

2022

## “Faking” is Neither Good Nor Bad, It Is a Misleading Concept: A Reply to Tett and Simonet (2021)

Bernd Marcus  
*University of Rostock*

Follow this and additional works at: <https://scholarworks.bgsu.edu/pad>

 Part of the [Human Resources Management Commons](#), [Industrial and Organizational Psychology Commons](#), and the [Other Psychology Commons](#)

[How does access to this work benefit you? Let us know!](#)

### Recommended Citation

Marcus, Bernd (2022) "“Faking” is Neither Good Nor Bad, It Is a Misleading Concept: A Reply to Tett and Simonet (2021)," *Personnel Assessment and Decisions*: Number 8 : Iss. 1 , Article 4.

DOI: <https://doi.org/10.25035/pad.2022.01.004>

Available at: <https://scholarworks.bgsu.edu/pad/vol8/iss1/4>

This Article is brought to you for free and open access by the Journals at ScholarWorks@BGSU. It has been accepted for inclusion in Personnel Assessment and Decisions by an authorized editor of ScholarWorks@BGSU.

# “FAKING” IS NEITHER GOOD NOR BAD, IT IS A MISLEADING CONCEPT: A REPLY TO TETT AND SIMONET (2021)

Bernd Marcus<sup>1</sup>

1. University of Rostock

## ABSTRACT

### KEYWORDS

faking, self-presentation,  
personality testing,  
personnel selection

This paper comments on Tett and Simonet’s (2021) outline of two contradictory positions on job applicants’ self-presentation on personality tests labelled “faking is bad” (FIB) versus “faking is good” (FIG). Based on self-presentation theory (Marcus, 2009) Tett and Simonet assigned to their FIG camp, I develop the ideas of (a) understanding self-presentation from the applicant’s rather than the employer’s perspective, (b) avoiding premature moral judgment on this behavior, and (c) examining consequences for the validity of applicant responses with a focus on the intended use for, and the competitive context of, selection. Conclusions include (a) that self-presentation is motivationally and morally more complex than assumed by proponents of the FIB view; (b) that its consequences for validity are ambivalent, which implies that simple credos like “FIB” or “FIG” are equally unjustified; and (c) that the label “faking” shall be abandoned from the scientific inquiry on the phenomena at hand, as it contributes to prejudiced and often erroneous conclusions.

In a recent position paper, Tett and Simonet (2021; hereafter referred to as T&S) created two “camps” of scholars of self-presentation on personality tests in personnel selection. One position, exemplified by earlier writings of T&S themselves (e.g., Tett & Simonet, 2011), among others, was labelled “faking is bad” (FIB) and represents the traditional view of psychometricians on faking as systematic bias that impairs construct validity and thus shall be controlled for or eliminated. The second camp was labelled “faking is good” (FIG) and, in T&S’s interpretation, holds that faking either does not matter or even contribute positively to the validity of personality tests in personnel selection. Based on particular definitions of faking and validity I return to later, T&S eventually paint a dystopic picture of the future of personnel selection if the FIG camp succeeds in this controversy, which “will lead ultimately to the demise of self-report personality assessment in hiring settings” (p.13).

T&S ascribe the foundation of the FIG camp primarily to socioanalytic theory (Hogan & Holland, 2003; Johnson & Hogan, 2006) and to my own self-presentation theory (Marcus, 2009; Marcus et al., 2020, the latter paper being cited as Marcus et al., 2019, by T&S). They almost equate the two theories and they completely equate the construct

of analytical skills in self-presentation theory (hereafter referred to as SPT) with Kleinmann and colleagues’ (2011) “ability to identify criteria” (ATIC). Although I consider those analogies gross over-simplifications at best, I will primarily focus on SPT for developing a position that sharply contrasts both with T&S’s understanding of my theory and with their general position on the value, the meaning, and the implications of the concept of faking for understanding applicants’ self-presentation and its consequences. I first propose three rules I sophomorically consider “golden” for that purpose and then conclude with a call for abandoning the concept of faking (not the phenomena described by this word) from further scientific inquiry in this area of research.

