hypothesized that individuals receiving ITP training would perform better in an SJT than others. As expected, SJT scores of the training group ($M = 8.77$, $SD = 4.97$) and the control group ($M = 6.09$, $SD = 5.66$) were significantly different, $t(129) = 2.88$, $p = .005$, with an equal variance assumed (i.e., based on Levene's Test, $F(1, 128) = 2.20$, $p = .14$). The effect size is small (Cohen's $d = .32$) for the SJT score difference of 2.68, 95% CI [.84, 4.52], between the two groups. The comparison of SJT scores between experimental and control groups is shown in Figure 3. Based on the results of t-test, *Hypothesis 1* of the current study on the effect of training on SJT performance was supported.
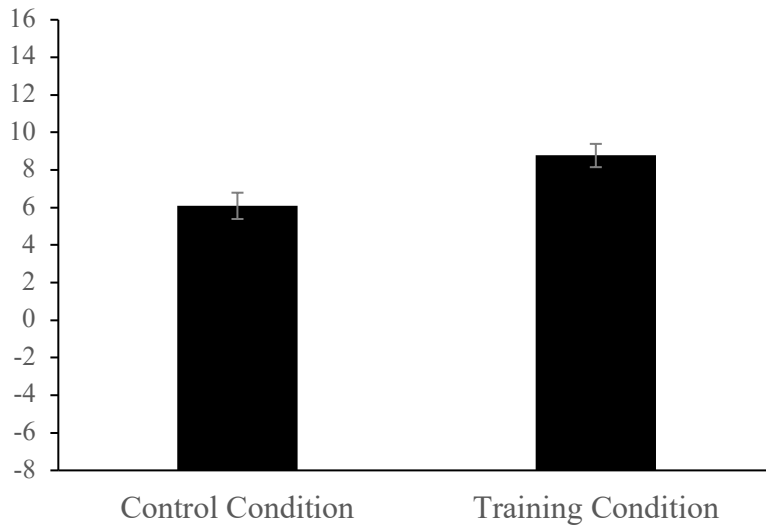
Figure 3. Average SJT scores in participants of the training versus control condition. Error bar indicates the standard error of sample mean.

Multiple Linear Regression analysis. A few multiple regression models were conducted through SPSS (2016) to test *Hypothesis 1*, *2*, *3*, and *4*. The corresponding assumptions were all checked; there was no violation to assumptions of linearity, normality (of residuals), and homoscedasticity. As shown in the Table 5, correlations among interested variables were all below .7, suggesting no issues of multicollinearity.

Table 5.

Correlations of All Measured Variables in All Participants.

	M	SD	1	2	3	4	5	6	7	8	9	10
1. Sex	1.55	0.52	-									
2. Condition	0.50	0.50	-0.08	-								
3. Extraversion	2.87	0.86	0.04	-0.02	-							
4. Agreeableness	3.88	0.70	0.28**	-0.01	-0.34**	-						
5. Conscientiousness	3.70	0.75	0.03	-0.04	0.27**	0.47**	-					
6. Emotional Stability	3.12	0.70	-0.16	-0.07	0.39**	0.27**	0.52**	-				
7. Openness	3.80	0.62	0.14	0.07	0.41**	0.52**	0.44**	0.32**	-			
8. TSDI Integrity	33.98	6.33	-0.05	-0.05	0.37**	0.33**	0.70**	0.47**	0.24**	-		
9. Situational Judgement Test	7.42	5.48	0.28**	0.25**	0.00	0.25**	0.15	0.00	0.17	0.01	-	
10. Cognitive Ability	5.81	3.20	-0.04	0.00	-0.12	-0.02	-0.14	-0.02	0.15	-0.19*	0.39**	-
11. Manipulation Check	3.64	1.24	0.16	0.51**	-0.12	0.09	0.00	-0.05	0.16	-0.13	0.56**	0.34**

Note. Pearson's Correlations are on the Diagonal. All the statistics above were calculated based on the data collected from all sources; details of each sample source are summarized in Table 3. * indicates $p < .05$; ** indicates $p < .01$. Sex: 1=Male, 2=Female, 3=Other; Condition: 0=Control, 1=Experimental.

To test the main effect of ITP training (*Hypothesis 1*), personality (i.e., the degree to which an individual's personality is aligned with the ITP; measured by a personality fit score in the current study; *Hypothesis 2a*), cognitive ability (*Hypothesis 3a*), and sex (*Hypothesis 4*) on predicting participants' SJT performance, a main-effects model with these four predictors was tested (Model 1). The statistics of each predictor in this model are displayed in Table 6. Consistent with the results of *t*-test, training condition significantly predicted SJT performance $b = 3.03$, $p < .001$, which supported *Hypothesis 1*. However, there was no significant main effect of personality (i.e., personality fit score) on SJT performance, $b = -.31$, $p = .093$, suggesting that our *Hypothesis 2a* was not supported. Though, there was a statistically significant main effect of cognitive ability $b = 0.69$, $p < .001$, and sex (i.e., male was coded as 1, female was coded as 2, and "other" was coded as 3), $b = 3.33$, $p < .001$; thus, both *Hypothesis 3a and 4* were supported. The scatterplot of cognitive ability predicting SJT performance and the bar chart comparing SJT scores of male versus female participants are shown in Figure 4 and 5.

Table 6.

Multiple Linear Regression Models Showing Hypothesized Predictors Predicting SJT

Performance

Models	Predictors	b	95%CI	R^2
1	(Intercept)	-1.08	[-5.06, 2.91]	
	Sex	3.33**	[1.78, 4.89]	
	Cognitive Ability	0.69**	[0.44, 0.94]	
	Training Condition	3.03**	[1.44, 4.63]	
	Personality	-0.31	[-0.68, 0.06]	
				0.325
2	(Intercept)	2.96	[-1.88, 7.76]	
	Sex	3.53**	[1.98, 5.07]	
	Cognitive Ability	0.72**	[0.38, 1.05]	
	Training Condition	-4.44	[-10.34, 1.60]	
	Personality	-0.96**	[-1.50, -0.41]	
	Personality × Training Condition	1.13**	[0.42, 1.85]	
	Cognitive Ability × Training Condition	-0.09	[-0.59, 0.40]	

0.375

Note. * indicates $p < .05$; ** indicates $p < .01$. *b* is used to represent the unstandardized beta in the multiple linear regression equation. Values in square brackets indicate the 95% confidence interval for each unstandardized beta. R^2 indicates model fit for each regression model. Personality was measured by a personality fit score, which referred to the degree to which an individual's personality profile was differed from the ITPs.
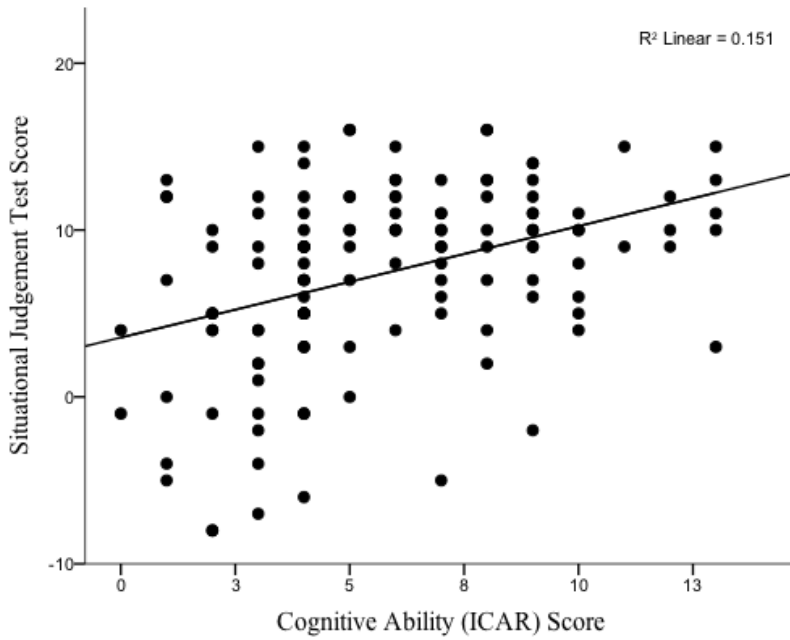
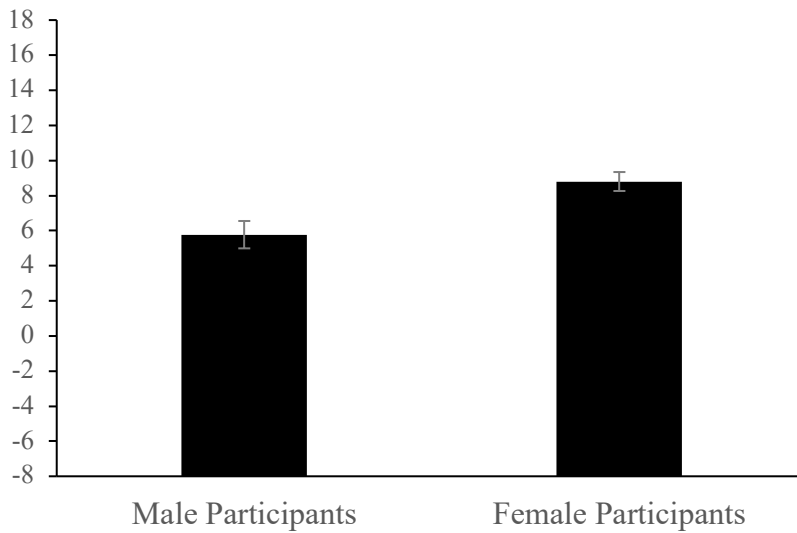*Figure* 4. SJT scores as a function of cognitive ability measured by ICAR.

Figure 5. Average SJT scores in male versus female participants. Error bar indicates the standard error of sample mean.

