



## Full Length Article

## False consensus in situational judgment tests: What would others do?

Janneke K. Oostrom<sup>a,\*</sup>, Nils C. Köbis<sup>b</sup>, Richard Ronay<sup>a</sup>, Myckel Cremers<sup>c</sup><sup>a</sup> Vrije Universiteit Amsterdam, The Netherlands<sup>b</sup> University of Amsterdam, The Netherlands<sup>c</sup> Dutch Ministry of Defense, The Hague, The Netherlands

## ARTICLE INFO

## Article history:

Received 30 March 2017

Revised 25 August 2017

Accepted 2 September 2017

Available online 7 September 2017

## Keywords:

Situational judgment test

Response instruction

Faking

False consensus

Validity

## ABSTRACT

We introduce an alternative response instruction to reduce the fakability of situational judgment tests. This novel instruction is based on the false consensus effect, a robust social psychological bias whereby people infer that the majority of other people's thoughts, attitudes, and behaviors are aligned with their own. In four studies, including both field and laboratory data (total  $N = 882$ ), we demonstrate that participants show a false consensus bias when asked what others would do in situational judgment tests. Furthermore, the situational judgment test based on the false consensus effect turned out to be relatively difficult to be fake, and produced scores that were meaningfully correlated with conceptually related traits, as well as both self-reported and behavioral outcomes.

© 2017 Elsevier Inc. All rights reserved.

## 1. Introduction

Situational judgment tests (SJTs) have been used for employee selection for about 90 years (e.g., McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001; Moss, 1926) and have become increasingly popular in research and practice over the past decades (Weekley & Ployhart, 2005). A SJT is a measurement method typically composed of a series of job dilemmas, followed by a list of plausible response options. Candidates are asked to evaluate each response option for either the likelihood that they would respond in that particular manner, or the effectiveness of the response, and are often instructed to either rate each of the response options or to pick the most and/or least adequate option (e.g., O'Connell et al., 2007; Oostrom, De Soete, & Lievens, 2015). HR professionals then assess the degree to which the candidate's responses map onto the kinds of knowledge, skills, abilities, or traits they are interested in. In general, the literature has supported the criterion-related validity of SJTs (e.g., Christian, Edwards, & Bradley, 2010). For instance, in a meta-analysis, McDaniel, Hartman, Whetzel, and Grubb (2007) found SJTs to have an average observed validity of 0.20 for predicting job performance. Furthermore, research has demonstrated that SJTs show smaller ethnic

score differences than cognitive measures (e.g., Chan & Schmitt, 1997; Whetzel & McDaniel, 2009) and have high face and content validity (Lievens, Peeters, & Schollaert, 2008), making SJTs an attractive selection tool.

Despite their popularity and criterion-related validity, SJTs have a clear limitation: they are easy to fake by candidates in high-stakes selection settings. For instance, Peeters and Lievens (2005) conducted a between-subjects study on the fakability of SJTs and found that candidates in the faking condition scored 0.89 *SD* higher than candidates in the honest condition. Faking on a selection test can be defined as a candidate's conscious distortion of their answers to score more favorably (e.g., McFarland & Ryan, 2000). Although there is an ongoing debate whether faking influences a selection test's criterion-related validity (e.g., Hough, 1998; Ones & Viswesvaran, 1998), researchers do agree that faking can have a significant effect on the selection decisions that are based on individual test scores. Indeed, Peeters and Lievens showed in their SJT study that, with a selection ratio of 0.25, an organization would largely hire fakers (i.e., 76%) – a potentially costly decision, given the negative consequences of work-place dishonesty (Ariely, 2012).

A seemingly easy way to prevent faking on SJTs is to change the response instruction. There are two common types of response instructions: should-do (i.e., knowledge-based) and would-do (i.e., behavioral tendency) instructions (McDaniel & Nguyen, 2001). Should-do response instructions ask the candidate to identify the best or correct course of action. Would-do response

\* Corresponding author at: Vrije Universiteit Amsterdam, Department of Management & Organization, School of Business and Economics, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.

E-mail address: [j.k.oostrom@vu.nl](mailto:j.k.oostrom@vu.nl) (J.K. Oostrom).

instructions ask the candidate to indicate how he or she would likely behave (McDaniel et al., 2007). Nguyen, Biderman, and McDaniel (2005) found that candidates can easily distort their answers on a would-do SJT. However, the results for the should-do SJT were inconsistent, due to the difficulty to fake knowledge. Faking even led to lower scores when candidates first answered honestly because they had already responded to the best of their ability the first time they were presented with the job dilemmas.

However, changing the response instruction to a should-do SJT is not possible without changing the constructs that are being measured with the SJT. Indeed, Ployhart and Ehrhart (2003) showed that construct validity is “dramatically affected by the type of instructions” (p. 11). In general, should-do SJTs tap more into ability and knowledge related constructs and would-do SJTs tap more into attitudes and personality related constructs (McDaniel et al., 2007). The question then is: How can we measure personality related constructs in a manner that is resistant to faking? In the present study, we offer a potential answer to this question by introducing an alternative response instruction based on a well-known social psychological phenomenon – the false consensus effect (Ross, Greene, & House, 1977). The false consensus approach would offer an alternative personality assessment, which seems to be particularly useful in situations when candidates have a strong tendency to present themselves in socially desirable ways (Alliger & Dwight, 2000; Becker, 1998).