Corresponding author:  
Bernd Marcus  
Author Email: bernd.marcus@uni-rostock.de

### Rule 1: If You Want To Understand Behavior, Look At it From the Actor's Perspective

One of the assumptions stated early in T&S's paper refers to the reason why organizations use personality tests in selection. Specifically, T&S assume that organizations "are making a good-faith attempt to distinguish among applicants with respect to their expected fit in meeting work demands" (p.7). Regardless of whether their assumption is correct (see my footnote 1 on that matter), this statement sets the cornerstone for the perspective the authors take throughout their paper, namely that of the employer. Consistent with this perspective, the entire paper seems (at least implicitly) to be based on the understanding that serving organizations to achieve this goal defines our primary role as scholars of personnel selection. In line with this perspective, T&S employ a definition of faking they adopted from Griffith et al. (2011). According to this definition, faking may take on four different forms, in descending order of deceptive intentionality from fraudulent to exaggerating to reactive response (i.e. meeting perceived expectations) to self-presentation (derived from socioanalytic theory as presenting one's desired reputation). Frankly, this ordinal definition of faking types reminds me of legal definitions of differing degrees of murder. I believe this association did not come to me by coincidence. Such conception of faking defines the roles of employer and applicant analogous to those of the judge and defendant in the courtroom (or maybe of the examiner and student in an exam).

By contrast, SPT rests on different definitions of the roles of scholars of personnel selection, as well as of the employer and the applicant. According to SPT, the first and foremost role of the scholar is *not* serving the interests or purposes of any of the parties involved but trying to understand what is going on in personnel selection. Serving either party's interests (including that of the applicant!) may well be a secondary goal of scientific inquiry, but this goal can only be achieved based on a proper understanding of what is going on in the first place. The adequate perspective of the scholar thus depends on what exactly we try to understand. Regardless of how we define faking or related concepts, it is the applicant's behavior to which such concepts refer. Hence, the only perspective that will take us closer to understanding this behavior is that of the applicant, not the employer.

Now, is the role of the applicant in selection adequately described analogous to that of a defendant in a courtroom or a student in an exam? Applicants may at times feel like that, but I still hold that the selection process in practice is not nearly as much guided by strict rules and regulations but rather by individual interests, mutual perceptions, and social norms that are subject to idiosyncratic interpretation. According to SPT, a useful metaphor for describing this situation is not that of a courtroom but rather that of a first

date candlelight dinner. Two interaction partners who do not know each other well meet to find out whether or not they want to begin a lasting relationship with each other. If they come to the conclusion that such relation is desirable, they then have to convince their prospective partner to arrive at the same conclusion about themselves. Those two tasks (selecting and attracting the potential partner) are complementary in that selecting in would not work without successful attraction, and attraction does not make sense with partners selected out. Moreover, the two partners are mutually dependent in their success, as one cannot initiate a relationship without consent of the other party. Finally, these tasks are perfectly equivalent for both partners. Although there may be an imbalance of power depending on what either side has to offer relative to other potential partners, the tasks are essentially the same for both partners, and roles are not distinguished between them such that one partner has more legitimate right to tell the other one what to do. Based on these definitions of roles, SPT suggests that applicants consider three things simultaneously, each of which has important implications for their behavior.

First, they need to find out what the employer expects from them, as this will guide any attempt to attract the employer. Success in this task depends on a formative construct I labelled analytical skills, which is similar, but not identical in meaning and especially in measurement, to Kleinmann et al.'s (2011) ATIC construct, as well as to the "ability to fake" proposed by several theorists T&S may assign to their FIB camp (e.g., Levashina & Campion, 2007; McFarland & Ryan, 2000, 2006). There may not be much of a fundamental controversy about this element of self-presentation, unless we turn to the moral evaluation of behavioral consequences, which I will pick up in the next section.

Second, according to SPT, applicants also consider their own needs and expectations, as well as their image of themselves. They will then judge the employer's attractiveness (i.e., the selection task from their perspective) based on perceptions of how far employment in the job offered will fulfill their needs and of the degree to which employer's expectations match their self-image. In my reading of the extant literature, this element is unique to SPT, and it leads to conclusions that are partially contrary to beliefs of the FIB camp, as outlined by T&S. For example, SPT proposes that completion of applicants' judgment based on the above considerations leads to what is labelled "informed motivation" (to self-present). Among other things, informed motivation is held to depend on the discrepancy between (honest) self-image and perceived employer's expectations such that larger discrepancies tend to *lower* informed motivation. By contrast, T&S, among other faking theorists, suggest that a conceptually similar discrepancy often called "opportunity to fake" (i.e. the discrepancy between an individual's honest score and the maximum score on a personality test) trans-

lates into *more* faking behavior. I shall note that opportunity to fake describes a potential, whereas faking describes the extent to which individuals make use of this potential. Concluding from the size of a potential to actual behavior creates a tautology, which is a serious theoretical flaw. T&S may nonetheless be right with their prediction, and the contradictory prediction of SPT may be wrong, which is an empirical question research has to answer. A recent study (Hummert et al., 2021) found that applicants with larger discrepancies between honest self-image and perceived ideals tended to restrict their engagement in self-presentation, thereby lending initial support to SPT's position on the matter.

Finally, applicants need to consider that they compete with other applicants for the job in question. Unfortunately for them, they typically possess little information on their competitive advantage or disadvantage. It is therefore wiser to assume that competitors are strong than to under-rate them. One consequence of competitiveness is that the psychometric logic of distinguishing between instruments designed to predict maximum performance (e.g., ability tests) and those designed to predict typical performance (e.g., personality tests) conflicts with a defining feature of the situation. In a competitive environment, all instruments used for making selection decisions call for maximum performance from the applicant's perspective (Marcus et al., 2020). Instructing applicants to switch from maximum performance in ability or work sample tests to typical behavior in personality tests is a bit like asking participants in the Olympic decathlon to show their best performance in some disciplines but to perform more like in a training session in others. You won't win the gold medal if you follow this instruction. This issue poses some problems on the interpretation of personality test scores and the conception of validity in this context, which I will also return to later.

To summarize, job applicants are faced with the situational demands to (a) analyze the employer's expectations, (b) assess the degree to which these expectations match one's own expectations and identity, and (c) meet these expectations in competition with other applicants. This already is a collection of complex and partially conflicting demands, but readers may miss the question of morality with which faking researchers are so concerned. This is what I turn to next.

### **Rule 2: If You Want To Understand Behavior, Look at It Without Prejudice**

It is about time to review the "faking" label. I have already noted that T&S adopted a definition of faking that covers almost every form of self-presentation. Of course, one may attach any label to a given construct space. Yet, labelling has real consequences, as every victim of discrim-

inatory language knows. The label "faking" is used as a technical term in the extant literature, but it is adopted from plain language. In everyday usage, "faking" unequivocally has a connotation of lying, fraud, and deceit (as in "fake news", for example). This connotation overlaps with the most extreme form of fraudulent faking in Griffith et al.'s (2011) typology, but *only* with this form. Fraud is a clear violation of widely accepted social norms, and it tends to trigger aversive emotional responses on the receiving end. No one wants to be lied at. If we'd ask a sample of employers whether they would want to hire applicants who are lying to them, we would probably see an endorsement rate close to zero.

By contrast, self-presentation is defined in SPT as the extent to which applicants are willing and able to adapt to the employer's expectations. If we ask the same sample of employers whether they would want to hire applicants willing and able to adapt to their expectations, I bet we would find a very different distribution of responses. This illustrates a mere labelling effect. The two labels have very different connotations but "faking" (as defined by Griffith et al., 2011) and "self-presentation" (as defined in SPT) are used in the scholarly literature to refer to largely overlapping forms of observable behaviors (see Levashina, 2018, for a broader discussion of the emergence of different forms of impression management in the extant literature). This raises the question which label and, more importantly, which effects implied by the use of respective labels are justified by what we are able to observe.

As researchers of either faking or self-presentation, we may design a study that allows us to observe, for example, that a test taker responds to an item on a conscientiousness scale (say, "I am always on time for appointments.") with "4" on a five-point endorsement scale in a neutral (or "honest," which is yet another value-laden label) condition, but shifts to "5" in an applicant setting. This is all we can observe directly (i.e., at the behavioral level). This observation per se does not tell us anything about differing intentions that underlie the forms specified in Griffith et al.'s (2011) definition of faking, not even whether these distinct forms actually exist. To understand what we observe, we need to infer the *meaning* of this response shift. Consider the following set of messages, each of which corresponds to one possible meaning of the observable response shift just described: (a) "I know I am sometimes late, but I won't tell you because otherwise you may not hire me"; (b) "I may at times be late for private appointments, but I am always on time for work"; (c) "I may not always have been punctual in my life, but if you hire me, I am willing to show up on time every day for a job like this." Message A refers to the logic of the faking paradigm; Message B corresponds to the logic underlying the contextualization of personality tests; and Message C illustrates what colleagues and I recently

referred to as the social meaning of item responses (Marcus et al., 2020). These different meanings have very different psychometric implications, as I will discuss in greater detail in the following section. The present section is concerned with moral implications, to which I turn now.

The key difference from a moral point of view is that Message A implies that, once hired, the applicant has no intention to actually show the behavior implied by his or her response (i.e. perfect punctuality), whereas both Message B and C imply that this behavior will actually be shown in the job question. I'd contend that this is the only thing the employer has a legitimate right to know when asking this question. Whether or not I show up on time anywhere else than for the job offered is not my employer's cup of tea. As there are possible meanings of the exact same observable act with completely contradictory implications, it is essentially unscientific to conclude one specific meaning simply from observing the behavior. However, as I had tried to demonstrate at the beginning of the section, this is essentially what labelling the observed behavior "faking" does. By attaching the faking label, it is *assumed* that the observed act is amoral<sup>1</sup>, whereas the label of self-presentation is open to all kinds of meanings and implications. It seemed quite evident to me that T&S fall in their own labelling trap at various places in their article (e.g., when they predict that masses of ruthless but skilled fakers will flood the job market after the FIG dystopia turned reality).

Following the path of science would dictate to infer the meaning of behavior either by deduction or induction. As the ambiguity of the behavior implies that simple deduction (logical reasoning) will not suffice in this case, researchers need to design studies allowing for more indirect or inductive inferences. This is typically done by correlating some measure of self-presentation or faking with outside variables. I will forego the manifold difficulties of measuring self-presentation independent of trait variance<sup>2</sup> and assume for now that we may draw some reliable conclusions at least on parts of the massive volume of research accumulated. T&S presented a selected review of this research, which I do not need to discuss at length, because I agree with them for the most part in that respect.

For my purpose here, two types of evidence are especially relevant. First, as reviewed by T&S, there is some evidence that a number of stable traits and abilities (e.g., self-monitoring, emotional intelligence, cognitive ability) correlate with self-presentation. Such findings shed some light on the psychological meaning of self-presentation, but the nomological net is still far from dense, results do not allow for distinguishing between types of self-presentation, and they certainly do not justify any moral judgment. A second stream of research either correlated measures of self-presentation with behaviors on the job or assessed the effect of self-presentation on the validity of personality tests for predicting job performance. I concur with T&S's

review that these studies yielded inconclusive results. I'd reiterate that work behavior and its prediction is what the employer is legitimately concerned with. Understanding the consequences of self-presentation for work behavior thus is key to clarifying the moral question as well. This is the issue I turn to next.

### **Rule 3: If You Want to Understand the Consequences of Behavior, Tell the Whole Story**

T&S present an understanding of (construct) validity they consider "unified" and contrast it with their interpretation of the FIG camp's understanding they consider "antiquated." According to T&S, FIG proponents return to the traditional threefold concept of validity in which construct validity is distinguished from criterion-related validity. T&S adopt their definition of validity from the *Standards* by the American Educational Research Association (AERA) as "support [for] the interpretation of test scores proposed by the test user ... [which] includes specifying the construct the test is intended to measure." (AERA 2014; p.11). T&S

1 Whereas T&S, like most faking researchers, are quick at assuming amoral intentions in applicants who alter their responses to personality items in selection, they seem to assume only the best intentions on part of the employer (see the "good-faith" quote on p.7 referred to earlier). This is an equally untested assumption, which I suspect may stem from defining the role of the scholar as serving the employer's interests. I doubt that this assumption is always correct. In line with my dating metaphor, both the applicant and the employer have reason to use self-presentation for their respective purposes. If the employer wants to recruit a particular applicant, or just wants to create a favorable image of the organization, there clearly is an incentive for positive self-presentation similar to the applicant's incentive. Moreover, it seems naïve to assume that all organizations unequivocally aim at hiring perfectly honest employees. On the contrary, it may well turn out that, if in conflict, the organization prefers loyalty over uncompromising honesty (e.g., in frontline employees dealing with customers' critical questions on the quality of products or services offered). Hence, the perfectly honest "non-faker" may turn out to be a misfit with the organization, yet not because of an unfitting general personality profile but just because of his or her uncompromising honesty. For the same reason, not all organizations are happy about whistleblowers driven by firm moral principles.

2 From a SPT perspective, one would ideally observe responses to each item under neutral and applicant conditions and then derive measures of self-presentation directly from observed within-person differences at the item level. Unfortunately, this is rarely possible in practice, especially with actual rather than simulated applicant conditions (see Hummert et al., 2021, for a realization in a high-fidelity simulation). It is more often possible to compare correlation coefficients between conditions (i.e., without measuring self-presentation as an individual-level variable), which leads to more indirect but in my view still potentially meaningful conclusions. I am much less enthusiastic about proxy measures of faking (e.g., desirability scales, response latencies, counts of "blatant extreme responding", etc.), but this topic belongs in the next section.

then go on by stating that “validity is the degree to which a test measures what it is purported to measure (i.e., a defined construct)” (p.8), which makes clear that they consider the word “construct” the key term in the AERA definition. Throughout their article, T&S also make clear that they understand “construct” completely independent of the situation and the purpose of test use. A conscientiousness scale is held to be valid to the degree it measures conscientiousness. Period. I must admit that I was surprised to see the author of what I consider one of the most important interactionist theories in applied psychology (Tett & Burnett, 2003) offer an entirely context-free understanding of validity.

My understanding of validity is in fact different. In my view, the key term in the AERA definition is “the interpretation ... proposed by the test user.” Perhaps even more relevant for selection research, the *SIOP Principles* (2018) state that the “essential principle in the evaluation of any selection procedure is that evidence be accumulated to support an inference of job relatedness.” (p.4). This definition is not in conflict with that of the *Standards*, as the interpretation proposed by test users selecting applicants is the prediction of job-related behavior. I will not discuss whose understanding is more “unified” or “antiquated,” which I would not consider a particularly fruitful undertaking. Rather, I discuss a number of implications of T&S’s understanding of validity vis-à-vis mine.

Recently, colleagues and I (Marcus et al., 2020) referred to the context-free understanding of validity as the psychometric perspective on selection. This view may largely overlap with T&S’s FIB camp, although I prefer our own label. According to this perspective, all evidence of validity has its roots in the target construct test authors had in mind when they designed the test. For inferring job relatedness by means of personality tests, one first has to identify job-relevant personality constructs and then to find or design tests that measure these constructs. Wernimont and Campbell (1968) described this process as the “sign” approach to selection. No doubt, the merits of this traditional approach are documented by an enormous volume of evidence. Yet, as discussed earlier, for the issue at hand the “sign” approach is faced with the problem that personality tests are typically designed with reference to typical behavior, whereas the selection context calls for showing maximum performance. As reviewed by T&S, there is evidence that this misfit between test authors’ intentions and situational demands may lead some test takers to alter their responses in the direction of perceived demands. From a traditional psychometric perspective, this change must be considered bias, as tests no longer measure the constructs in the same way as under neutral (or typical performance) conditions. Contrary to T&S’s assignment of SPT to the FIG camp, this problem is explicitly recognized and discussed in SPT’s original formulation (Marcus, 2009).

However, SPT also explicitly addresses possible pos-

itive effects of the situational context of selection. More recently, we referred to this view as the social perspective on selection (Marcus et al., 2020). In line with the psychometric view, the social perspective holds that a systematic component, which was not intended by test authors, is entered to personality test scores in selection. SPT shares with socioanalytic theory a focus on this social side of testing, but whereas Hogan and colleagues seem to consider testing almost universally as communication of social reputation, SPT suggests to distinguish generalizable components of personality test scores (in line with the psychometric view) from components specific to the situation. What must be considered bias from the psychometric perspective may turn out to contain potentially *meaningful* information for the particular *use* (sic!) of test scores in this context. We may write these differing views as variations of the basic axiom of classic test theory:  $x$  (observed score) =  $t$  (true score) +  $e$  (random error). The variant of the psychometric (FIB) view would read:  $x = t + b$  (systematic bias) +  $e$ . By contrast, SPT would formulate:  $x = t + b + s$  (social meaning) +  $e$ . According to SPT, the inherent problem of personality testing applied to selection is that tests are exclusively designed to measure  $t$ , whereas  $s$  tends to be entirely neglected. From the perspective of the FIB camp, this neglect is consistent with fundamental psychometric principles, because  $s$  converges on  $b$  and thus is a component that detracts from construct validity regardless of the use of test scores. These differing views on the nature of situation-specific variance components lead to essentially opposing conclusions on how research and practice of selection should deal with these components.

Congruent with the logic of the FIB camp, all measures proposed to control for or to eliminate faking are aimed at simulating a testing situation equivalent to typical performance. Essentially, these measures try to produce scores of job applicants that may be interpreted as if test takers were *not* applying for a job. If successful, scores would reproduce the standard axiom of psychometrics ( $x = t + e$ ) by eliminating the  $b$  component. Although numerous (Burns & Christiansen, 2011) and still growing in number (e.g., Levashina et al., 2014), in my reading of the literature these measures still await demonstrating impressive effects, at least in terms of improvements of the job relatedness of test scores.

SPT and related approaches suggest taking a very different approach. Instead of trying to force the situation to become congruent with the test authors’ intentions, SPT suggests to adapt to the actual situation by developing a scoring that captures the  $s$  component. In technical terms, this approach is aimed at developing two independent scores for the interpretation of personality tests in selection. In addition to (not in place of!) the traditional score aimed at measuring the target construct ( $x_c = t + b + e$ ), which inevitably contains bias, a second score aimed at measuring

the social component is proposed:  $x_s = s + e$ .

Two practical solutions aligning with this complementary approach have been proposed in the literature. Fifteen years ago, König and colleagues (2006) already applied Kleinmann et al.'s (2011) ATIC concept to personality testing. More recently, Marcus et al. (2020) proposed and tested a measure labelled "ideal employee coefficient" (IEC), which is explicitly based on SPT. There are a number of important differences between ATIC and IEC measures. For example, ATIC refers to the construct level, it is focused on the ability or skill component of social meaning, and it requires some additional (and qualitative) data collection for obtaining scores. By contrast, the IEC is based on responses to single items, it combines skill and motivation elements of self-presentation, and it may be computed directly from test takers' regular responses. Whereas ATIC has also been applied to interviews and assessment centers (Kleinmann et al., 2011), the IEC so far is restricted to personality testing, although SPT aims at covering all selection instruments. Despite these differences, both approaches are based on the idea to complement regular personality scores with an additional score tapping  $s$  as a component akin to the maximum performance nature of the setting. Ironically, T&S explicitly called for an independent measure of ATIC, although ATIC already is and always has been measured independently. Perhaps most importantly, both ATIC (Klehe et al., 2012; Kleinmann et al., 2011) and IEC (Marcus et al., 2020) were shown to improve the prediction of job performance beyond regular scores, in case of the IEC including a large-scale field study in a high-stakes setting.

After having contrasted anti-faking measures aimed at eliminating the social component with measures aimed at measuring and utilizing this component, I may need to discuss a third approach to which T&S also briefly referred. T&S describe the contextualization of personality items (often realized by simply appending the phrase "at work") as "offering promise" (p.15) for the FIB camp, and they list contextualization along with anti-faking measures as a candidate for making personality tests "relatively immune to faking" (p.15). I hold that "promising" is an understatement in this case, as a considerable volume of research supports the incremental validity of contextualization for predicting job performance (Shaffer & Postlethwaite, 2012). Yet, I doubt that contextualization follows the logic of anti-faking measures or that of the FIB camp.

The theoretical rationale for contextualization is clarifying the frame of reference for test takers (Schmit et al., 1995). This way, job applicants no longer need to guess whether the behavior mentioned in items refers to unspecific tendencies or to the work context specifically (see message #2 in my earlier illustration of possible meanings of responses). This helps eliminating an unwanted source of variance, as individual test takers may otherwise employ different frames of reference, which contributes to the

component in test scores psychometrically. Contextualization thus is still rooted in the traditional psychometric view on testing, but it is not at all aimed at reducing faking. On the contrary, the logic of the FIB camp would imply that contextualization *facilitates* faking by offering some guidance to potential "fakers" who may otherwise lack the skills to analyze the employer's intentions. This is perhaps most evident by observed effects of contextualization on item means, which are shifted in the *more* desirable direction (Schmit et al., 1995). This approach thus takes the exact opposite route as anti-faking measures aimed at disguising the purpose of measurement or warning applicants of (imaginary!) means of detection to make them respond as if they were not applying for a job. Notably, those anti-faking measures yield lower means but fail to yield positive effects on validity (e.g., Converse et al., 2008), whereas contextualization leads to both higher means and to better validity. This pattern is inconsistent with the logic of the FIB camp. I still describe contextualization as a third way independent of, though not contradictory to, the social view, as the context referred to is not that of the testing situation (i.e. selection) but that of behavior shown at work after being hired.

To summarize, SPT suggests that validity in selection is to be defined with reference to the purpose of predicting job-related behavior and that selection tools shall be tailored to the competitive nature of the context rather than trying to make it disappear. Based on these premises, SPT calls for complementing the traditional scoring of personality tests with context-specific additional scores tapping the social element in applicant responses. The value of these measures, like that of any other tool proposed for selection including anti-faking measures, should be evaluated based on evidence of job relatedness. Mere effects on mean scores are entirely useless for that purpose, as is most evident from findings on the contextualization of personality tests.

### Concluding Remarks

In three sections labelled as (golden) "rules," I tried to contrast the understanding of the nature and consequences of applicants' self-presentation on personality tests from a SPT perspective with both T&S's interpretation of that perspective, which they assigned to a FIG camp they created, and with the same authors' understanding of faking from their perspective labelled FIB. Although in my view the two positions do not constitute polar opposites as implied by T&S's labels, they differ substantially and in many respects lead to contradictory predictions. I will not reiterate the list of controversial issues discussed in this paper, which I by no means consider exhaustive.

Rather, I'd encourage researchers interested in this topic to take those controversial issues as starting points for designing studies aimed at resolving the controversies. In applied psychology, there generally is a paucity of research

testing contradictory positions in one study. The bulk of empirical studies I have seen as a reader, a reviewer, an editor, and—admittedly—at times also as an author of research papers, is aimed at testing some kind of theoretical “model” based on more or less eclectic collections of plausible and hardly controversial theoretical rationales. Typically, those models consist of constructs placed on the left side of a panel from which arrows point to other constructs positioned on the right side, often interrupted by still other constructs placed somewhere in the middle. Those kinds of models typically hurt nobody, and I have no doubt that they often lead to interesting insights. However, the history of science has shown that real progress is often made the harder way after lively debates between scholars disputing each other’s incompatible views. There are not many fields in applied psychology with more potential for controversy than “faking” or self-presentation. Although I disagree with T&S on more issues than I agree with them, it was fun reading their paper, and it was also fun writing this reply. The next logical step would be clarifying the controversial matters by means of empirical research, which I hope may be inspired by debates like the present one.

Qualifying my above call to some extent, there is also one particular area of research for which I’d call for a halt. This area is the further development and test of any anti-faking measures. This area seems to have engaged the creativity of assessment researchers like hardly anything else. We have seen tons of “validity” or “social desirability” scales, item formats of various kinds designed only to disguise the testing purpose and outsmart test takers, bold interpretations of test takers’ behavior ranging from counts of “blatant extreme responses” to measuring response latencies in milliseconds, psychometric gymnastics so sophisticated that nonuse in practice seems almost guaranteed, warnings against faking, which are themselves a perfect fake (Dear test taker: No, we cannot detect faking, so lay back and relax), and many, many more like these. I am not aware of any area in our field where so much of investment yielded so little return in terms of knowledge accumulated and practical improvements. Enough of this!

On a more personal note, I’d add that the entire faking paradigm seems to be based on a paternalistic attitude toward job applicants, who are seen as objects of measurement rather than as clients with their own legitimate, and perhaps sometimes illegitimate, needs and interests. In my view, our position as scholars of these phenomena should be that of attentive, empathetic, but also distal and impartial observers of both the employer and the applicant. Both employers and applicants are to be considered equally valued clients, regardless of who pays our bills. I therefore hold that it is not our role to tell applicants what to do (i.e. being “honest,” for whatever that means), but rather to try to understand why they do what they do, what the consequences of their behavior are, and eventually to find ways of bal-

ancing the interests of all parties involved in the selection process. One of the things I find most disturbing in faking research is the presumption of a judge’s role on the moral value of our clients’ behavior, which is something we are not entitled to in my view.

I also believe that we have no obligation to prevent “the demise of self-report personality assessment in hiring settings” (T&S, p.13). Personality tests are means, not ends, for our actual task of predicting job-related behavior. Even designing “a test [that] measures what it is purported to measure” (T&S, p.8) is only a means toward that end. If it turns out that personality tests do not serve this purpose, we may abandon them, and the world will continue to turn. However, I have also tried to show that, ironically, a few suggestions potentially useful for preventing that from happening came from the very “camp” T&S blamed for working on the demise of personality testing for selection.

In conclusion, T&S are probably right that there are many scholars in this field whose research is based on the premise that “faking is bad.” There are also researchers, although probably much fewer in number, who doubt this premise, at least in its unbalanced form of a credo. I am one of those latter researchers, as shall be clear for every reader of this paper. I don’t think that sceptics like me constitute anything like a coherent “camp,” though, let alone subscribing to a credo as stupid as “faking is good.” We shall continue studying the phenomena described by this label, but we shall stop using the label.

## REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing* (4th ed.). American Educational Research Association.
- Burns, G. N., & Christiansen, N. D. (2011). Methods of measuring faking behavior. *Human Performance*, 24(4), 358–372.
- Converse, P. D., Oswald, F. L., Imus, A., Hedricks, C., Roy, R., & Butera, H. (2008). Comparing personality test formats and warnings: Effects on criterion-related validity and test-taker reactions. *International Journal of Selection and Assessment*, 16(2), 155–169.
- Griffith, R. L., Lee, L. M., Peterson, M. H., & Zickar, M. J. (2011). First dates and little white lies: A trait contract classification theory of applicant faking behavior. *Human Performance*, 24(4), 338–357.
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, 88(1), 100–112.
- Hummert, H., Traum, A., & Marcus, B. (2021). Behaviour in selection situations as an adaptation to external expectations: Testing a theory of self-presentation. *European Journal of Work and Organizational Psychology*. Advance online publication: <https://doi.org/10.1080/1359432X.2021.1981866>

- Johnson, J. A., & Hogan, R. (2006). A socioanalytic view of faking. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 209-231). Information Age.
- Klehe, U. C., Kleinmann, M., Hartstein, T., Melchers, K. G., König, C. J., Heslin, P. A., & Lievens, F. (2012). Responding to personality tests in a selection context: The role of the ability to identify criteria and the ideal-employee factor. *Human Performance*, 25(4), 273-302.
- 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. *Organizational Psychology Review*, 1(2), 128-146.
- König, C. J., Melchers, K. G., Kleinmann, M., Richter, G. M., & Klehe, U. C. (2006). The relationship between the ability to identify evaluation criteria and integrity test scores. *Psychology Science*, 48(3), 369-377.
- Levashina, J. (2018). Evaluating deceptive impression management for personnel selection and job performance. In R. Rogers & S. D. Bender (Eds.), *Clinical assessment of malingering and deception* (4th edition, pp. 530-551). Guilford Press.
- Levashina, J., & Campion, M. A. (2007). Measuring faking in the employment interview: Development and validation of an interview faking behavior scale. *Journal of Applied Psychology*, 92(6), 1638-1656.
- Levashina, J., Weekley, J. A., Roulin, N., & Hauck, E. (2014). Blatant extreme responding in high-stakes selection. *International Journal of Selection and Assessment*, 22(4), 371-383.
- Marcus, B. (2009). "Faking" from the applicant's perspective: A theory of self-presentation in personnel selection settings. *International Journal of Selection and Assessment*, 14(4), 417-430.
- Marcus, B., Goldenberg, J., Fine, S., Hummert, H., & Traum, A. (2020). Self-presentation in selection settings: The case of personality tests. *Journal of Business and Psychology*, 35, 557-571.
- McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across non-cognitive measures. *Journal of Applied Psychology*, 85(5), 812-821.
- McFarland, L. A., & Ryan, A. M. (2006). Toward an integrated model of applicant faking behavior. *Journal of Applied Social Psychology*, 36(4), 979-1016.
- Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion related validity. *Journal of Applied Psychology*, 80(5), 607-620.
- Shaffer, J. E., & Postlethwaite, B. E. (2012). A matter of context: A metaanalytic investigation of the relative validity of contextualized and noncontextualized personality measures. *Personnel Psychology*, 65(3), 445-494.
- Society for Industrial and Organizational Psychology, Inc. (2018). *Principles for the validation and use of personnel selection procedures* (5th ed.). Author.
- Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, 88(3), 500-517.
- Tett, R. P., & Simonet, D. V. (2011). Faking in personality assessment: A "multisaturation" perspective on faking as performance. *Human Performance*, 24(4), 302-321.
- Tett, R. P., & Simonet, D. V. (2021). Applicant faking on personality tests: Good or bad and why should we care? *Personnel Assessment and Decisions*, 7(1), Article 2.
- Wernimont, P., & Campbell, J. P. (1968). Signs, samples, and criteria. *Journal of Applied Psychology*, 52(5), 372-376.

RECEIVED 07/29/21 ACCEPTED 11/29/21