To test *Hypothesis 2b* and *3b*, based on the previous model, Model 2 added the two-way interactions between training condition and personality (i.e., personality fit score) and between training condition and cognitive ability (i.e., Personality×Training

condition and Cognitive ability×Training condition interaction terms) to the previous

model. As presented in Table 6, the results indicated that there was a significant

interaction effect between personality (i.e., personality fit score) and training condition on

SJT performance, $b = 1.13$, $p = .002$. Although *Hypothesis 2b* did expect an interaction

effect, such a result was in the opposite direction of our original hypothesis. For example,

whereas Hypothesis 2b suggested that individuals with personality profile close to the ITP

would benefit more from the training, our result suggested the opposite. Participants with

personality profile different from the ITP performed better in SJT after training. Also,

there was no effect of interaction between cognitive ability and training condition $b =$

$-.09$, $p = .709$, so *Hypothesis 3b* was not supported.

The interaction between personality and training condition was plotted in Figure

6. The slope for personality predicting SJT performance was $b = -.95$, $p = .001$ in the

training condition, whereas the slope for personality predicting SJT performance was b

$= .18$, $p = .444$. In other words, according to the interaction plot, higher personality fit

scores predicted higher SJT scores in participants of the control group, but there was no

such relationship between personality fit scores and SJT performance in the training
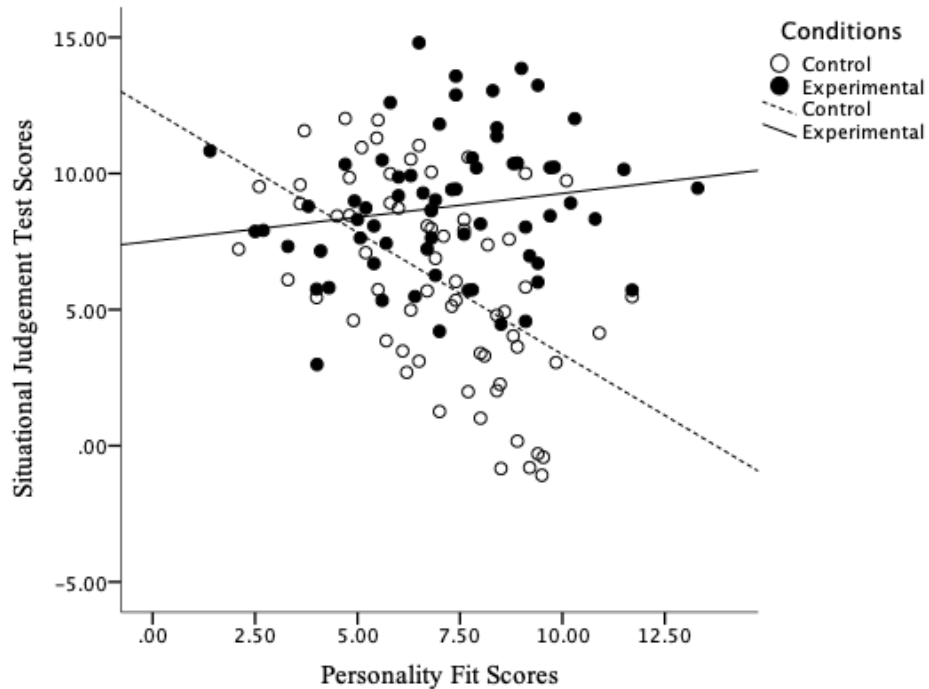
condition.

Figure 6. Interaction effect between participants' personality (i.e., measured by personality fit score) and training condition on SJT scores. The SJT scores were predicted by training condition, personality, sex, cognitive ability, and interaction between personality and training condition.

To specify how each hypothesized factor predicted SJT scores within each condition, the unstandardized beta coefficients of sex, cognitive ability, and personality (i.e., personality fit score) were also specified in Models 3 and 4, respectively, for the training and control condition, as shown in Table 7. However, as noted in the beta coefficients, unlike personality, both sex and cognitive ability had similar prediction effects on SJT performance in two conditions.

Table 7.

Sex, Cognitive Ability, and Personality as Predictors of SJT Performance in Separate

Conditions

Predictors	Training Condition (Model 3)		Control Condition (Model 4)	
	b	95%CI	b	95%CI
(Intercept)	-1.87	[-7.62, 3.88]	3.22	[-2.03, 8.48]
Sex	3.75**	[1.48, 6.02]	3.33**	[1.17, 5.49]
Cognitive Ability	0.63**	[0.27, 1.00]	0.73**	[0.37, 1.06]
Personality fit	0.18	[-0.29, 0.66]	-0.95**	[-1.51, -0.40]
R^2	0.241		0.406	

Note. * indicates $p < .05$; ** indicates $p < .01$. b is used to represent the unstandardized beta in the multiple linear regression equation. Values in square brackets indicate the 95% confidence interval for each unstandardized beta. R^2 indicates model fit for each regression model.

To better understand the relationship between personality and SJT performance,

Multiple linear regression models (i.e., Models 5, 6, and 7) where each of the Big-Five

personality factors (i.e., Extraversion, Agreeableness, Conscientiousness, Emotional

stability, and Openness) predict individuals' SJT performance in the training, control, and

both conditions were fitted to the data. The statistics of each predictor of the three models

are displayed in Table 8. Different results from Models 5 and 6 confirmed that personality

factors from the Big-Five domains predicted SJT performance differently in two

conditions. Specifically, higher Agreeableness significantly predicted higher SJT

performance in the control condition, $b = 3.15$, $p = .044$, and both conditions together, $b =$

2.02, $p = .031$. However, within the training condition, it was the lower Emotional

Stability that significantly predicted higher SJT scores, $b = -2.89$, $p = .009$.

Table 8.

Big-Five Personality Factors as Predictors SJT Performance in each Condition.

Predictors	Training Condition (Model 5)		Control Condition (Model 6)		Both Conditions (Model 7)	
	b	95%CI	b	95%CI	b	95%CI
(Intercept)	8.79	[-0.13, 17.70]	-12.24*	[-22.94, -1.54]	-1.49	[-8.64, 5.67]
Extraversion	-0.13	[-1.85, 1.58]	-1.01	[-2.80, 0.77]	-0.69	[-1.96, 0.59]
Agreeableness	1.76	[-0.34, 3.87]	3.15*	[0.09, 6.21]	2.02*	[0.19, 3.85]
Conscientiousness	1.25	[-0.80, 3.29]	0.01	[-2.36, 2.38]	0.54	[-1.06, 2.14]
Emotional Stability	-2.89**	[-5.03, -0.76]	1.73	[-0.73, 4.19]	-0.71	[-2.39, 0.96]
Openness	-0.59	[-3.12, 1.94]	0.91	[-2.03, 3.84]	0.84	[-1.14, 2.81]
R^2	0.147		0.193		0.085	

Note. * indicates $p < .05$; ** indicates $p < .01$. b is used to represent the unstandardized beta in the multiple linear regression equation. Values in square brackets indicate the 95% confidence interval for each unstandardized beta. R^2 indicates model fit for each regression model.

To test *Hypothesis 5*, Pearson's correlation was calculated between participants' TSDI

integrity scores and SJT scores. There was no statistically significant correlation between

the TSDI and SJT measures of integrity, $r = .01$, $p = .901$, 95% CI [-.14, .16]. Thus,

Hypothesis 5 was not supported. The scatterplot showing the association between

participants' SJT and TSDI integrity scores is shown in Figure 7. As such a result was

unexpected, we also took a further step to look at the Pearson's correlation within

conditions. The results revealed no significant correlation in either the control $r = .05$, p

$= .674$, 95% CI [-.20, .30] or the training condition $r = -.00$, $p = .975$, 95% CI [-.18, .18].
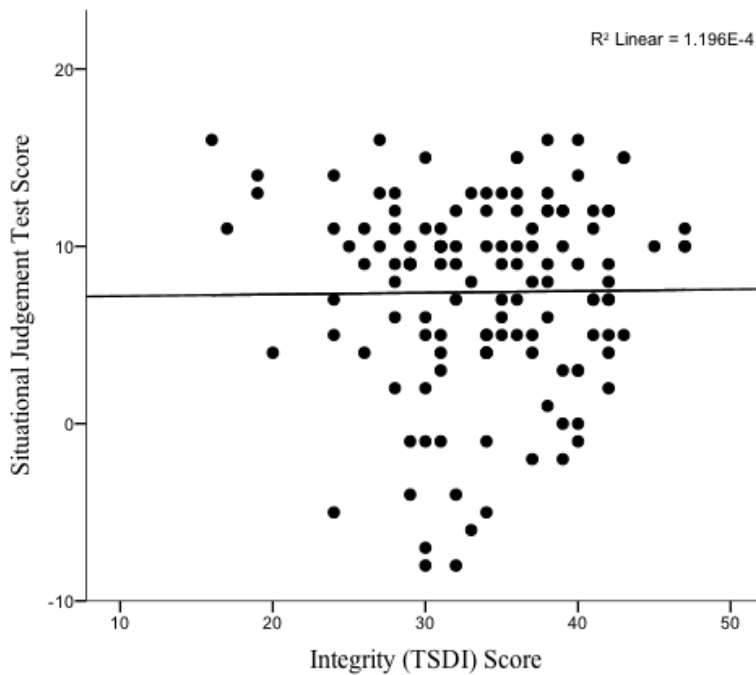

Figure 7. SJT scores as a function of TSDI integrity scores.

Discussion

Effect of Training through ITP

The current study primarily aimed at examining if individuals' SJT performance

would be improved by ITP-based training (*Hypothesis 1*). We tested this hypothesis by

comparing the SJT scores between participants in the training and control conditions.

Participants were randomly assigned to learn the ITP that is favored by the subsequent

SJT (i.e., training condition) or to learn only the basics of Big Five personality theory

(i.e., control condition). Based on *Hypothesis 1*, we predicted that participants in the

training condition would perform better in SJT. The results were as predicted; participants

in the training condition had significantly higher SJT scores, indicating that the ITP based

training had effectively improved participants' SJT performance.

Our finding agreed with Motowidlo and Beier (2010)'s new conceptualization of

SJT. These authors proposed that there are two components underlying the scores of SJT

(i.e., general domain knowledge and specific job knowledge components). This notion

differed from the traditional view that SJT only assesses context-dependent knowledge. In

this view, individuals' performance on SJTs would only reflect their procedural

knowledge of how to respond to specific job situations, and individuals would also be

required to have relevant work experience to perform on SJTs (Lievens and Motowidlo,

2016). However, the newer SJT model proposed that some judgements in SJTs may not

require individuals to have actual experience in job specific situations. According to

Motowidlo and Beier (2010), in addition to the traditionally emphasized specific job

knowledge (i.e., acquired through relevant job experience), a more general and context-

independent knowledge, which could be learned during the general process of

socialization, may also allow individuals to respond effectively to SJT situations.

Motowidlo and Beier (2010) presented evidence to support this notion through comparing

the SJT scoring keys prepared by university students and SMEs. They found that the

scoring key prepared by university students successfully predicted job performance

although they did not have specific job knowledge.

Although Motowidlo and Beier (2010) had discussed some empirical evidence

and demonstrated how their own research findings would support the new SJT

framework, such a reconceptualization of SJTs is relatively new. Through a controlled

experiment, the current study found evidence to support their position. Specifically, the

current study adopted student samples, in which all individuals were recently registered

as full-time students in post-secondary institutions. This group of individuals were likely

to have limited specific job experience available to respond the SJT items, meaning that

the current study had a relatively effective control for experience. Then, by training them

in general domain knowledge (i.e., the ITPs), we found a statistically significant

improvement in SJT performance in participants of the training condition. In other words,

the current study was able to demonstrate that, although being short of specific job

experience, individuals who have been through brief general knowledge training on ITPs

were able to perform better on an SJT. Such a result supported the existence of the

general knowledge component (i.e., that is context independent) in SJT and that

individuals' performance corresponding to this component can be improved through

training.

Our finding did not only agree with but also added more evidence to the new SJT

framework. Motowidlo et al., (2006) suggested that the knowledge of ITPs is the general

domain knowledge component underlying individuals' SJT scores. Unlike specific job

knowledge that would be hard to summarize into brief words but needs to be obtained

through day-to-day experience at specific jobs, ITP is more explicit knowledge about how

to express personality traits that can be summarized into short rules. Because such a new

conceptualization of SJT has only been proposed recently, the notion that ITP is the

general domain knowledge component of SJT was still left to be tested. The current study

approached this issue through developing the training material specifically focusing on

ITP. The rules of trait expression in Becker's SJT was summarized into Big Five

personality factors. Because ITP is supposed to be the general domain knowledge, the training materials we developed were general and not specific to certain job situations. Our results demonstrated that those who had learned through the ITP based training material performed better than those who did not. Although our finding cannot directly prove that the knowledge of ITPs is the general knowledge component of SJT in general, but we might be able conclude that this type of knowledge is at least one important general knowledge component underlying the performance on Becker's SJT (2005). This finding would provide a premise for future works that attempt to define the content of the general knowledge component of SJT.

Results of the current study would also help to explain why the study by Cullen and colleagues (2006) found a significant training effect in one SJT but not the other. Specifically, this previous study examined the effect of training on two SJTs. One of the SJTs (i.e., CSQ) consisted of more straight forward general domain knowledge, and the other SJT (i.e., SJI) consisted of more complex items with more context-specific situations. The authors found a significant training effect on undergraduate participants' performance on CSQ but not on SJI. Our finding helped to explain why the effect of training was significant on one SJT but not the other. On the one hand, the SJI mainly focused on the specific job knowledge, which was more context-dependent and required relevant experience to perform well. Without experiencing the specific job situations, the participants could not benefit from a verbally communicated training. On the other hand, the CSQ mainly focused on the general domain knowledge, which could be easily trainable through a verbal communication on trait expression rules. Due to its nature, the general knowledge was communicated easily in words and could be applied to the situations in the CSQ. As a result, training improved participants' performance in this

particular SJT. However, the current findings also lead us to further questions that need to be addressed in the future. For example, whether or not different SJTs have different levels of trainability, and whether such trainability can be determined by the composition of SJTs in terms of the general knowledge and the specific job knowledge (e.g., the less focus on specific job knowledge, the higher trainability of an SJT). These questions are critical to be addressed in the future to provide us with more insights into SJT as a selection tool.

Individual Differences

Personality. The current study also attempted to examine how individual factors, such as personality, cognitive ability, and sex would impact participants' performance on SJT. The current study hypothesized that the participants who have personality traits aligned with the corresponding ITPs would overall perform better than others in the SJT (*Hypothesis 2a*). This hypothesis was partly supported by our results. Specifically, the tendency that the closer their personality aligned with the ITPs the higher SJT scores did not apply to all participants but only to participants of control condition. On the one hand, individuals with personality close to the ITPs would naturally perform better in the SJT, which would support the personality element in Motowidlo and Beier's (2010) new SJT conceptualization. On the other hand, with the presence of the ITP based training, individuals with closer personality profiles to the ITPs would no longer have such an advantage. Thus, in essence, we might conclude that personality do play an important role in SJT performance, but its influence can be moderated by a training on the knowledge of ITPs.

If individuals had a chance to learn about the ITPs underlying the SJT, their own personality predisposition did not matter. This demonstrated a possibility that individuals

who are not naturally equipped with the ideal personality profile to a job can still learn the rules of trait expression and perform as well as their counterparts. Although individuals' personality traits were relatively stable and hard to change, they could still learn certain ways of interpretation and behaviours to achieve desired performances. ITP-based training can be beneficial to them, because those who did not receive such training performed less well than their counterparts who had natural ideal personality profile. It is an important finding because it may allow us to conclude that ITP-based training can improve SJT performance in individuals who do not have a "perfect" personality profile. However, it could also raise a question that requires future investigations – how the improved SJT scores would predict candidates' actual work performance after their personality difference was minimized. Thus, the current finding merely provided an initial understanding of the effect of ITP training, but the following steps will need be taken to reach further conclusions.

However, different from what was hypothesized, individuals with personality profiles closer to the ITPs did not benefit more from the ITP training (*Hypothesis 2b*); rather, they did not benefit from it at all. These individuals did not have SJT scores higher than those who had similar personality profile but did not receive training (i.e., their counterparts in the control condition) or those who received training but did not have ITP aligned personality (i.e., their counterparts in the training condition). Thus, our findings do not support McCrae and Costa's (1996) notion of characteristic adaptation and Motowidlo's (2003) notion of disposition fit. Specifically, in this point of view, individuals' acquisition of various domains of knowledge, including ITPs, comes from interactions with their environments, and such a learning process would be influenced by their natural tendencies (i.e., personality and personal interests). In other words, being

prepared with an ideal personality profile would facilitate the acquisition of ITPs; these individuals were expected to show even more improvement in the current study. However, our results did not support this view. Nonetheless, the current study did not present direct evidence against these previous notions; rather, it is possible that, to perform on an SJT, those who were naturally equipped with the ITP aligned personality had already performed at their best and that anymore training on this particular type of knowledge would not benefit (i.e., ceiling effect). For example, the knowledge of ITPs might only account for a limited part of SJT scores, so a training intervention on ITPs would not benefit individuals who are already good at it.

However, there might be other ways to explain the current finding, and the previous notions (McCrae & Costa, 1996; Motowidlo, 2003) might remain true. First, it is possible that individuals who had personality aligned with the ITP would learn better in general, but our short training study did not capture their advantage over other individuals. In other words, during their process of characteristic adaptation, the predisposition fit did play an important role in day-to-day learning experience and it might still have contributed to their better SJT performance. Yet, due to the nature of our experiment, such a daily adaptation process could not be captured in a one-shot observation. Thus, further explorations on ITP learning process through daily socialization might be conducted in the future to clarify how personality and acquirement of ITP knowledge would interact. Second, the current finding may only apply to the particular SJT that we used. The ITP in Becker's SJT (2005) was relatively straightforward and intuitive, which may limit the room of improvement for those who have already equipped with behavioural tendency consistent to the ITP. An SJT with more complicated ITP may be tested in the future to explore this individual difference.

Additionally, although not hypothesized, we took a further step and analyzed the relationship between each of the Big-Five personality domain and SJT performance in the current sample. Previous studies found that Agreeableness, Emotional Stability, and Conscientiousness significantly predicted scores of integrity measures. In the current study, participants across two conditions and specifically in the control condition had SJT scores significantly predicted by Agreeableness. This result agreed with previous literature (Neuman & Baydoun, 1998; Sackett & Wanek, 1996), confirming that Agreeableness is an important predictor of integrity, although Becker's SJT adopted a different format than other integrity measures. However, neither high Emotional Stability or Conscientiousness had significantly predicted higher SJT scores in the current study. It remains unclear why our findings of these two domains were different from the previous studies (Ones et la., 1994; Neuman & Baydoun, 1998; Sackett & Wanek, 1996). However, note that the combination score calculated from all Big-Five domains (i.e., personality fit score) did significantly predict SJT scores in participants who received no ITP training (i.e., the control condition) in the current study. Thus, it is also possible that such a finding was due to a lack of power (i.e., only participants in the control condition could be used to examine the relationship between SJT and personality, without being influenced by the ITP training). Further investigations might be needed to clarify how Becker's SJT can be predicted by each of the Big-Five domains.

Further, interestingly, the current study found that in the training condition, participants' lower Emotional stability significantly predicted higher SJT scores. It was unexpected, because previous studies showed positive relationship between Emotional Stability and integrity measures; the negative direction found in the current study seemed counterintuitive. One plausible explanation is that, although in natural settings (i.e., in the

control condition), Emotional stability did not predict participants' SJT scores, for people

who have low Emotional stability, the ITP training had helped them to learn the

importance of managing one's emotional impulses and guided them to respond to the SJT

situations more effectively. However, since such a finding is relatively new, future studies

will be needed to explore the underlying explanations. As well, because Becker's SJT is

not only an SJT but also a measure of integrity, it is unknown if such an effect was

specifically on an SJT (as a test format), on a measure of integrity (as a construct), or on

an integrity SJT (as a combination). Future studies may attempt to address this issue by

examining the effect of Emotional stability on SJTs measuring other constructs or on

integrity measures in other formats.

Cognitive Ability. The current study also hypothesized that individuals with

higher cognitive ability would outperform others in SJT (*Hypothesis 3a*). Results of the

current study indicated that higher cognitive ability scores significantly predicted higher

SJT scores, which supported our hypothesis. This finding also agreed with and added

evidence to Motowidlo and Beier's (2010) new SJT framework, such that cognitive

ability is an important personal predictor of SJT performance. The relationship between

cognitive ability and job performance can be explained by our previous knowledge on

how cognitive ability influence individuals' information processing and acquisition.

Specifically, individuals with higher cognitive ability would have higher capacity to

acquire and process information (Motowidlo & Beier, 2010). Hence, in daily experience,

these individuals may acquire the knowledge of ITPs better and perform better on SJT.

However, individuals with higher cognitive ability might not benefit more than

others from a brief training such as the one in the current study. Our hypothesis on the

interaction effect between cognitive ability and training (*Hypothesis* 3b) was not

supported by our results. In other words, in the current study, although there was a training effect across participants with various levels of cognitive ability, such an effect was not stronger in those who with higher levels of cognitive ability. Similar to our finding on personality, it is possible that the ITP training and SJT in the current study lacked higher complexity so the individuals who were equipped with higher cognitive ability could not take such an advantage (i.e., the materials were simple enough so that individuals with various levels of cognitive ability had similar levels of learning). Alternatively, there might be different knowledge acquisition processes involved in ITP learning between specific training and during daily life. For example, while individuals with higher levels of cognitive ability may acquire ITPs more effectively in daily life, when such an information was extracted, summarized into digestible materials, and presented in a formal training session, other individuals may also be able to acquire such a knowledge and achieve the same result. If that is the case, the current experiment was not able to capture such individual difference. Thus, future investigations may be conducted to clarify if and why there might or might not be differences in acquisition of the knowledge of ITPs among individuals with different levels of cognitive ability.

The current study also hypothesized a sex difference. Specifically, we hypothesized that sex is also an important predictor of SJT performance, such that female candidates would overall perform better than male participants (*Hypothesis 4*). In the current study, indeed, female participants had a higher average SJT scores, which supported our hypothesis. This finding was congruent with the previous conclusion made by Whetzel et al. (2008) through a literature review and meta-analysis (i.e., women tend to achieve slightly higher SJT scores than men when an SJT is correlated with Conscientiousness and Agreeableness). In the current study, the Becker's (2005) SJT of

integrity indeed favored high Conscientiousness and Agreeableness. Thus, with a natural tendency of higher Conscientiousness and Agreeableness, the female participants in the current study had an overall higher SJT score by responding to the SJT items toward high level of Conscientiousness and Agreeableness. However, the sex difference in SJT performance may not be entirely explained through personality differences between men and women, because sex showed an incremental validity (i.e., $b = 3.53$ after personality was held constant) in predicting SJT scores even when personality was measured and presented in the same regression model. However, it remained relatively unknown what might have contributed to such a sex difference other than personality. As well, based on the current data, none of the other measures (i.e., cognitive ability and integrity) were associated with sex. Thus, these two variables are unlikely an explanation to the sex difference in the SJT performance. Future investigation might be conducted to explore the alternative explanations (e.g., the interpersonal nature of the situations in SJTs).

Integrity Measures

The current study also aimed to examine the relationship between different measures of integrity. The current study used Becker's (2005) SJT to examine the effect of ITP-based training. This test was developed to measure employees' integrity using SJT format. Unlike the other more overt measures of integrity, Becker (2005) argued that the integrity SJT would be more realistic (i.e., using situations at work) and less transparent (i.e., respondents may not identify that they were being tested on integrity). This author has validated such a measure; however, subsequent studies examining its concurrent validity with other integrity measures was lacking. The current study hypothesized that participants' SJT scores would be correlated with their scores on the TSDI integrity measure, developed by (Catano et al., 2018). However, our results did not find a

significant correlation between Becker's SJT (2005) and TSDI (Catano et al., 2018). Our *Hypothesis 5* was not supported.

Although the current study failed to demonstrate correlations between two integrity measures, both of Becker's SJT or the TSDI integrity measure might still be valid because they tend to focus on different facets of integrity. First, development of these two assessments were based on slightly different definitions and revealed different underlying factors. On the one hand, TSDI (Catano et al., 2018) is a purely personality-based integrity measure that defined the construct of integrity as reliability, dependability, honesty, and Conscientiousness. Factor analysis underlying this integrity test revealed three personality factors, namely, Conscientiousness, Agreeableness, and Emotional stability. On the other hand, although a good amount of personality related information could be extracted from in Becker's SJT (2005), the author saw integrity as a more complex construct that consists of higher-level combinations of personality and other related principles (i.e., honesty, Conscientiousness, etc.). Becker (1998) also opposed to the view that integrity was a linear composite of Big-Five personality traits; he argued that the construct of integrity should also integrate an objective code of morality. For example, according to the author, a person with integrity should not only be honest but also act on a rational and morally justifiable code. Viewing integrity as a more complex construct, the development and validation of Becker's SJT revealed four higher level factors than simple personality traits, namely, positive relationships, career potential, leadership, and in-role performance. Thus, although both measures were imbedded with personality related information, their construct definitions and underlying factors may have contributed to their uncorrelated scores.

The factors underlying Becker's SJT (2005) also seem to have a more specific emphasis on the workplace. Indeed, the second major difference between these two integrity measures lies between their purposes of application. Specifically, Becker (2005) intended to apply such an integrity measure to settings where candidates' responses to complex moral situations at work might be used to predict promotion, career progress, and leadership status. Thus, Becker (2005) presented complex moral dilemmas instead of directly assessing candidates' integrity related personality traits, because these traits may not be able to directly reflect candidates' tendency and behaviour in real work situations that demand integrity. In contrast, the authors of TSDI (Catano et al., 2018) envisioned their integrity measure as a selection tool for security focused organizations. These organizations need employees who are unlikely to engage in dishonest, unreliable and counterproductive work behaviours. These behaviours tend to be less complex than employee performance for promotion and leadership. Thus, a more overt trait assessment that focuses on reliability, dependability, honesty, and Conscientiousness without introducing complex moral dilemmas may be better targeted on organizations such as police or military. Therefore, such different focuses for application may contribute to the low correlation between the two integrity measures.

Practical Implications

There are some important practical implications based on the current findings. Generally speaking, for HR practitioners and decision-makers, the current study has provided them a better understanding of SJT, which they can utilize (i.e., there are more than one knowledge components underlying SJT and the general domain knowledge component can be trainable in a short time) when making hiring decisions. For example, when interpreting candidates' SJT scores, it is important for them to acknowledge that

these scores may actually reflect two types of knowledge, the specific job knowledge and the general domain knowledge, and that the latter one is improvable by test preparation.

The trainability of SJT performance may raise concerns in test developers. For example, if candidates could be trained on the knowledge of ITPs and perform better on an SJT, would this test still maintain its validity in predicting job performance? Should test developers implement strategies to avoid such training effect? Findings from both of the current and previous SJT training studies would suggest that these concerns might not be necessary. First, previous studies examining SJTs' criterion-related validity found that there was minimal or no impact from the effect of training (Cullen et al., 2006; Stemig et al., 2015). Based on this evidence, SJT's trainability may not threaten test developers' interest in predicting candidates' work performance. However, future studies are encouraged to accumulate more solid evidence.

Second, compared with the previous study (Cullen et al., 2006), the current training on the knowledge of ITPs emphasized a stronger connection between training and its application in actual work situations. Cullen et al.,'s (2006) training represented an explicit purpose of improving SJT performance by using test-taking strategies (i.e., avoid extreme responses). For example, they used phrases such as "answer as a responsible person" and showed participants "good answers" and "poor answers" (p.146). Such language might have created an impression in participants that the training would only benefit test performance. In contrast, the current training presented ITPs as an explanation for good workplace behaviours instead of test-taking strategies, and our training material was carefully phrased to represent the idea of how to become a high-performing employee. For example, we introduced participants to the ITPs through phrases such as "knowing certain rules about how to behave properly at work may improve your

competitiveness". Thus, our training would have a better chance to link the participants' test preparation with their actual job performance. Based on this comparison, even though the current study did not measure job performance and could not draw a solid empirical conclusion, we would speculate that such training may have an even smaller impact on an SJT's criterion-related validity than the previous study.

Third, SJT is a knowledge-based test, which assesses the result of learning instead of the process of learning. In other words, through which experience have the candidates learned the ITPs might not be important. For example, whether this type of knowledge is learned during everyday life or from a training session, both learning experiences would provide the mental resource that allows candidates to perform on an SJT and potentially at work. Also, if an SJT is developed to test candidates' knowledge, identifying the underlying learning processes may not provide obvious advantages to test developers. Thus, the fact that some candidates may achieve improved SJT scores due to training might not pose any obvious negative effect.

Additionally, Cullen et al., (2006) revealed that the SJT scoring keys based on a discrepancy approach (i.e., candidates' judgements of behaviours' effectiveness on a Likert-scale were compared with the responses provided by an expert group) might threaten an SJT's utility in test developers' interest. In their study, simple test-taking strategies such as to avoid extreme ratings had resulted in a large-scale increase of SJT scores. These authors, therefore, suggested that test developers may be cautious about using this approach, because candidates may exploit its feature to achieve instantly higher scores. In the current study, we used Becker's SJT (2005) that asks respondents to select the best answers from the specific behavioural responses provided by SMEs, so such an issue was avoided. However, because our main focus was not on simple test-taking

strategies, the current work could not provide further insights into this issue. If test developers hope to avoid it, we would agree with Cullen et al. that cautions may be taken in the use of discrepancy-based marking approach.

In the current study, without further investigations of how ITP-based training would influence actual job performance, practical implications based on the training effect may be more obvious to individuals (i.e., job seekers). In general, individuals who receive ITP-based training may achieve better performances on SJTs. More specifically, however, some individuals might utilize such training to overcome their natural weaknesses. For example, without training, individuals with personality profiles aligned with the ITP might perform better than others on certain SJTs and jobs. However, ITP-based training would minimize this difference and benefit individuals who do not have the ideal personality profiles for certain SJTs. Thus, these individuals may participate in such training and improve their SJT scores. Moreover, based on the composition of the current sample, individuals with more ideal personality profiles to an SJT might only represent a minority. Thus, the current study may also indicate that ITP-based training can be beneficial most of the population. More practically, this type of intervention might be integrated into career development programs to help job seekers to achieve more success in job search.

Similar to the advantage of having certain personality profile, individuals with higher cognitive ability would also outperform others in SJTs. The fact that higher cognitive ability predicted both SJT and job performance may indicate that they may share common underlying factors that are related to cognitive ability. Although it would be relatively impractical to recommend individuals to improve their cognitive ability in order to improve SJT performance, our findings on the relationship between cognitive

ability and SJT performance may be valuable for HR practices. For example, HR practitioners may need to acknowledge that, whether or not other types of tests such as an SJT would be conducted, cognitive ability tests should always be emphasized during selection process. In a circumstance that adopting an SJT paradigm is not feasible, a cognitive ability test might always be confidently administered and used for hiring decisions. This is because first, according to our well-established understanding, cognitive ability is the best predictor of job performance (Schmidt & Hunter, 1998), and second, the results of cognitive ability test may yield a similar decision as the ones based on SJT scores. However, based on the current findings, the correlation between one's cognitive ability and SJT score was only medium, which would suggest that cognitive ability cannot explain the SJT performance entirely. Indeed, SJTs are often used in selection processes because organizations are interested in how candidates would respond to certain work situations. Although cognitive ability tends to predict candidates' SJT score, such a measure would not provide information on candidates' judgements and responses to work related situations. Thus, SJTs are still considered as an important tool in selection practices and would provide more insights about candidates than cognitive ability measures.

The current study also confirmed previous findings on sex difference in SJT performance. Female candidates tend to have an advantage in performing on SJTs, especially those that favour high Conscientiousness, Agreeableness, and Emotional Stability. As well, the fact that sex remained as a significant predictor for SJT controlling for the personality factor may suggest that sex was not a redundant factor for SJT performance in addition to personality. Thus, in practice, when SJT scores are used as a selection criterion, decision makers might need to be cautious about the difference

between sex, because SJTs might naturally favor female candidates. Although the underlying reasons need to be further researched, the current finding suggested that simply comparing candidates' SJT scores across sexes might introduce an adverse effect on male candidates.

The current study failed to demonstrate a positive correlation between Becker's SJT (2005) and TSDI (Catano et al., 2018) integrity measure would call for caution when applying these integrity measures. In addition to their use of different testing formats and degree of transparency, there were also differences underlying their definitions of integrity. Although the elements within their construct definitions are mostly overlapping, Becker's SJT focused on a higher complexity of its form, so candidates who wrote this test achieved different scores. The unrelated scores are also in line with different application purposes (i.e., to be used different types of organizations, jobs, and predictions), which may largely influence the tests' design and development. For example, with a purpose of helping practitioners to make decisions on promotion, Becker's SJT focuses on employees' behavioural tendency when encountering social situations that involve moral dilemma. In contrast, to be applied to organizations that emphasize security, such as police stations, TSDI consists of questions that can screen out individuals who may engage counterproductive work behaviours. In practice, therefore, depending on the reasons why organizations choose to assess candidates' integrity, their needs might need be carefully considered before deciding which integrity measure will be used.

However, the current study may have other practical implications than the current recommendations. In a long run, the current study initiated a first step to test the effect of ITP-based training; we demonstrated that such a training had a positive influence on

individuals' performance on SJT. However, as a selection tool, SJT was often used to predict job candidates' work performance, thus the eventual practical application of ITP-based training should be related to candidates' actual performance at work. In other words, whether or not SJT can maintain its criterion-related validity after training and individuals with improved SJT scores would also have improved work performance would help us to speculate the ultimate effect of ITP-based training on organizations. Although there was an expectation that an improved ITP knowledge would be associated with an improvement of actual work performance (Motowidlo & Beier, 2010), whether or not the SJT scores can maintain its ability to predict work performance remains to be clarified.

Considering the findings of ATIC literature, which suggested that a reduction in test transparency had improved individuals' test performance but their test scores did not maintain the previous level of criterion-related validity (Kleinmann, 1997; Smith-Jentsch, Salat, & Brannick, 2001), one might infer that improving candidates' knowledge of ITPs would also decrease the SJT's validity in predicting candidates' actual work performance. Nonetheless, important differences between these transparency studies and the training current study need to be considered as well. The transparency studies and the current training study did not share the same research purpose. On the one hand, the research idea behind the transparency studies was that reducing the advantage of ATIC in some individuals would provide further evidence to support ATIC as an essential component in the employment tests that predicted job performance. Thus, the manipulation of making performance criteria transparent had reduced the difficulty of assessment hence candidates who were not able to identify these criteria was also able to perform well (Kleinmann et al., 2011). On the other hand, the research idea behind the current ITP

training study was that improving candidates' knowledge of ITPs and observing their improvements on SJT would provide an evidence to support the general knowledge component underlying SJT.

Unlike manipulations in the transparency studies that made the tests easier for candidates, the current training study tried to improve candidates' SJT performance through improving their relevant knowledge. Specifically, in the transparency studies, candidates who were not able to identity criteria achieved higher test scores because the tests were made easier, but these tests scores would not reflect their true ability and could not be used to predict their work performance. In contrast, in the current training study, there was no manipulation of the test; rather, the participants were better equipped with the relevant knowledge, and such knowledge was then reflected by their scores on the corresponding knowledge test (i.e., the SJT). Because there was no modification on the test, it is likely to maintain the ability to reflect the participants' true knowledge.

In fact, in the ATIC literature, when the test was kept original while candidates' ability was improved, research evidence did suggest a maintenance of criterion-related validity. For example, Schmit and Ryan (1993) studied candidates who had identified the underlying performance criteria of a personality inventories. These participants altered their responses to represent an ideal profile scores when taking the measures, but their scores did maintain the criterion-related validity. These authors explained that this might be because both performance on employment tests and on jobs require the candidates' ability to identify the demanded type of personality. Unlike transparency studies, Schmit and Ryan's (1993) study has an important similarity to the current study because both had a focus on the improvement in candidates' ability of knowledge, instead of modifying the measurement itself. Thus, when avoiding modifications to the SJT, the current ITP

training that focused on improving individuals' cognitive resources may not pose an influence on validity of the test. Such learned knowledge might be carried over to actual work and therefore maintain the SJT's validity and even increase employees' job performance. The recent studies (Stemig et al., 2015; Cullent et al., 2006;) provided evidence to support such a notion. These authors found that the criterion-related validity of an SJT was minimally decreased by training. However, in these studies, the coaching or training programs were not specifically designed to train ITP knowledge; rather, they were either a general commercial program without a clearly defined methodology and underlying construct or a training intervention on direct test-taking strategies. Thus, their findings would only provide indirect evidence to support our speculation on the effect ITP training.

Therefore, without a direct investigation, how ITP-based training would influence SJT's criterion validity and employees' actual job performance remains an unresolved question. It is important to be addressed in the future, and the related practical implications for organizations might then be speculated. If SJT's validity would be decreased by such training, employers need to be cautious about the training effect. However, if such training does not influence the test's validity and candidates' improved performance can be transferred into actual work situations, such an intervention might also provide potential benefit to organizations. For example, employers do not need to concern the effect of training on candidates' SJT scores; rather, they can be more confident that candidates' performance on SJTs would be a valid predictor of their future performance. In this case, the benefits of such training might not be limited to organizations' selection practices (i.e., candidates with satisfying SJT performances would also perform well in jobs) but also in terms of training. For instance, when there is

a higher cost associated with hiring new employees, organizations may instead consider training the current employees, if the area of improvement is relevant (i.e., employees need to improve on how they handle social situations). Thus, based on the current findings, further examinations on how SJT's criterion validity would be influenced by ITP-based training may introduce a broader range of practical implications, which are particularly beneficial to organizations.

Strengths, Limitations, and Future Directions

The current study has a few important strengths. First, the controlled experiment allowed us to test the training effect as an isolated variable and to reach causal conclusions about the effect of training on SJT scores. Second, because the current study was conducted on a student sample, who had relatively less work experience, findings from the current study might have meaningful application to the same population, especially those who will graduate soon and are looking for employment. Unlike other job seekers, candidates with less work experience may perform less well on SJT, which supposedly favors candidates with richer experience in doing particular jobs. However, for those who are relatively inexperienced but would be competitive candidates otherwise (e.g., having strengths in all other aspects that are required for their applied jobs), they may benefit from ITP-based training and experience greater success in job search.

Nonetheless, our use of undergraduate samples also has limitations, especially in the aspect of HR selection practices. While the current findings may be easily generalized to students or inexperienced job candidates, conclusions about other populations remain unclear. For example, whether or not more experienced candidates would have similar response to ITP-based training, and if practitioners can interpret their SJT scores and the underlying components as the current findings would suggest for the inexperienced

candidates. This limitation may prevent our current knowledge about SJTs and the effect of training from being generalized to the larger pool of job candidates. In personnel selection processes, the composition of candidate pool typically does not only include new graduates but also includes individuals with various job experiences. Some candidates may have rich experience to the jobs that they are applying, while some others may have other work experience but irrelevant to the jobs that they are applying. Thus, in practice, depending on jobs, organizations, current economics, and other factors, practitioners and decision makers may be facing various types of candidates. In this case, knowledge of SJT and the effect of training among different populations can be beneficial. Thus, to have broader influences on selection practices, the current study needs to be conducted with samples from the other populations.

Furthermore, in terms of the training effect, it still remains unknown how ITP-based training would influence SJT's ability to predict job performance; however, it may have further benefits if its criterion-related validity is maintained. As discussed in more detail in the previous section, the ITP-based training could potentially benefit organizations (i.e., trainees may transfer their knowledge of ITP to judge and respond effectively to real work situations and therefore achieving better job performance). Again, however, the current study is only the first step in exploring the effect of ITP training; how ITP-based training would influence SJT's criterion related validity needs to be addressed through empirical research. While the current study has a strength of allowing causal conclusions about the training effect on test performance, the limitation is that we cannot make further conclusions on actual work performance. Thus, we recommend future investigations that follow up the current study to address this limitation.

In addition, the fact that the current study scheduled the personality and integrity (i.e., TSDI) measures before the ITP training may have left some unresolved questions. With such a design, it remains unclear if participants would respond to these measures differently if they were scheduled after training. In other words, there might be a chance that our findings could be altered if personality and integrity measures were taken after training. Because both of these measures were relevant to the training material, the current study decided to arrange these measures before training so that we could collect participants' baseline information (i.e., the baseline information allowed us to describe the participants more accurately in a natural setting, so the results could be easily summarized and generalized to similar individuals). However, this design probably had only captured the relationships between participants' post-training SJT performance and baseline personality or integrity. Although the similar relationships when all variables were at baseline could be inferred based on our data in the control condition, how the variables would relate to each other if they were all measured after training remains unknown and may need to be revealed in the future.

Future studies are also needed to provide further explanations to why individuals with advantageous personality profile or cognitive ability did not show corresponding advantage in the ITP-based training in the current study. As discussed in previous sections, there are a few possible explanations, such as that the low complexity of materials was not capable to differentiate individual differences in training and that different knowledge acquisition processes may have involved in acquisition of the knowledge of ITPs between everyday learning and a training setting. These alternative explanations need to be examined in future studies, for example, through testing the

training effect with more complex ITPs or monitoring individuals' learning process in more naturalistic settings.

One might question that, if future studies examine other SJTs with more complex ITPs, whether or not the training effect would still be present. However, such training effect should be expected because of the following reasons. First, based on a controlled experiment, the current study holds a relatively confident conclusion on the causal relationship between the training intervention and participants' improvement in SJT performance. In fact, such a finding supports not only the effectiveness of training but also the trainability of the general domain knowledge. Other training interventions focusing on the same knowledge component would also have the potential to improve SJT performance.

Second, because the current study was only a preliminary investigation of the effect of ITP-based training, our training intervention had a relatively simple design and was presented a low-fidelity format. Unlike others, such as the more comprehensive training or coaching programs that typically involve trainer-trainee interactions and practice questions, the current training only involved a short reading material. In this material, only the most critical information was delivered, and the participants were not given any chance to ask questions and address confusions. Thus, if future studies are conducted to examine SJTs with more complex ITPs, we recommend a more comprehensive intervention to deliver such complex knowledge. For example, it might include trainer-trainee interactions, multi-media presentations, and practice questions to ensure participants' understanding. Compared with the one in the current study, such an intervention might be more likely to reveal a training effect.

Finally, SJTs with more complex ITPs may imply more room of improvement. As discussed in previous sections, individuals with advantageous personality profiles may have experienced a ceiling effect, so they did not show any improvement from the training. Such a ceiling effect was reasoned as a result of the low complexity of the ITPs. Thus, when the ITPs become more complex, a training intervention may become more beneficial and lead to higher degrees of improvement in these individuals. Therefore, future studies examining SJTs with higher complexity of ITPs might be unlikely to find no training effect; rather, if they invest in a better-designed training intervention with comprehensive contents and effective delivery media, it might reveal a stronger training effect.

Conclusion

In short, the current study has examined the effect of ITP-based training on SJT performance and suggests that there might be a training effect. Our finding has also provided evidence to support Motowidlo and Beier's (2010) new SJT model, which proposed that there is a general knowledge component (i.e., the knowledge of ITPs) in SJT scores. At the same time, we also found individuals with personality profiles closer to the ITP or with higher cognitive ability might outperform others in SJTs. Though, in the current study, these types of individuals did not benefit more from the ITP training, thus we call for further investigations to clarify and explain this phenomenon. As well, women tended to outperform men in Becker's SJT (2005). We might explain this tendency through two sexes' different personalities. However, because sex showed incremental validity in predicting SJT performance with the presence of personality, we recommend future explorations on the alternative explanations to the sex difference in SJT performance. Finally, the current study found that individuals' scores in the two integrity

measures, Becker's SJT (2005) and TSDI (Catano et al., 2018), might not be correlated. Such low concurrent validity might be explained by their slightly different construct definitions and the particular practice purposes that the two selection tools were developed for.

Our findings might have a variety of practical implications, such as to benefit inexperienced job candidates during job search and HR practitioners during decision-making process. However, more importantly, the current study initiated a first step in testing the effect of ITP-based training. Although this study would only allow us to draw conclusions on the training effect in terms of test performance, future investigations can follow up to look for such effect on actual job performance. If ITP-based training does not influence SJTs' criterion-related validity and has extended improvement effect on employees' responses to actual work situations, such training may have broader influence on selection and training practices.

References

Ackerman, C. (2017). The big five personality theory: the five factor model explained. Retrieved from https://positivepsychologyprogram.com/big-five-personality-theory/

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*(2), 227-257. doi:10.1016/S0160-2896(96)90016-1

Ashton, M. C., Lee, K., & de Vries, R. E. (2014). The HEXACO Honesty-Humility, agreeableness, and emotionality factors: A review of research and theory. *Personality and Social Psychology Review*, *18*(2), 139–152. https://doi.org/10.1177/1088868314523838

Becker, T. E. (1998). Integrity in organizations: Beyond honesty and conscientiousness. *The Academy of Management Review*, *23*(1), 154–161. https://doi.org/10.2307/259104

Becker, T. E. (2005). Development and Validation of a Situational Judgment Test of Employee Integrity. *International Journal of Selection & Assessment*, *13*(3), 225-232. doi:10.1111/j.1468-2389.2005.00319.x

Catano, V. M., Brochu, A., & Lamerson, C. D. (2012). Assessing the reliability of situational judgment tests used in high-stakes situations. *International Journal of Selection and Assessment*, *20*(3), 333-346. doi:10.1111/j.1468-2389.2012.00604.x

Catano, V. M., O'Keefe, D. F., Francis, R. E., & Owens, S. M. (2018). *Construct-Based Approach to Developing a Short, Personality-Based Measure of Integrity*. Unpublished Manuscript.

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*(1), 83–117. https://doi.org/10.1111/j.1744-6570.2009.01163.x

Costa, P. T., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*(2), 322–331. doi:10.1037/0022-3514.81.2.322

Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, *43*, 52-64. doi:10.1016/j.intell.2014.01.004

Cullen, M. J., Sackett, P. R., & Lievens, F. (2006). Threats to the Operational Use of Situational Judgment Tests in the College Admission Process. *International Journal of Selection & Assessment*, *14*(2), 142-155. doi:10.1111/j.1468-2389.2006.00340.x

Fine, S., Goldenberg, J., & Noam, Y. (2016). Beware of those left behind: Counterproductive work behaviors among nonpromoted employees and the moderating effect of integrity. *Journal of Applied Psychology*, *101*(12), 1721–1729. doi:10.1037/apl0000156

Fine, S., Horowitz, I., Weigler, H., & Basis, L. (2010). Is good character good enough? The effects of situational variables on the relationship between integrity and counterproductive work behaviors. *Human Resource Management Review*, *20*(1), 73–84. doi:10.1016/j.hrmr.2009.03.010

Goldberg, L. R. (1992). The development of markers for the Big-Five factor
     structure. *Psychological Assessment*, *4*(1), 26-42. doi:10.1037/1040-3590.4.1.26

Harper, D. (1990). Spotlight abuse-save profits. *Industrial distribution*, *79*(3), 47-51.

International Personality Item Pool. (2017a). Administering IPIP measures, with a 50-
     item sample questionnaire. Retrieved from https://ipip.ori.org/New_IPIP-50-item-
     scale.htm

International Personality Item Pool. (2017b). Interpreting Individual IPIP Scale Scores.
     Retrieved from https://ipip.ori.org/InterpretingIndividualIPIPScaleScores.htm

John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History,
     measurement, and theoretical perspectives. In L. A. Pervin, O. P. John, L. A.
     Pervin, O. P. John (Eds.), *Handbook of personality: Theory and research, 2nd
     ed* (pp. 102-138). New York, NY, US: Guilford Press.

John, O. P., Robins, R. W., & Pervin, L. A. (2008). *Handbook of personality: Theory and
     research* (3rd ed.). New York: Guilford Press.

Johnson, J. A. (2014). Measuring thirty facets of the five factor model with a 120-item
     public domain inventory: Development of the IPIP-NEO-120. *Journal of
     Research in Personality*, *51*, 78-89. doi:10.1016/j.jrp.2014.05.003

Kleinmann, M. (1997). Transparenz der Anforderungsdimensionen: Ein Moderator der
     Konstrukt- und Kriteriumsvalidität des Assessment-Centers = Transparency of the
     required dimensions: A moderator of assessment centers' construct and criterion
     validity. *Zeitschrift Für Arbeits- Und Organisationspsychologie*, *41*(4), 171–181.
     Retrieved from
     http://library.smu.ca:2048/login?url=http://search.ebscohost.com/login.aspx?direct
     =true&db=psyh&AN=1997-43211-001&site=ehost-live

Kleinmann, M., Ingold, P. V., Lievens, F., Jansen, A., Melchers, K. G., & König, C. J.

    (2011). A different look at why selection procedures work: The role of candidates'

    ability to identify criteria. In G. J. Boyle, J. G. O'Gorman, & G. J. Fogarty

    (Eds.), *Work and organisational psychology: Research methodology; Assessment*

    and selection; Organisational change and development; Human resource and

    performance management; Emerging trends:

    Innovation/globalisation/technology., Vols. 1-5. (pp. 289–310). Thousand Oaks,

    CA: Sage Publications, Inc. Retrieved from

    http://library.smu.ca:2048/login?url=http://search.ebscohost.com/login.aspx?direct

    =true&db=psyh&AN=2016-48704-031&site=ehost-live

Krumm, S., Lievens, F., Hüffmeier, J., Lipnevich, A. A., Bendels, H., & Hertel, G.

    (2015). How 'situational' is judgment in situational judgment tests?. *Journal of*

    Applied Psychology, *100*(2), 399-416. doi:10.1037/a0037674

Lee, K., Ashton, M. C., & de Vries, R. E. (2005). Predicting Workplace Delinquency and

    Integrity with the HEXACO and Five-Factor Models of Personality

    Structure. *Human Performance*, *18*(2), 179–197.

    doi:10.1207/s15327043hup1802pass:[_]4

Lievens, F., Buyse, T., Sackett, P. R., & Connelly, B. S. (2012). The effects of coaching

    on situational judgment tests in high-stakes selection. *International Journal of*

    Selection and Assessment, *20*(3), 272–282. https://doi.org/10.1111/j.1468-

    2389.2012.00599.x

Lievens, F., & Motowidlo, S. J. (2016). Situational judgment tests: From measures of

    situational judgment to measures of general domain knowledge. *Industrial and*

Organizational Psychology: Perspectives on Science and Practice, *9*(1), 3-22. doi:10.1017/iop.2015.71

Lievens, F., & Sackett, P. R. (2006). Video-based versus written situational judgment tests: A comparison in terms of predictive validity. *Journal of Applied Psychology*, *91*(5), 1181-1188. doi:10.1037/0021-9010.91.5.1181

Lounsbury, J. W., Sundstrom, E., Loveland, J. M., & Gibson, L. W. (2003). Intelligence, 'Big Five' personality traits, and work drive as predictors of course grade. *Personality and Individual Differences*, *35*(6), 1231-1239. doi:10.1016/S0191-8869(02)00330-6

Maurer, T. J., Solamon, J. M., Andrews, K. D., & Troxtel, D. D. (2001). Interviewee coaching, preparation strategies, and response strategies in relation to performance in situational employment interviews: An extension of Maurer, Solamon, and Troxtel (1998). *Journal of Applied Psychology*, *86*(4), 709-717. doi:10.1037/0021-9010.86.4.709

Maurer, T., Solamon, J., & Troxtel, D. (1998). Relationship of coaching with performance in situational employment interviews. *Journal of Applied Psychology*, *83*(1), 128-136. doi:10.1037/0021-9010.83.1.128

McCrae, R. R., & Costa, P. J. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins, J. S. Wiggins (Eds.), *The five-factor model of personality: Theoretical perspectives* (pp. 51-87). New York, NY, US: Guilford Press.

McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, *9*(1-2), 103-113. doi:10.1111/1468-2389.00167

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. I. (2007). Situational

    judgment tests, response instructions, and validity: A meta-analysis. *Personnel*

    Psychology, *60*(1), 63-91. doi:10.1111/j.1744-6570.2007.00065.x

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*(4), 730-740.

    doi:10.1037/0021-9010.86.4.730

McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment tests: A review of

    practice and constructs assessed. *International Journal of Selection and*

    Assessment, *9*(1-2), 103-113.

Motowidlo, S. J. (2003). Job performance. In W. C. Borman, D. R. Ilgen, & R. J.

    Klimoski (Eds.), *Comprehensive handbook of psychology: Vol. 12. Industrial and*

    organizational psychology (pp. 39–53). New York: Wiley.

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*(2), 321-333. doi:10.1037/a0017975

Motowidlo, S. J., Crook, A. E., Kell, H. J., & Naemi, B. (2009). Measuring procedural

    knowledge more simply with a single-response situational judgment test. *Journal*

    of Business and Psychology, 24, 281–288. doi:10.1007/s10869-009-9106-4

Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006). Implicit policies about

    relations between personality traits and behavioural effectiveness in situational

    judgment items. *Journal of Applied Psychology, 91,* 749– 761. doi:10.1037/0021-

    9010.91.4.749

Mumford, S. (2015). *The Situational Judgment Test: Cognition, Constructs and Criterion Validity* (Doctoral dissertation). Retrieved from http://etheses.whiterose.ac.uk/9597/1/S%20Mumford%20Thesis%20July%20201 5.pdf

Anderson, N. (1994). Honesty in the Workplace. *Journal of Occupational & Organizational Psychology*, *67*(3), 279–280. Retrieved from http://search.ebscohost.com.library.smu.ca:2048/login.aspx?direct=true&db=aph &AN=9501231853&site=ehost-live

Neuman, G. A., & Baydoun, R. (1998). An empirical examination of overt and covert integrity tests. *Journal of Business and Psychology*, *13*(1), 65–79. doi:10.1023/A:1022971016454

Ones, D. S., Schmidt, F. L., & Viswesvaran, C. (1994). Correlates of preemployment integrity tests. In *23rd International Congress of Applied Psychology, Madrid, Spain*.

Sackett, P. R., & Devore, C. J. (2001). Counterproductive behaviours at work. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran. (Eds.), *Handbook of industrial, work & organizational psychology: Volume 1: Personnel psychology*. (pp. 145-164). Sage.

Sackett, P. R., & Wanek, J. E. (1996). New developments in the use of measures of honesty, integrity, conscientiousness, dependability, trustworthiness, and reliability for personnel selection. *Personnel Psychology*, *49*(4), 787–829. doi:10.1111/j.1744-6570.1996.tb02450.x

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of

research findings. *Psychological Bulletin*, *124*(2), 262-274. doi:10.1037/0033-2909.124.2.262

Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology*, *78*(6), 966–974. doi:10.1037/0021-9010.78.6.966

Smith-Jentsch, K. A., Salas, E., & Brannick, M. T. (2001). To transfer or not to transfer? Investigating the combined effects of trainee characteristics, team leader support, and team climate. *Journal of applied psychology*, *86*(2), 279.

Stemig, M. S., Sackett, P. R., & Lievens, F. (2015). Effects of organizationally endorsed coaching on performance and validity of situational judgment tests. *International Journal of Selection and Assessment*, *23*(2), 174–181. doi:10.1111/ijsa.12105

Thompson, N. (2015). Work out how likely an assessment result could have been achieved by chance. Retrieved from https://en.commtap.org/additional-resource/work-out-how-likely-assessment-result-could-have-been-achieved-chance

Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: how good are single-item measures?. *Journal of applied Psychology*, *82*(2), 247.

Wells, J. T. (1999). A fistful of dollars. *Security Management*, *43*, 70-75.

Whetzel, D. L., McDaniel, M. A., & Nguyen, N. T. (2008). Subgroup differences in situational judgment test performance: A meta-analysis. *Human Performance*, *21*(3), 291-309. doi:10.1080/08959280802137820

Appendix A
Implicit Trait Policy Training Content

Today, the "Big Five" theory is a very popular way of how people describe personalities. In this theory, there are five primary factors: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to experience. When applying for jobs, knowing certain rules about how to behave properly at work may improve your competitiveness during recruitment and selection. Those rules can largely be summarized to expressions of certain personality traits. For example, be humble to listen to manager's suggestions, which can be summarized to an expression of the personality trait, Agreeableness. Below are details about each personality factor and the essential personality expression rules that might improve your competitiveness. Try your best to learn these rules as they may help you to get high scores in employment tests.

Agreeableness:
What is it?
This trait briefly refers to individuals' tendency in getting along with others. This spectrum ranges from prosocial and communal orientation to antagonism. For example, individuals that are high in Agreeableness tend to be sympathetic, kind, humble, patient, loyal, amiable, etc. In contrast, individuals that are low in Agreeableness tend to be callous and antagonistic.
What would high-performing employees tend to do?
High-performing employees tend to express moderate Agreeableness at work. They are humble and patient in general. However, the degree to which they are kind and loyal to other people depends on what type of issues are involved in a situation. When things are going well, they are nice and supportive. When knowing that someone might be harmed or is facing a hardship, they would be sympathetic and amiable. However, when conflicts or different opinions happen, they do not agree with others just because they want to be friendly and humble or to avoid conflicts. These employees would bring up disagreement and be ready to maintain their own point of view if necessary. More importantly, in some difficult situations, where there are conflicts between Agreeableness and Conscientiousness, these employees will always choose to be Conscientious rather than being Agreeable. Many employers like this type of employees because they go along with other people but also, they would bring up issues that need to be solved.
Behavioural Example
In an example of a high-performing employee, Lauren is working in a small team. Recently, a new employee, Rebecca, joined her company and was assigned to her team. However, probably because all the other team members are in this company for at least 6 years, Lauren has noticed that Rebecca is experiencing a discrimination. In recent meetings, whenever Rebecca proposed her ideas, other members immediately denied her ideas. Rebecca looks quite embarrassed and frustrated. Laura often saw her crying at her own desk. One day during a meeting, Rebecca proposed her ideas and was denied again, but it was actually a good idea. The situation was obviously unfair. Lauren explained to the other members that why Rebecca's idea was a good one. One of the other member immediately became irritated and spoke aggressively toward Rebecca. Lauren then

brought up her concern of discrimination in recent days and suggested that everyone in this team should be respected whether or not he or she was new to the team.

Conscientiousness:
What is it?
This trait briefly describes individuals' tendency to have good impulse control and behave toward goals in socially acceptable way. For example, individuals that are high in Conscientiousness tend to be persistent, self-disciplined, reliable, hard-working, and consistent; they obey norms and rules, delay gratification, think before acting, plan and organize, etc. In contrast, individuals that are low in Conscientiousness tend to procrastinate and lose control of their impulses.
What would high-performing employees tend to do?
High-performing employees tend to express <u>high Conscientiousness</u> at work. In other words, they are honest and hard workers that always do what they are supposed to do. They are also responsible and care about other people. Unlike lazy workers, who do not care about their negative impact, the high-performing employees put efforts and follow the rules of workplace even though no one is overseeing them. They are good at controlling impulses, avoid temptations, and delaying gratification toward work goals. As well, they are organized and good at planning and prioritizing work tasks. Also, as noted previously, in some difficult situations, where there are conflicts between Agreeableness and Conscientiousness, these employees will always choose to be <u>Conscientious rather than being Agreeable</u>. Many employers prefer this type of employees because they usually follow rules and work hard.
Behavioural Example
In an example of a high-performing employee, Tom is given several tasks to do in a typical work day. A few of them are quite interesting, but others are very boring. However, he puts more effort into the tasks that are mostly important no matter whether they are interesting or not.

Neuroticism:
What is it?
This trait briefly describes individuals' emotional stability and even-temperedness. For example, individuals that are high in Neuroticism are more easily to be nervous, moody, anxious, fearful, sensitive, unstable, etc. In contrast, individuals that are low in Neuroticism are more certain, confident, and adventurous.
What would high-performing employees tend to do?
High-performing employees tend to express <u>low Neuroticism</u> at work. In other words, they maintain a stable temper at work. When encountering problems, they tend to behave calmly and confidently in solving problems rather than having an emotional burst. As well, in front of conflicts, their decisions are usually rational rather than self-defensive, but they are not afraid of expressing opinions and concerns. Also, these employees are okay to admit their mistakes.
Many employers favor this type of employees because they would remain calm and solve the problem it if occurs.
Behavioural Example

In an example of a high-performing employee, Nicole is working on a project plan that her team is trying to settle. During the meeting, she advocates for the plan that she thinks would work best, but the other team members advocate for a very different plan. Nicole calmly expressed her opinions and concerns toward the project. Then, instead of being emotional and irritated, she listened patiently when other members were presenting.

Extraversion:
What is it?
This trait, as self-explanatory as its name, refers to the spectrum of Extraversion and introversion. It describes how energetic individuals interact with the social and material world. For example, individuals that are high in Extraversion are more sociable, assertive, outgoing, talkative, socially confident, etc. In contrast, individuals that are high in introversion (i.e., low in Extraversion) tend to be quiet, reserved, introspective, and thoughtful.
What would high-performing employees tend to do?
High-performing employees tend to express high Extraversion. They are assertive and socially confident. When necessary, they talk about their concerns and persuade others toward their point of view. Many employers prefer this type of employees because they are easy to communicate with and would contribute to a workplace with a joyful atmosphere.
Behavioural Example
In an example of high-performing employee, James was making coffee in the office's kitchen. He met Laura, who came to heat her sandwich. They started to chat. James shared a funny experience that he had in this morning. Although the experience seemed a little bit embarrassing, he was relaxed and easy with it. He made Laura laugh and also laughed himself as talking about a funny story.

Openness to experience:
This trait refers to individuals' mental life and experiences in terms of depth, breadth, complexity, and originality. Individuals that are high in Openness to experience tend to be imaginative, daring, creative, curious, and are willing to try new things. In contrast, individuals that are low in Openness to experience tend to be more conservative.
What would high-performing employees tend to do?
High-performing employees tend to express high Openness to experience. As its self-explanatory name, this type of employees has open mind to new things. For example, they are able to think outside the box to find the best solution to problems. As well, unlike closed-minded employees, they are willing to consider and embrace different ideas, even though previously they viewed their own ideas as the best. Also, when change happens, they are easy to accept it. Many employers want this type of employees because they are creative and flexible at work.
Behavioural Example
In an example of high-performing employee, Kevin is an experienced technician for maintaining printers in a company. He has been doing well in this job for the same company for 12 years. One day a printer stopped working, he did trouble shooting and tried everything he knew to fix the problem. However, the printer would not start working no matter what he tried. Such a problem had never happened before, and Kevin had

always been able to fix printer's issues on his own. One of his colleague, Maggie, who is a secretary, suggested that he might need to contact the manufactory to see what to do. Kevin had never considered this opinion before. To be honest, he was quite confident on how much he knew about printers and it was a bit frustrating that he was encountering such a dubious problem. However, he considered this option and admitted that it might actually be a smart thing to try. Thus, instead of insisting on his own expertise, he took the secretary's suggestion and called the manufactory.

Appendix B
Control Task Content

When applying for jobs, knowing aspects about personality may improve your competitiveness during recruitment and selection because many employers have become interested in knowing candidates' personalities before making hiring decisions. However, what is personality? Today, the "Big Five" theory is the most popular basis of how people describe personalities. Learn more about this theory may help you prepare for employment tests and interviews. Here we have prepared an easy-to-digest material for you. In the "Big Five" theory, there are five primary factors of personality: Agreeableness, Conscientiousness, Neuroticism, Extraversion, and Openness to experience. Below are more details about each of the five factors.

Agreeableness:
What is it?
This trait briefly refers to individuals' tendency in getting along with others. This spectrum ranges from prosocial and communal orientation to antagonism. For example, individuals that are high in Agreeableness tend to be sympathetic, kind, humble, patient, loyal, amiable, etc. In contrast, individuals that are low in Agreeableness tend to be callous and antagonistic.
Behavioural Example
Rebecca is low in Agreeableness. She always has a hard time understanding the views that other people have. She sometimes feels irritated of how people are different.
Lauren is high in Agreeableness, she is kind and always goes along with others.

Conscientiousness:
What is it?
This trait briefly describes individuals' tendency to have good impulse control and behave toward goals in socially acceptable way. For example, individuals that are high in Conscientiousness tend to be persistent, self-disciplined, reliable, hard-working, and consistent; they obey norms and rules, delay gratification, think before acting, plan and organize, etc. In contrast, individuals that are low in Conscientiousness tend to procrastinate and lose control of their impulses.
Behavioural Example
Tom is low in Conscientiousness. He does not always follow the rules, and he often decides to do things that he personally thinks great.
Carrie is high in Conscientiousness. She is always hard working, and she tries to maintain her responsibility.

Neuroticism:
What is it?
This trait briefly describes individuals' emotional stability and even-temperedness. For example, individuals that are high in Neuroticism are more easily to be nervous, moody, anxious, fearful, sensitive, unstable, etc. In contrast, individuals that are low in Neuroticism are more certain, confident, and adventurous.
Behavioural Example

Mark is low in Neuroticism. He is always confident about himself and his ideas. He loves sharing different ideas to his work team even though sometimes he receives criticism. Nicole is high in Neuroticism. She often concerns about her work. For example, she sometimes worries that her small mistakes might have negative impacts on her company.

Extraversion:
What is it?
This trait, as self-explanatory as its name, refers to the spectrum of Extraversion and introversion. It describes how energetic individuals interact with the social and material world. For example, individuals that are high in Extraversion are more sociable, assertive, outgoing, talkative, socially confident, etc. In contrast, individuals that are high in introversion (i.e., low in Extraversion) tend to be quiet, reserved, introspective, and thoughtful.
Behavioural Example
James is quite introverted. He spends most of his time working quietly at his office desk. He prefers to keep his feelings inside.
Kyle is quite extraverted. He loves talking, even with someone he meets at the subway station.

Openness to experience:
This trait refers to individuals' mental life and experiences in terms of depth, breadth, complexity, and originality. Individuals that are high in Openness to experience tend to be imaginative, daring, creative, curious, and are willing to try new things. In contrast, individuals that are low in Openness to experience tend to be more conservative.
Behavioural Example
Kevin is low in Openness to experience. He prefers to follow the tradition and avoid risk taking.
Maggie is high in Openness to experience. She loves trying new things and always embraces new ideas at work.

Appendix C
Situational Judgement Test

Instruction:
Now, pretend you are attending a job interview. Before the interview, you are given a paper test to do. This type of test is quite popular in HR practices. Here is the test, please answer the following questions by selecting the best answer. Remember, you want to do it well so that you would be selected to the job. By the way, the material you have just read and the associated examples were to help you do well in this type of test. Feel free to apply what you have learned when choosing your answers.

Nineteen SJT items were then presented following this instruction.

Note. The original content of Becker's SJT can be found in "Appendix: Integrity Scenarios" (Becker, 2005, p. 229). The original items were written in a work or business context, which might be hard for students to associate with. Thus, to help participants of the current study easily understand and respond to the SJT, we presented this test with minor modifications to represent a university context. For example, in one of the items, instead of presenting "You are an architect, and have been asked to work with a group of three other architects to design a new building" from the original test, we presented the situation as "You and three other students in your engineering class are working on a project to design a new building". In another example, instead of "You're retiring from a successful business that you started", we presented "Your term as president of Student Council is coming to an end". However, the modifications were only contextual, the dilemma or conflicts presented by the original test was not changed.

Appendix D
Manipulation Check

Remember the materials you have read before. In that material, you have learned about the basics of the "Big Five" theory. Now, think about the five personality traits and choose the best answer for each of the following questions. Remember, try to answer the questions <u>based on what you have read</u> in the previous material.

Question 1. High-performing employees tend to express _____ at work.
A. High Agreeableness
B. Moderate Agreeableness
C. Low Agreeableness

Question 2. High-performing employees tend to express _____ at work.
A. High Conscientiousness
B. Moderate Conscientiousness
C. Low Conscientiousness

Question 3. High-performing employees tend to express _____ at work.
A. High Neuroticism
B. Moderate Neuroticism
C. Low Neuroticism

Question 4. High-performing employees tend to express _____ at work.
A. High Extraversion
B. Moderate Extraversion
C. Low Extraversion

Question 5. High-performing employees tend to express _____ at work.
A. High Openness to experience
B. Moderate Openness to experience
C. Low Openness to experience

Answer key:
Question 1. B
Question 2. A
Question 3. C
Question 4. A
Question 5. A

Appendix E
Integrity Test

Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Indicate for each statement whether it is 1. Very Inaccurate, 2. Moderately Inaccurate, 3. Neither Accurate nor Inaccurate, 4. Moderately Accurate, or 5. Very Accurate as a description of you.

Ten TSDI items were then presented following this instruction.

Note. The original TSDI items can be found in "Table 1" of Catano, et al. (2018, p. 48).

Appendix F
IPIP Personality Measure

Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Indicate for each statement whether it is 1. Very Inaccurate, 2. Moderately Inaccurate, 3. Neither Accurate nor Inaccurate, 4. Moderately Accurate, or 5. Very Accurate as a description of you.

Fifty IPIP items were then presented following this instruction.

Note. The original IPIP items can be found in the section "How Accurately Can You Describe Yourself?" (International Personality Item Pool, 2017a, p. 10).

Appendix G
The International Cognitive Ability Resource (ICAR) Measure

Now, you have a 16-question test. You have 10 minutes (maximum) to complete it. After 10 minutes, you will be moved to the next page automatically. Remember to answer all of the questions. If you are not sure, trust your intuition.

Sixteen ICAR items were then presented following this instruction.

Note. The original ICAR items can be found in "Appendix A: ICAR Sample Test" (Condon & Revelle, 2014, Supplementary Materials p. 2).