Although we position this research within the context of employee selection, any assessment of personality and attitudes, be it in the service of theory testing or in the pursuit of applied goals, invites a consideration of how to manage the potentially confounding effects of socially desirable responding (Greenwald & Banaji, 1995; Greenwald et al., 2002). Because self-report measures, such as typical SJTs, are based on information that is introspectively accessible to respondents at the time of measurement, responses might reflect information about the person that emerges in response to self-presentation concerns, rather than from the construct under investigation. Implicit measures, such as we position our alternative response instruction, are based on information that is not intentionally given as self-informant (Uhlmann et al., 2012) and as such should appeal to researchers and practitioners interested in investigating personality, attitudes, and beliefs, in the absence of such self-presentation concerns.

### 1.1. False consensus effect: origins and theoretical perspectives

Already in 1931, Katz and Allport observed that students who admitted to having cheated on exams expected other students to have cheated too. This phenomenon of overestimating the percentage of others in the population who share one's own characteristic (s), labeled the “false consensus effect” (Ross et al., 1977), has been demonstrated for numerous abilities, beliefs, and traits (see Mullen et al., 1985 for a meta-analysis). Ross et al. (1977) defined the false consensus effect as people's tendency to “see their own behavioral choices and judgments as relatively common and appropriate to existing circumstances while viewing alternative responses as uncommon, deviant, or inappropriate” (Ross et al., 1977, p. 280).

There are several theoretical accounts of the false consensus effect. Marks and Miller (1987) grouped these into four perspectives. The first, *selective exposure and availability*, suggests that perceptions of similarity are affected by the ease with which instances of similarity between self and others come to mind (Tversky & Kahneman, 1973). Such instances are readily available because people typically associate with others who are more similar than dissimilar to themselves (Berscheid & Walster, 1978; Bishop, 2008). The second, *salience and focus of attention*, suggests that consensus arises from a focus of attention on one's preferred position,

which then becomes the only position in immediate consciousness (Marks & Miller, 1987). The third, *logical information processing*, suggests that active reasoning and rational processes underlie perceptions of similarity. People consider themselves and others to be similarly rational beings who are affected by the situation in the same manner. From this perspective, the false consensus effect is thus a manifestation of the tendency of people to attribute behavior to situational forces (Gilovich, Jennings, & Jennings, 1983). The fourth is the perspective of *motivation*, which suggests that perceiving similarity has functional value – increasing perceived social support, validating the correctness of positions, maintaining self-esteem, maintaining or restoring cognitive balance, or reducing tension because of anticipated social interactions. Marks and Miller (1987) concluded that there is evidence for each of these perspectives; many of the false consensus findings are open to multiple interpretations and the proposed mechanisms often overlap or operate simultaneously or in concert.

Even though the exact underlying mechanisms remain unclear, what is clear is that people have a strong tendency to base their estimates of others' characteristics on their own characteristics (Marks & Miller, 1987; Mullen et al., 1985). This tendency is very robust; persisting even when people are educated about the bias, or about actual consensus rates surrounding the issue in question (Alicke & Largo, 1995; Krueger & Clement, 1994). For example, Krueger and Clement (1994) informed participants just before making their prevalence estimates that such estimates are often biased by respondents' own characteristics, and still found no reduction in false consensus. Thus, when people are asked about others' characteristics, they seem unable to avoid revealing information about themselves, even when aware of the phenomenon they then exhibit.

### 1.2. SJTs as measures of individual differences through false consensus

In most social psychology experiments, the false consensus effect is measured at the group level; the *standard test* of the false consensus effect is whether the mean consensus estimate provided by people who endorse an item is greater than the mean estimate provided by those who do not endorse the item (e.g., Ross et al., 1977). However, with a *SJT based test* of the false consensus effect – one that instructs candidates to choose the response option that reflects what *other people* would do, instead of what *they themselves* would do (from hereon called ‘FC SJT’) – we move from group level to individual level measurement, with the potential for inferring traits and predicting behavior.

There have been a few earlier attempts at operationalizing false consensus at the individual level (see De la Haye, 2000; Krueger & Clement, 1994). So far, individual level measures of the false consensus effect have been used to confirm the existence of the effect, and to explore the boundary conditions surrounding it. To our knowledge, using the paradigm to directly assess individual level personality with the intention to predict future (job) behavior, has been limited to integrity tests such as the Reid Report (Ash, 1971; Cunningham, Wong, & Barbee, 1994).

We reasoned that the FC SJT is an implicit measure of personality; a measure that minimizes candidate's awareness of what is being measured and/or their ability to control their responses (Uhlmann et al., 2012). Implicit measures are especially useful in evaluative situations in which participants are unwilling to admit their attitudes to others (Sackett & Lievens, 2008; Uhlmann et al., 2012). In the personality domain, several promising implicit measures have been introduced to address the faking concerns of self-report measures (McDaniel, Beier, Perkins, Goggin, & Frankel, 2009).

Our idea of using FC response instructions to measure personality builds on the work by Motowidlo, Hooper, and Jackson (2006),

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات