

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of
Business

Lee Kong Chian School of Business

4-2018

Assessing personality dynamics in personnel selection

Joanna SOSNOWSKA

Vrije University

Joeri HOFMANS

Vrije University

Filip LIEVENS

Singapore Management University, filiplievens@smu.edu.sg

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



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

Citation

SOSNOWSKA, Joanna; HOFMANS, Joeri; and LIEVENS, Filip. Assessing personality dynamics in personnel selection. (2018). *The handbook of personality dynamics and processes*. 1-38.

Available at: https://ink.library.smu.edu.sg/lkcsb_research/6223

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

Assessing personality dynamics in personnel selection

Joanna Sosnowska, Joeri Hofmans, & Filip Lievens

Reference:

Sosnowska, J., Hofmans, J., & Lievens, F. (in press). Assessing personality dynamics in personnel selection. In J. F. Rauthmann (Ed.). *The handbook of personality dynamics and processes*. Elsevier

Correspondence address

Joanna Sosnowska: Vrije Universiteit Brussel, Faculty of Psychology and Educational Sciences, Department of Work and Organizational Psychology (WOPs), Pleinlaan 2, 1050 Brussel, Belgium; email: joanna.sosnowska@vub.be

Acknowledgements:

This research was supported by Fonds Wetenschappelijk Onderzoek – Vlaanderen (FWO) (grant number G024615N).

Abstract

Recently, there have been repeated calls in the literature for an integrative approach to personality, in which both between- and within-person fluctuations are simultaneously considered. Although the integrative approach to personality offers a compelling extension of the traditional trait approach, one of the major challenges is its applicability in applied settings. In the present chapter, we address this challenge for the domain of personnel selection, showing that an integrative approach to personality assessment in selection settings is possible through careful consideration of available theories and selection methods. By explaining and delineating how existing concepts can be used and how existing selection methods can be adjusted and expanded to measure these dynamic personality constructs, the present chapter contributes to a better assessment and understanding of personality in selection contexts, which in turn should result in better predictive validities.

Keywords: personality dynamics, personality assessment, personnel selection, Situational Judgement Tests, Assessment Centers

Assessing personality dynamics in personnel selection

In recent years, there have been several calls in personality psychology to supplement the traditional trait approach to personality—focusing on how people feel, think, and behave on average—with a more dynamic approach—focusing on how these feelings, thoughts, and behaviors fluctuate within one individual. In Industrial and Organizational psychology, this dynamic approach to personality has also made some inroads, with several studies showing that within-person fluctuations in personality states relate to a variety of work outcomes, such as job performance (Debusscher, Hofmans, & De Fruyt, 2014, 2016a, 2016b, 2016c, 2017; Huang & Ryan, 2011), learning and transfer (Huang & Bramble, 2016), mood and job satisfaction (Judge & Ilies, 2002), and work motivation (Judge, Simon, Hurst, & Kelly, 2014). Despite this evidence of within-person fluctuations being predictive of a range of important work outcomes, trait assessments still dominate the selection field. One important reason is that dynamic personality constructs are challenging to measure in a selection setting (Lievens et al., in press). To address this issue, the current chapter shows that assessing personality dynamics in a personnel selection setting is possible when one carefully considers available theories and methods. Moreover, we aim to give direction on how to use recent advances in personality research by introducing, explaining, and discussing how existing personnel selection methods can be adjusted to assess dynamic personality constructs.

In what follows, we start by providing a brief overview of how personality assessment is typically conducted in personnel selection. Next, we argue how a dynamic approach to personality can address some of its limitations. To this end, we review dynamic personality constructs that are relevant to personnel selection and that are feasible to apply in a selection setting. These dynamic personality concepts range from macro-level concepts (i.e., within-person stability and change) to micro-level concepts (i.e., how specific interactions between person characteristics and situations arise). The final part discusses how those theoretical

concepts can be assessed using adjusted versions of selection methods that exist already in the repertoire of selection psychologists. An overview of these dynamic concepts and the associated selection procedures can be found in Table 1.

Traditional personality assessment in personnel selection

The aim of personnel selection is to assess whether a candidate has the knowledge, skills, and abilities that are necessary to perform effectively in a particular job. To do so, the selection procedure typically includes a wide variety of assessments, with personality assessment being a commonly used one.

There are several reasons why personality tests are prevalent in personnel selection. First, research shows that personality adds incremental value above and beyond general mental ability or bio-data when predicting work performance (e.g., Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Hogan & Holland, 2003; Judge & Zapata, 2015). Moreover, most personality traits show little or no variation as a function of ethnic or racial group (Roth, Bobko, & Buster, 2013), which is particularly important for social and legal reasons. So, personality assessments are generally conceived of as a valid, reliable, and legally sustainable way for assessing candidates' potential to be perform effectively at work.

In a selection setting, a number of different methods can be used to assess candidates' personality, ranging from self-assessments (i.e., psychometric tests or Situational Judgement Tests [SJT]) to observer ratings (i.e., interviews, structured letters of reference, or assessment centers [ACs]). Despite the fact that these methods show many differences, they are all based on the same principle: they aim at measuring general and stable predispositions that are believed to be predictive of job-related behavior. According to this logic, applicants' estimated trait level on the basis of the selection procedure is an indicator of behavior in the future work context because the traits are considered to be underlying determinants of behavior across a wide variety of different contexts.

Although traditional personality assessment contributed to the effectiveness of selection decisions and prediction, it ignores that trait-related behaviors may differ not only across contexts (i.e., selection versus work), but also within one context across situations and time. Indeed, instead of dismissing within-person fluctuations as measurement error, there exists now relative consensus that the lack of stability in behavioral manifestations is meaningful because it results from the interaction between personal characteristics and people's perceptions of the situation (Funder, 2006, 2016; Reis, 2008). Moreover, as several studies showed that these dynamic elements of personality are predictive of a wide variety of job outcomes (e.g., Judge et al., 2014; Lievens et al., in press), it is essential to go beyond stable traits and acknowledge that also within-person fluctuations provide meaningful information about one's future behaviors.

Towards a dynamic approach to personality in selection

In the last decades, an increasing number of studies revealed that within-person fluctuations in trait-relevant behaviors are not random, but instead represent meaningful within-person variability (Dalal, Meyer, Bradshaw, Green, Kelly, & Zhu, 2015; Debusscher et al., 2016a, 2016b, 2016c, 2017; Fleeson, 2001; Hofmans, Debusscher, Doci, Spanouli, & De Fruyt, 2015; Judge et al., 2014). Hence, the personality field is adopting a dynamic perspective on personality, maintaining that personality can be conceptualized as a dynamic system that consists of both between-person stability and within-person variability (e.g., as seen in Whole Trait Theory: Fleeson & Jayawickreme, 2015).

A core idea of the dynamic perspective on personality is that dispositional and situational factors jointly influence behavior (Cervone, 2005; Fleeson, 2004; Fournier, Moskowitz & Zuroff, 2009; Furr, 2009; Mischel & Shoda, 1995; Shoda, LeeTiernan & Mischel, 2002; Smith, Shoda, Cumming, & Smoll, 2009). More specifically, the dynamic perspective holds that the expression of traits is situation-bound, with psychologically active

characteristics of situations triggering trait-relevant behavior (Mischel, 1973; Mischel & Shoda, 1995, 2008; Sherman, Rauthmann, Brown, Serfass, & Jones, 2015; Tett & Burnett, 2003; Tett & Guterman, 2000). Moreover, the dynamic perspective on personality does not assume a one-directional relation from situations to personality-related behaviors, but holds that people create their situations as much as they are affected by them (Bandura, 1978; Schneider, 1987).

As opposed to the traditional trait approach, the dynamic perspective acknowledges that personality cannot be reduced to one's average level of behavior, feeling, and thinking. As people respond in different ways to situational triggers, they differ not only in their average level of behavior, feeling, and thinking, but also in other respects, such as the variability of their everyday behavior across situations. Moreover, when accepting the situation-specificity of traits, one can also inspect how the interaction between characteristics of the individual and situational variables result in idiosyncratic trait manifestations and behavioral patterns. In sum, by taking into account the situation-specificity of personality manifestations, the dynamic perspective on personality has the potential to further improve the predictive validities of personality measures in personnel selection.

Personality dynamics – a review of relevant concepts

Even in the earliest writings on the Five Factor Model (FFM), the dynamic nature of personality has been acknowledged. McCrae and Costa (1999), for example, argued that the FFM of personality should be conceived of as a dynamic psychological organization that coordinates our experience and action. Yet, only recently personality psychology has started to shift from a descriptive, trait-based approach to an explanatory and dynamic perspective. To enhance our understanding of the dynamics of personality, a wide range of topics has been studied, including stability and change in personality, dynamic interactions between

person and situation variables, mechanisms underlying trait expression, and contextual aspects of personality.

In this chapter, we focus on three main areas of research on personality dynamics, moving from a macro- to a micro-level. First, at the macro-level, we review how stability and change combine when considering behavioral manifestations of traits (i.e., research on within-person variability as a predictor of work-related behavior). Second, on a meso-level, we focus on how within-person variability moderates the personality-performance relation (i.e., research on traitedness and personality strength). Lastly, we cover the dynamic interaction between characteristics of the individual and of the situation, thereby reviewing research on personality signatures (i.e., Fournier, Moskowitz, & Zuroff, 2008, 2009; Smith et al., 2009), the cognitive-affective processing system (CAPS) model (Mischel, 2004; Mischel & Shoda, 1995, 2008), and trait activation theory (i.e., Tett & Burnett, 2003; Tett & Gutterman, 2000).

Within-person personality stability and change

As personality exhibits both stability and within-person variability, traits and personality states are fundamental to understanding personality (Judge et al., 2014). One of the best known and most popular approaches to the integration of traits and states is the density distribution approach of Fleeson (2001). This approach draws on the idea that, although traits are useful in predicting behavior over longer periods of time, in their day-to-day behavior people actively display a wide range of state levels (Fleeson & Jayawickreme, 2015; Fleeson & Nofle, 2009). The direct result of this observation is that the average level of these state-behaviors does not represent the entire spectrum of the individual's behaviors and is therefore an incomplete indicator of personality trait. Instead, the density distribution approach suggests that, because behavioral manifestations of the trait form a distribution of

states over time, looking at the entire distribution of state levels provides a richer picture of one's trait.

In line with this idea, people were found not only to differ from each other in their average level of behavior (Dalal et al., 2015; Fournier, Moskowitz, & Zuroff, 2009); the level of behavioral variability (Debusscher et al., 2016c), the most frequent behavior (mode), the starting level of the behavior (minimum), and the behavior where individuals unfold their maximum value may differ between individuals (Fleeson, 2001; Fleeson & Gallagher, 2009).

Moreover, and of key importance to personnel selection, these density distribution parameters were shown to predict important work-related outcomes. For example, variability in the interpersonal circumplex relates to social relations, with high variability on the communion and agency axes (as measured by interpersonal spin) being associated with low closeness of social relations (Coté, Moskowitz, & Zuroff, 2011). These results suggest that people with high interpersonal circumplex variability were less likely to form close relations at work, possibly because their behavior is difficult to anticipate which affects liking, trust, and coordinated performance (Maddux, Mullen, & Galinsky, 2008). At the same time, research also shows that intra-individual variability in personality states was positively linked with performance, predicting peer-rated academic performance over and beyond mean trait scores (Lievens et al., 2016), which indicates that behavioral variability can be an adaptive feature in a workplace setting. In line with this idea, within-person variability was linked to self-rated functional flexibility, showing that people who are able to appropriately adjust their behavior to situational demands exhibited high behavioral variability across time and situations (Lievens et al., 2016). In personnel selection, behavioral variability might therefore serve as an indicator of how well the candidate is able to adapt to dynamic and complex workplace interactions, which is crucial in 21st century organizations (Huang, Ryan, Zabel, & Palmer, 2014).

Traitedness and personality strength

Apart from the direct effect of within-person variability on several relevant (work-related) outcomes, within-person variability also emerged as a boundary condition for trait personality-outcome relations, thereby providing an explanation for the moderate to weak relations between personality and work behaviors. This idea is particularly apparent in the literatures on traitedness (Baumeister & Tice, 1988) and on personality strength (Dalal et al., 2015).

The concept of traitedness draws on the assumption that people differ in the consistency of their trait-relevant behaviors, with higher levels of consistency being indicative of high levels of traitedness (Baumeister & Tice, 1988). The direct consequence of the existence of between-person differences in behavioral consistency is that the same average trait level is not equally informative for individuals with different levels of traitedness. For those high in traitedness, their average trait level is more representative of their everyday behaviors than for people low on traitedness. Hence, trait scores are expected to have a higher predictive validity for people high in traitedness than for people low in traitedness.

Although drawing on fundamentally the same idea, traitedness was operationalized in different ways across studies. Some studies conceptualized traitedness as the extent to which the different behavioral indicators of one personality dimension (i.e., the different items in the personality questionnaire) co-vary within one and the same individual (Tellegen, 1988). Others referred to traitedness as the extent to which an individual behaves in a similar way when being in the same situations. Finally, yet others conceptualized traitedness by examining the variability in personality states across different situations, which equates traitedness with within-person variability as measured by the density distribution approach

(Debusscher et al., 2016c). In the present chapter, this latter interpretation is also how we conceptualize traitedness.

Personality strength (Dalal et al., 2015) is a concept that strongly resembles traitedness. Personality strength taps into one's coherence of trait-relevant behavior and is in fact the person-related counterpart of situation strength (which represents cues provided by environmental forces with respect to the appropriateness or desirability of potential behaviors). A strong personality thus promotes similar behaviors across different situations, reflecting a strong consistency of internal cues associated with a particular trait. As with traitedness, this consistency is independent of the individual's standing on a particular trait. For example, this means that two individuals may have the same moderate level of trait neuroticism, but the person with a strong neuroticism trait will exhibit less variability around his/her moderate trait neuroticism score and therefore act in a similar, moderate neurotic manner across situations, whereas the person with a weak neuroticism trait will display more variability resulting from situational impacts, acting in both highly neurotic and less neurotic states.

In summary, an examination of behavioral consistency is crucial in a selection process because the selection decisions are based on the assumption that applicants' behavior (and/or his/her self-reports of such behavior) during the selection procedure is a valid predictor of his/her future work behavior. Importantly, focusing on traitedness or personality strength through the assessment of within-person variability allows assessing to what degree the trait assessments made in the selection procedure are predictive for the candidate's future behavior in the work context. Thus, as trait-related behaviors will be a more accurate predictors of work performance for highly traited individuals than for those who are less traited, examining traitedness or personality strength in a selection setting has the potential to enhance the predictive validity of personality assessment.

Dynamic interactions between personality and situations

In what follows, we shift the focus from cross-situational consistency towards clusters of regularities in contextualized trait expressions. The idea underlying this approach is that people change their behavior in response to situations. This interactionist principle is often referred to as trait activation. According to trait activation theory, the behavioral expression of a particular trait requires stimulation of that trait through a relevant situational cue (Tett & Burnett 2003). Hence, the link between personality and trait-relevant behavior is stronger in situations that are highly relevant to the trait under consideration (i.e., situations that contain more cues for trait-relevant behavior), which suggests that situational trait relevance moderates the relation between personality and behavior. For example, in a work context, it means that one's score on a particular trait is more predictive for one's performance when the job or task is highly relevant for that particular trait, such as agreeableness and extraversion in interpersonal situations, or neuroticism and conscientiousness in task-related situations (Kell, Rittmayer, Crook, & Motowidlo, 2010).

This conditional approach to the person - situation interaction is also reflected in the cognitive-affective processing system (CAPS) model of personality (Mischel & Shoda, 1995, 1998, 1999). According to the CAPS model, behavior is dependent on how a person processes the situational characteristics (s)he is confronted with in a particular situation. That is, people encode the psychological characteristics of the situation, which in turn activate their cognitions and emotions which in turn manifest in his/her behavioral reactions. Due to individual differences in these situation-cognition/emotion-behavior links, each individual can be characterized by a unique profile that consists of stable clusters of "if... then..."-prepositions, which represent how specific situational characteristics ('if...') trigger specific behavioral responses ('then...'). In the context of the CAPS model, these prepositions are also denoted to as behavioral signatures (Furr, 2009). Drawing on the idea of behavioral

signatures, Minbashian, Beckmann, and Wood (2010) demonstrated that people differ in the extent to which they vary their level of conscientiousness as a function of task demands at work, and that these individual differences in “if...then...”-contingencies predict adaptive performance on a lab task, over and beyond the level of trait conscientiousness. Moreover, contextualized situation-behavior patterns seem to be more stable than average trait levels (Furr & Funder, 2004), with the predictive validity of the contextualized situation-behavior patterns increasing when situations become more similar (Furr & Funder 2004; Sherman, Nave, & Funder, 2010). Therefore, assessing how personality is expressed as a function of the situational demands that the candidate will be confronted with in his/her future job may yield better insights in the personality-performance relation and has the potential to substantially improve the predictive validity of personality assessment.

Assessing personality dynamics using personality inventories

As noted above, in a selection context, personality has been mostly measured using self-report personality questionnaires. In personality questionnaires, the candidate is typically asked to report on how s(he) behaves, feels, and thinks in general, across a wide range of situations. Whereas it has been shown that well-constructed personality measures are valid predictors of overall job performance, recent advances in personality research suggest that going beyond one's average trait levels might lead to a more profound understanding of one's personality and therefore to a better prediction of one's future work performance. In the next sections, we discuss how existing personality measures can be adapted to allow assessment of dynamic components of personality.

Assessing within-person stability and change using personality inventories

One way to capture average trait levels as well as the extent to which people vary in their trait expressions across situations via a personality questionnaire, is to replace the traditional Likert-type response format with a frequency-based response format (i.e., Edwards

& Woehr, 2007). Frequency-based measurement is based on a distributional assessment model (Kane, 1986, 2000), which requires participants to report how frequently particular levels of the behavior occurred within a specified period of time. For example, instead of asking individuals to indicate to what extent each item in the questionnaire describes them or applies to them, the frequency-based measurement format asks them to indicate the relative frequency with which each of a set of response categories (e.g., very inaccurate, neither accurate nor inaccurate, and very accurate) reflects their behavior in the past six months. This frequency-based measurement format thus collects people's answers that describe the perceived distribution of their behavior, thereby capturing both the average level and the variability of trait expressions. For example, using frequency-based measurement, the average (trait) level is obtained by computing a weighted sum of the percentages (e.g., $(.01 \times \% \text{ very inaccurate}) + (.03 \times \% \text{ neither accurate nor inaccurate}) + (.05 \times \% \text{ very accurate})$). The variability of trait expressions is computed via a standard deviation across the three percentages per item (within-item *SD*), after which they can be averaged across all behavioral indicators of the personality dimension.

A frequency-based response format has several advantages over a Likert-type response format. First, it appears to be cognitively easier to recall frequencies of behaviors than mentally estimating an average behavioral level across time (Cosmides & Tooby, 1996). This is an important advantage because research revealed that frequency estimations are highly correlated with actual frequency counts in everyday life (Kane & Woehr, 2006). Second, and of key importance in a selection setting is that personality inventories based on frequency estimations are not only less susceptible to rating errors, but that they are also less vulnerable to faking (Fleisher, Woehr, Edwards, & Cullen 2011). Third, and central to the dynamic approach to personality, the frequency-based format allows capturing temporal stability and change in a single testing administration. Research based on the frequency-based format

showed that this procedure indeed increases the predictive validity of personality because within-person variability moderates the relation between personality traits and work outcomes (Edwards & Woehr 2007; Fleisher, Woehr, Edwards, & Cullen, 2011). Hence, a frequency-based response format provides more precision in predicting behavior than a traditional Likert-type response format.

Assessing person-situation interactions using personality inventories

As research showed that people behave differently in different contexts (e.g., at home or at work), another option to increase the predictive validity of personality questionnaires consists of contextualizing personality questionnaires (Bing, Whanger, Davidson, & VanHook, 2004; Schmit, Ryan, Stierwalt, & Powell, 1995). Typically, this is done by inserting a work-related frame-of-reference in the instructions (e.g., ‘How confident do you feel *at work*?’). Such a work-related frame-of-reference aims to ensure that all applicants answer the questions in relation to their self-perceived propensity for the traits at work specifically. So, there should be less ambiguity about which frame-of-reference to adopt when answering the items.

The empirical evidence in favor of these contextualized personality inventories is promising. First, research revealed that contextualized personality inventories increase reliability because they indeed reduce between-person differences in the measured variables as well as inconsistencies within individuals in the frames-of-reference that are used to respond to generic personality items (Lievens, De Corte, & Schollaert, 2008). Other studies found that the factor structure of contextualized and generalized personality ratings was invariant, but that error variances were smaller in the contextualized form (Robie, Schmit, Ryan, & Zickar, 2000; Schmit, Ryan, Stierwalt, & Powell, 1995; see also Lievens et al., 2009).

Second, meta-analytic research showed that the validity of such contextualized personality scores was nearly double the size of those of generalized inventories for four of the Big Five traits (Shaffer & Postlethwaite, 2012). These validity results are not trivial because a broad criterion (job performance) was used here. In other words, the prediction improvements cannot be ascribed to both the predictor and the criterion becoming more narrow and similar.

Third, contextualized personality inventories seem to improve applicants' perceptions of personality inventories. In one experimental study, students favored the contextualized variant over the generalized one for its job-relatedness (Holtz, Ployhart, & Dominguez, 2005). However, the difference in perceptions between the two conditions did not reach statistical significance. Another study did find significant differences in perceptions of perceived validity and liking in favor of contextualized variants (Holtrop, Born, de Vries, & de Vries, 2014).

Assessing personality dynamics using Situational Judgement Tests

In the last two decades, Situational judgement tests (SJTs) have become well-established and popular selection procedures. In an SJT, candidates are presented with a series of work-related situations and are asked how they would behave in these situations. Typically, the situation descriptions (the SJT items) contain information about what is at issue in the situation, the people who are involved in the situation, the situation's novelty (has this situation happened before?), and where the situation takes place (Lievens et al., 2017). Most commonly, SJTs are in a paper-and-pencil format, but they are also available in verbal, computer-based, video-based, 3D-animated, and even avatar-based formats. There also exist various response options: picking one response option from a list, ranking the response options from most to least likely, or rating all response options. Apart from these closed-ended formats, there also exist open-ended SJTs where the candidate is not provided with a

list of response options but explains orally or in writing how (s)he would behave in each of the situations (e.g., Rockstuhl, Ang, Ng, Lievens, & Dyne, 2015).

Although most SJTs are not specifically developed for assessing personality, recently researchers have started designing SJTs that specifically target personality traits (see, e.g., Mussel, Gatzka, & Hewig, 2016 for a good example). Such a construct-driven SJT is developed by asking subject-matter experts (i.e., experts on personality psychology) to generate for each SJT item response options that reflect different trait levels (Lievens, 2017). In terms of scoring, candidates then typically receive a higher trait score when they select response options that are considered to reflect higher trait levels.

Research has shown that personality-related SJTs can be considered alternative and viable assessment methods of personality (Lievens et al., 2017). For example, Mussel et al. (2016) obtained an average convergent correlation of .59 between five narrow traits measured by an SJT and those same traits measured using self-reports (correlations ranged between .41 and .70). Moreover, they also found that the SJT scores were equally good in predicting a range of relevant criteria as compared to traditional self-reports. In addition to these findings, other advantages are that SJTs have low adverse impact (against specific ethnic subgroups), are well accepted among test takers, are less susceptible to faking, and are less dependent upon the candidate's ability to engage in introspection (see Lievens & De Soete, 2012, for a review). Thus, SJTs have been proposed as a promising alternative for self-reported personality questionnaires.

Yet, SJTs have other unique strengths that have not yet been fully appreciated (Lievens, 2017; Lievens et al., 2017). That is, in SJTs standardized situations are presented to people. In addition, one might manipulate specific situational features in these SJT situations. Accordingly, one can examine not only how someone generally behaves across situations, but also assess dynamic personality aspects, namely (a) within-person variability and (b)

person-situation contingencies. Importantly, the situations in an SJT are always work-related and standardized, which means that SJTs allow capturing pure within-context variability that is not confounded by cross-context variability (Geukes, Nestler, Hutteman, Kűfner, & Back, 2017). This is an important quality of SJTs as within-context and cross-context variability appear to be driven by both shared and unique processes (Geukes et al., 2017).

Assessing within-person fluctuations using SJTs

It is important to develop a construct-driven SJT when one aims to measure personality and within-person variability in particular (Lievens et al., in press). Traditional (non-construct-driven) SJTs are typically designed to optimally predict future job performance and measure a myriad of constructs (Christian, Edwards, & Bradley, 2010). Conversely, construct-driven SJTs are explicitly designed to measure a single construct (Lievens, 2017).

Although the construct-driven approach has traditionally been used to reduce the impact of unintended, non-job-relevant constructs, it can also be used to examine within-person fluctuations. That is, by looking at one's responses across the whole set of situations, one can capture both the average trait level as well as within-person variability (i.e., how much the indicated trait levels vary across situations). So far, research that capitalized on this potential benefit is scarce. Lievens et al. (2017) developed a construct-driven SJT and showed in three studies that an assessment of within-person variability significantly added to the prediction of job performance above and beyond average levels of the constructs of interest (sociability, dutifulness, and personal initiative). Moreover, they also showed that within-person variability in the responses across the written situations predicted actual variability in state traits measured 2 years later through an experience sampling study. These findings are promising because they suggest that the within-person variability as measured by

situational inventories is representative of one's level of within-person variability in everyday life.

To capture the full range of within-person variability with a single construct-driven SJT, some admonitions are in order. It is crucial to critically reflect on the situations to be included. The best way to ensure an adequate range of situations consists of drawing on a situational taxonomy. There are several newly developed situational models that can be used to generate situation descriptions, such as the DIAMONDS model (Rauthmann et al., 2014), specifying eight psychologically meaningful situational characteristics (e.g., duty, intellect, adversity), or the CAPTION model (Parrigon, Woo, Tay, & Wang, 2017), a seven-factor situational taxonomy based on lexical analysis (e.g., complexity, importance, humor). Apart from using situational taxonomies to guarantee that the most important situational dimensions are included, one needs to construct SJT items in such a way that the situational variance per situational dimension is maximized (Dalal et al., 2015). To this end, one needs to (1) select enough SJT items per situational dimension and (2) make sure that the situations cover the whole spectrum of the dimension. For example, in case of three SJT items per situational dimension, one SJT item should be low on the situational dimension, one should be moderate, and one should be high. As strong situations can limit the applicants' behavioral responses, the use of items with different levels of situational strength enable capturing the full extent of applicants' behavioral reactivity to situations.

Assessing person-situation interactions using SJTs

Unlike in experience sampling studies where situational characteristics might be *measured* by asking respondents to report on their perceptions of the situation, SJTs allow *manipulating* rather than just measuring characteristics. Hence, SJTs allow examining which situational characteristics are related to which behavioral trait manifestations. By doing so,

SJTs offer a more refined way of inspecting situation-behavior contingencies, which potentially increases the predictive value and utility of personality in a selection setting.

To assess such dynamic interactions between personality and situations in a meaningful way, it is pivotal to incorporate situational taxonomies, such as the DIAMONDS model (Rauthmann et al., 2014) or the CAPTION model (Parrigon et al., 2017) in the SJT items (see Horstmann, Rauthmann, & Sherman, in press for a review of several situational taxonomies). Using an established situational taxonomy guarantees that the most relevant situational characteristics are included, allowing for the comprehensive examination of meaningful trait-behavior situational contingencies.

Moreover, SJTs also allow examining individual differences in the perception of situational characteristics. As people differ in their selective attention and processing of situational characteristics, assessing how individuals perceive situations and their characteristics might provide valuable insights into the cognitive processes underlying their behavioral reactions. One way to do so is by using verbal protocol analysis (Ployhart, 2006), which involves asking applicants to describe their mental decision-making process. Using this procedure, Rockstuhl et al. (2015) demonstrated that situational judgments assessed via a verbal protocol predicted task performance and organizational citizenship behavior above and beyond typical response judgments. Another way of tapping into inter-individual differences in situation perception is to develop an SJT that simultaneously assesses the perception of the situation as well as behavioral intentions by asking the respondents to rate both. For example, Ziegler (2017) developed such an SJT (Big Five of Personality in Occupational Situations, B5PS). This SJT presents the candidate with 211 situational vignettes, such as “You just had your annual appraisal interview with your manager, in which you received a lot of detailed feedback.” For each situational vignette, the candidate is asked to report on a Likert-scale how (s)he perceived the situation (e.g., “I perceive this situation as challenging”) and how

(s)he would behave in this situation (e.g., “I reflect on the feedback”). To make sure that the situational vignettes assessed the most important personality traits and the most prominent situation dimensions, the vignettes combine the Big Five and Situation Five (a situational taxonomy based on a lexical approach).

Despite these advantages, an important downside of SJTs is that the situations are preselected and linearly presented, which means that they are identical for all candidates (Judge, Hofmans, & Wille, 2017). By standardizing the situations across participants, one fails to take into account that in real life people actively select, modify, and create situations (Rauthmann, Sherman, Nave, & Funder, 2015). Indeed, research by Bolger and Schilling (1991) showed that as much as one third of the within-person variance in personality-related behavior is due to situation selection (whereas the other two thirds are due to differential reactivity to situations). One solution might be to adopt interactive (or nonlinear/branched) SJTs (e.g., Kanning, Grewe, Hollenberg, & Hadouch 2006). In such interactive SJTs, applicant’s responses to the previous items determine which situations (or items) come next. In other words, the behavioral choices made by the candidate change the way the situation develops and evolves. Hence, interactive SJTs to some extent have the potential of capturing and reflecting the reciprocal effects between person and situation.

Assessing personality dynamics using Assessment Center Exercises

Similar to SJTs, assessment centers (ACs) are a popular method in the selection domain because they generate a wealth of behavioral information about applicants in a relatively short amount of time. ACs typically include a variety of situational exercises, such as role-plays, in-basket exercises, group discussions, oral presentations, and fact-finding exercises (Thornton & Mueller-Hanson, 2003). Generally, these AC tasks are also referred to as behavioral stimulations, which means that they resemble actual job-related tasks that should enable making predictions about applicants’ proficiency in these job-related areas. As

the exercises in ACs only pertain to the work domain, ACs tap into within-context variability, but not cross-context variability (Geukes et al., 2017). In other words, very much like SJTs, the variability measures obtained by ACs are not confounded by cross-context variability.

Much like SJTs, AC exercises tend to tap into more than one personality trait, with the observed behavior in each of the AC exercises being the result of the dynamic interaction of the situation and personality traits. In terms of differences between ACs and SJTs, it is important to note that AC exercises should not be confused with SJT situations because different AC exercises are often composed of several situational dimensions simultaneously. As ACs are high-fidelity simulations, they have some advantages over SJTs because they allow for direct observation of behavior rather than eliciting self-reported behavioral intentions. The AC methodology can be used to measure personality dynamics in two different ways, which we detail below.

Assessing within-person stability and change using ACs

Traditionally, inconsistencies in applicants' scores on the same dimensions across exercises were considered measurement error and were therefore believed to undermine AC's convergent and discriminant validity (see Lance, 2008). Only in the last ten years, it was recognized that people's behavioral inconsistency across AC exercises might actually provide meaningful information about applicants because it might indicate how applicants adjust their behavior to differing situational demands (Gibbons & Rupp, 2011). Indeed, Gibbons and Rupp (2011) pointed out that consistency should not be expected from applicants' scores and suggested that researchers and practitioners should instead focus on incorporating patterns of behavioral consistencies in ACs to get a better grip on the dynamics of how applicants' traits interact with the situations in the AC exercises and affect their performance.

Interestingly, methodologies that strongly resemble the AC procedure have already been applied in personality research on within-person stability and change (Lievens, 2017).

In particular, Fleeson and Law (2015) invited participants to the lab and asked them to perform a series of lab tasks on several occasions. This allowed for the direct observation of behavior in a controlled environment, while the large series of lab tasks allowed assessing the effects of situational variability on stability and change of the individual's behaviors.

Although the direct observation of behavior in a controlled setting such as a laboratory or an AC is certainly a way forward, an issue with most ACs is that they are based on a small number of exercises due to time and cost constraints (seven appears to be the average number of AC exercises; Gibbons & Rupp, 2009). As a result, ACs often do not yield sufficient observations per applicant to reliably assess within-person variability. In response to this issue, Brannick (2008) and Herde and Lievens (2016), suggested to use a multiple speed assessment procedure, which consists of a large number of short AC exercises. For example, instead of organizing a traditional 30-minute role-play task, one might organize 5 shorter role-plays. Such an adaptation to the AC procedure allows obtaining a larger sample of observations per candidate across independent tasks. If each task in a multiple speed assessment is specifically designed to evoke behavior related to a particular trait, the large series of tasks has the potential to provide valuable information about the applicants' behavioral variability. A less complex but potentially also less accurate alternative of assessing the consistency of the applicants' behavior consists of asking assessors to directly rate candidates' variability/consistency (e.g., at the end of the AC; Gibbons & Rupp, 2009).

As with SJTs, one should try to maximize the situational variance when developing AC exercises that measure personality dynamics. This can again be achieved by (1) constructing AC exercises that assess situational types inspired by a validated situational taxonomy and (2) ensuring that one has enough AC exercises to allow for exercises that cover the full spectrum (from low, over moderate, to high) of each situational dimension

(Dalal et al., 2015). Accordingly, it should be possible to develop an AC to measure the full range of applicants' work-related variability.

Assessing person-situation interactions using ACs

Scrutinizing person-situation contingencies across AC exercises offers yet another way to study personality dynamics in an AC context. The aim is then to focus on the overall pattern of behavioral responses in the AC exercises and on "if... then..."- contingencies, indicating under which conditions (that is: in which situations) an individual engages in specific behaviors. Along these lines, Gibbons and Rupp (2009) argued that connecting applicants' behavior with appropriate contexts allows examining their proficiency signatures, which represent individual differences in successful performance-situation contingencies (e.g., effective communication of someone in one-on-one but not in group settings).

The biggest challenge to obtain person-situation contingencies in ACs is that one has limited control over the situation. Unlike in SJTs, it is virtually impossible to control how the situations in most AC exercises develop. For example, a group discussion or role-play might be influenced by people taking part in the group discussion or role-play. Hence, these exercises might take different turns and swerve in different directions, generating different situational characteristics. Thus, to be able to study person-situation contingencies, it is necessary to not only measure candidates' behavior, but to also *measure* the situational characteristics in a systematic way. To this end, one might ask assessors to include ratings of situational characteristics alongside ratings of the candidate. Or perhaps even better, one might also use coders afterwards to code the situational characteristics on the basis of the videotaped candidate performances. Measurement of situational characteristics is preferably based on the aforementioned situational taxonomies. Accordingly, the effects of situational characteristics emerging during a particular AC exercise or a series of AC exercises might be used to evaluate "if ... then"- patterns.

Using role-player prompts/cues or technological advances in ACs represents an alternative way for dealing with the limited control over emerging situations in a typical AC (Oliver, Hausdorf, Lievens, & Conlon, 2016; Lievens, Schollaert, & Keen, 2015). In particular, Oliver and Lievens (2014) suggest using virtual adaptive stimulations and games, which allow better control over the situational characteristics that are presented to candidates. Hence, they permit studying how specific situational characteristics trigger trait-relevant behavior, even in dynamic settings that are difficult to control in real life (i.e., teamwork tasks). Rayburn (2007) used a similar technology to examine interpersonal dynamics between game players, showing it is a promising venue to explore in selection settings.

Quantification of within-person variability

So far, we have outlined several methods that allow capturing within-person variability across a range of situations. However, this deals with only one side of the coin. It is equally important to discuss how such within-person variability can be quantified. In the past, statistical indices and models that can be used for within-person variability were subject of some debate (e.g., Baird, Le, & Lucas, 2006).

Most of the existing research on within-person variability has operationalized within-person personality variability using the within-person standard deviation, computed as the *SD* of one's trait-relevant behavior scores across situations. However, this approach has several limitations. First, the within-person standard deviation is limited by the individual's standing on the trait (Cole, Bedeian, Hirschfeld, & Vogel, 2011): People who have very high or very low average trait levels also have a restricted range of the *SD* because of ceiling or floor effects, respectively (the *SD* approaches zero when the average is really high or really low; Baird et al., 2006). Second, besides capturing the true within-person variability, the within-person *SD* may also reflect other sources of variability such as measurement error or biases in the use of the rating scale (e.g., extreme answering).

As computing an *SD* based on ratings across trait-related items or situations is potentially problematic, several adjustments have been proposed. For example, Mestdagh, Pe, Pestman, Verdonck, Kuppens, and Tuerlinckx (2017) proposed a relative variability index, which is the proportion of variability that is observed relative to the maximum possible variability that can be observed given a certain trait score. By computing the variability relative to the maximum possible variability, this index controls for the fact that the average trait score and the variability are confounded. Mestdagh and colleagues tested the performance of this relative variability index in a simulation study, showing that their measure is indeed independent from the average trait score, thereby allowing researchers to study how variability predicts relevant criteria or moderates the trait-outcome relation.

Another approach to computing within-person variability is through item-response-theory (IRT). In these IRT models, the information is split into several sources of between-person variability: individual differences in the average trait level, individual differences in item difficulty, and individual differences in variability. For example, Lievens et al. (2017) used Böckenholt's three pseudo-item model (Böckenholt, 2012) to examine sociability and dutifulness. Using pseudo-items, this model mimics the following decision-making process: first the subject decides whether (s)he should give a mid-category answer or not (Pseudo-item 1). Then (s)he decides whether his/her answer is positive or negative (i.e., 'agree' or 'disagree'; Pseudo-item 2). Finally, (s)he decides on the strength of his/her answer ('agree' or 'strongly agree', Pseudo-item 3). By differentiating between these three decision-making steps, variability is captured in both the direction and strength of the answers. Moreover, the model can separate true within-person fluctuations from other sources of variability, such as measurement error and response biases, thereby obtaining a purer estimate of within-person variability.

As the traditional measures of within-person variability (i.e., within-person *SD*) and the extensions that have been discussed so far require many repeated observations, and because ACs typically only contain a limited number of exercises (note that this is not the case for the multiple speed assessment procedure), Gibbons and Rupp (2007) proposed an alternative method for measuring consistency in ACs (being the inverse of variability). Consistency assesses the extent to which an individual gets the same rating on the same dimension across exercises. To measure consistency, Gibbons and Rupp (2007) created an index of pattern similarity across AC exercises. This index is obtained by first computing the squared difference between ratings on the same dimension in two different exercises, then summing these squared differences across dimensions, and then taking the square root of this sum. This process is repeated for all possible pairs of exercises, after which an average pairwise index is computed across all pairs. In a simulation study, Gibbons and Rupp (2007) demonstrated that their consistency index indeed allows capturing individual differences in consistency in ACs. The major disadvantage, however, is that because of the collapsing across AC dimensions, the index does not provide information regarding consistency on individual dimensions (Gibbons & Rupp, 2009).

Choosing one or the other operationalization of within-person variability is not an easy task, particularly because there are no simulation studies available that allow for a direct comparison of the available alternatives. We therefore propose to choose the concrete operationalization based on the specifics of the data at hand. If one wants to compute an index of within-person variability based on few repeated AC observations, the Gibbons and Rupp (2007) consistency index is probably the best option, although one should realize that it assumes that the raters agree with each other in their ratings. When more repeated measures data are available but one suspects the data to be distorted by response biases, the Böckenholt's three pseudo-item model (Böckenholt, 2012) is probably best suited. Note that

this model can also be used to take into account between-item differences in item difficulty in an SJT (Lievens et al., in press). Finally, in the absence of such biases, the relative variability index of Mestdagh et al. (2017) might be a viable alternative. In sum, in the absence of simulation studies that show the relative performance of each of these measures in a wide range of conditions, the choice for one or the other measure depends on the researcher, who is tasked with balancing model complexity and the assumptions about the data (e.g., response biases) (s)he is willing to make.

Conclusion

In this chapter, we argued that small adjustments to existing selection procedures allow capturing not only average trait levels, but also assessing more dynamic personality concepts, such as within-person variability and "if... then..."- contingencies. By detailing how and where personality questionnaires, situational judgment tests, and ACs need to be adapted, and by reviewing how one can statistically capture within-person variability, we hope to pave the way for selection psychologists to start including personality dynamics assessments in their selection procedures and selection decisions.

References

- Baird, B.M., Le, K., & Lucas, R.E. (2006). On the nature of intraindividual personality variability: Reliability, validity, and associations with well-being. *Journal of Personality and Social Psychology, 90*, 512-527.
- Bandura A. 1978. The self-system in reciprocal determinism. *American Psychologist, 33*, 344– 58.
- Baumeister, R.F., & Tice, D.M. (1988). Metatraits. *Journal of Personality, 56*, 571-598.
- Barrick, M.R., & Mount, M.K. (1991). The big 5 personality dimensions and job-performance - a meta-analysis. *Personnel Psychology, 44*, 1-26.
- Barrick, M.R., Mount, M.K., & Judge, T.A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*, 9–30.
- Bing, M.N., Whanger, J.C., Davidson, H.K., & VanHook, J.B. (2004). Incremental validity of the frame-of-reference effect in personality scale scores: A replication and extension. *Journal of Applied Psychology, 89*, 150–157
- Bolger, N., & Schilling, E.A. (1991). Personality and problems of everyday life: the role of neuroticism in exposure and reactivity to daily stressors. *Journal of Personality, 59*, 355–386
- Brannick, M.T. (2008). Back to basics of test construction and scoring. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 131–133.
- Cervone, D. (2005). Personality architecture: Within-person structures and processes. *Annual Review of Psychology, 56*, 423–452.
- Cole, M. S., Bedeian, A. G., Hirschfeld, R. R., & Vogel, B. (2011). Dispersion-composition models in multilevel research: A data-analytic framework. *Organizational Research Methods, 14*, 718-734.

- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*, 1-73.
- Dalal, R. S., Meyer, R. D., Bradshaw, R. P., Green, J. P., Kelly, E. D., & Zhu, M. (2015). Personality strength and situational influences on behavior: A conceptual review and research agenda. *Journal of Management*, *41*, 261–287.
- Debusscher, J., Hofmans, J., & De Fruyt, F. (2014). The Curvilinear Relationship Between State Neuroticism and Momentary Task Performance. *PloS ONE*, *9*. DOI: 10.1371/journal.pone.0106989
- Debusscher, J., Hofmans, J., & De Fruyt, F. (2016a). The effect of state Core Self-Evaluations on task performance, organizational citizenship behavior, and counterproductive work behavior. *European Journal of Work and Organizational Psychology*, *25*, 89-104.
- Debusscher, J., Hofmans, J., & De Fruyt, F. (2016b). From State Neuroticism to Momentary Task Performance: A Person x Situation Approach. *European Journal of Work and Organizational Psychology*, *25*, 301-315.
- Debusscher, J., Hofmans, J., & De Fruyt, F. (2016c). Do personality states predict momentary task performance? The moderating role of personality variability. *Journal of Occupational and Organizational Psychology*, *89*, 330-351.
- Debusscher, J., Hofmans, J., & De Fruyt, F. (2017). The multiple face(t)s of state conscientiousness: Predicting task performance and organizational citizenship behavior. *Journal of Research in Personality*, *69*, 78-85.
- Edwards, B.D., & Woehr, D.J. (2007). An examination and evaluation of frequency- based personality measurement. *Personality and Individual Differences*, *43*, 803- 814.

- Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*, 1011- 1027
- Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology, 97*, 1097–1114.
- Fleeson, W., & Jayawickreme, E. (2015). Whole trait theory. *Journal of Research in Personality, 56*, 82-92.
- Fleeson, W., & Law, M. K. (2015). Trait enactments as density distributions: The role of actors, situations, and observers in explaining stability and variability. *Journal of personality and social psychology, 109*, 1090.
- Fleeson, W., & Nofhle, E. E. (2009). In favour of the synthetic resolution to the person-situation debate. *Journal of Research in Personality, 43*, 150–154.
- Fleisher, M.S., Woehr, D.J., Edwards, B.D., & Cullen, K.L. (2011). Assessing within-person variability via frequency estimation: More evidence for a new measurement approach. *Journal of Research in Personality, 45*, 535-548.
- Fournier, M. A., Moskowitz, D. S., & Zuroff, D. C. (2008). Integrating dispositions, signatures, and the interpersonal domain. *Journal of personality and social psychology, 94*, 531.
- Fournier, M.A., Moskowitz, D.S., & Zuroff, D.C. (2009). The interpersonal signature. *Journal of Research in Personality, 43*, 155–162
- Funder, D.C. (2006). Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality, 40*, 21-34.

- Funder, D. C. (2016). Taking situations seriously: the situation construal model and the riverside situational Q-sort. *Current Directions in Psychological Science*, 25, 203-208.
- Furr, R.M. (2009). Profile analysis in person-situation integration. *Journal of Research in Personality*, 43, 196–207
- Furr, R.M., & Funder, D.C. (2004). Situational similarity and behavioral consistency: Subjective, objective, variable-centred, and person-centred approaches. *Journal of Research in Personality*, 38, 421–447.
- Geukes, K., Nestler, S., Hutteman, R., Kűfner, A.C.P., & Back, M.D. (2017). Trait Personality and State Variability: Predicting Individual Differences in Within- and Cross-Context Fluctuations in Affect, Self-Evaluations, and Behavior in Everyday Life. *Journal of Research in Personality*, 69, 124-138.
- Gibbons, A. M., & Rupp, D. E. (2007, April). *Inconsistency in assessment center performance: A meaningful individual difference?* Paper presented at the 22nd annual meeting of the Society for Industrial and Organizational Psychology, New York.
- Gibbons, A.M., & Rupp, D.E. (2009). Dimension consistency as an individual difference: A new (old) perspective on the assessment center construct validity debate. *Journal of Management*, 35, 1154-1180
- Herde, C.N., & Lievens, F. (2016, April). The added value of speed assessment for assessing leadership: In. P. V. Ingold, & B. J. Hoffman (Chair), AC, Quo Vadis? Novel assessment center design and validation strategies. Symposium conducted at the Annual Conference of the Society for Industrial and Organizational Psychology, Anaheim, CA, USA.
- Hofmans, J., Debusscher, J., Doci, E., Spanouli, M. & De Fruyt, F. (2015) The curvilinear relationship between work pressure and momentary task performance: the role of state

and trait core self-evaluations. *Frontiers in Psychology*, 6.

doi: [10.3389/fpsyg.2015.01680](https://doi.org/10.3389/fpsyg.2015.01680)

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, 88, 100–112.

Horstmann, K., Rauthmann, J. F., & Sherman, R. A. (in press). Measurement of situational influences. In V. Zeigler-Hill (ed.), *The SAGE Handbook of Personality and Individual Differences*. Sage.

Huang, J.L., & Bramble, R.J. (2016). Trait, state, and task-contingent conscientiousness: Influence on learning and transfer. *Personality and Individual Differences*, 92, 180–185.

Huang, J. L., & Ryan, A. (2011). Beyond personality traits: A study of personality states and situational contingencies in customer service jobs. *Personnel Psychology*, 64, 451-488.

Judge, T.A., & Zapata, C.P. (2015). The person-situation debate revisited: Effect of situation strength and trait activation on the validity of the big five traits in predicting job performance. *Academy of Management Journal*, 58, 1-31.

Judge, T.A., Simon, L. S., Hurst, C., & Kelley, K. (2014). What I experienced yesterday is who I am today: Relationship of work motivations and behaviors to within-individual variation in the Five-Factor model of personality. *Journal of Applied Psychology*, 99, 199-221.

Judge, T.A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology*, 87, 797–807.

- Judge, T.A., Hofmans, J. & Wille, B. (2017) Situational Judgement Tests and Personality Measurement: Some Answers and More Questions. *European Journal of Personality*, 31, 463-464.
- Kane, J.S. (1986). Performance distribution assessment. In R. A. Berk (Ed.), *Performance assessment: Methods and applications* (pp. 237-273). Baltimore: Johns Hopkins University Press.
- Kane, J.S. (2000). Accuracy and its determinants in distributional assessment. *Human Performance*, 13, 47-84.
- Kane, J.S, Woehr, D.J. (2006). Performance measurement reconsidered: An examination of frequency estimation as a basis for assessment. In W Bennett, Jr., C.E. Lance, D.J. Woehr (Eds.), *Performance measurement: Current perspectives and future challenges*. Mahwah, NJ: Lawrence Erlbaum.
- Kanning, U.P., Grewe, K., Hollenberg, S., & Hadouch, M. (2006). From the subjects' point of view: Reactions to different types of situational judgment items. *European Journal of Psychological Assessment*, 22, 168-176.
- Kell, H.J., Rittmayer, A.D., Crook, A.E., & Motowidlo, S.J. (2010). Situational content moderates the association between the Big Five personality traits and behavioral effectiveness. *Human Performance*, 23, 213–228.
- Lance, C. E. (2008). Why assessment centers do not work the way they are supposed to. *Industrial and Organizational Psychology-Perspectives on Science and Practice*, 1, 84-97.
- Lievens, F. (2002). Trying to understand the different pieces of the construct validity puzzle of assessment centers: An examination of assessor and assessee effects. *Journal of Applied Psychology*, 87, 675– 686.

- Lievens, F., De Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference effect in personality scale scores and validity. *Journal of Applied Psychology, 93*, 268- 279.
- Lievens, F., & De Soete, B. (2012). Simulations (pp.383-410). In N. Schmitt (Ed.) *Handbook of Assessment and Selection*. Oxford University Press.
- Lievens, F., Lang, J., De Fruyt, F., Corstjens, J., Vandervijver, M. & Bledow, R. (in press) A closer look at the assessment and predictive power of people's intra-individual variability across situations, *Journal of Applied Psychology*.
- Lievens, F. (2017). Construct-driven SJTs: Towards an agenda for future research. *International Journal of Testing, 29*, 1-7.
- Lievens, F. (2017). Assessing Personality-Situation Interplay in Personnel Selection Towards More Integration into Personality Research. *European Journal of Personality, 31*, 424-440.
- Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review, 80*, 252–283.
- Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology, 55*, 1–22.
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualising situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review, 102*, 246– 268.
- Mischel, W., & Shoda, Y. (1998). Reconciling processing dynamics and personality dispositions. *Annual Review of Psychology, 49*, 229–258.
- Mischel, W., & Shoda, Y. (1999). Integrating dispositions and processing dynamics within a unified theory of personality: The cognitive–affective personality system. In L. A.

- Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed.) (pp. 197–218). New York: Guilford Press.
- Mischel, W., & Shoda, Y. (2008). Toward a unified theory of personality: Integrating dispositions and processing dynamics within the cognitive-affective processing system. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 208–241). New York, NY: Guilford.
- Mussel, P., Gatzka, T., & Hewig, J. (2016). Situational judgment tests as an alternative measure for personality assessment. *European Journal of Psychological Assessment*
- Oliver, T., & Lievens, F. (2014). Conceptualizing and assessing interpersonal adaptability: towards a functional framework. In *Individual adaptability to changes at work: new directions in research* (pp. 52-72). Taylor & Francis Group.
- Oliver, T., Hausdorf, P., Lievens, F., & Conlon, P. (2016). Interpersonal dynamics in assessment center exercises: Effects of role player portrayed disposition. *Journal of Management*, 42, 1044-1074.
- Parrigon, S., Woo, S. E., Tay, L., & Wang, T. (2017). CAPTION-ing the situation: A lexically-derived taxonomy of psychological situation characteristics. *Journal of Personality and Social Psychology*, 112, 642–681
- Ployhart R. (2006). Staffing in the 21st century: New challenges and strategic opportunities. *Journal of Management*, 32, 868–897.
- Rayburn, E.M. (2007). Applying simulation experience design methods to creating serious game-based adaptive training systems. *Interacting with Computers*, 19, 206-214.
- Rauthmann, J.F., Gallardo-Pujol, D., Guillaume, E. M., Todd, E., Nave, C.S., Sherman, R.A., Funder, D. C. (2014). The situational eight DIAMONDS: A taxonomy of major dimensions of situation characteristics. *Journal of Personality and Social Psychology*, 107, 677-718.

- Rauthmann, J.F., Sherman, R.A., Nave, C.S., & Funder, D.C. (2015). Personality-driven situation experience, contact, and construal: How people's personality traits predict characteristics of their situations in daily life. *Journal of Research in Personality, 55*, 98-111.
- Reis, H.T. (2008). Reinvigorating the concept of situation in social psychology. *Personality and Social Psychology Review, 12*, 311–329
- Rockstuhl, T., Ang, S., Ng, K.Y., Lievens, F., & Van Dyne, L. (2015). Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *Journal of Applied Psychology, 100*, 464–480
- Roth, P. L., Bobko, P., & Buster, M. A. (2013). Situational judgment tests: The influence and importance of applicant status and targeted constructs on estimates of Black–White subgroup differences. *Journal of Occupational and Organizational Psychology, 86*, 394-409.
- 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*, 607–620.
- Schneider, B. (1987). The people make the place. *Personnel Psychology, 40*, 437–453.
- Sherman, R.A., Nave, C.S., & Funder, D. C. (2010). Situational similarity and personality predict behavioral consistency. *Journal of Personality and Social Psychology, 99*, 330-343.
- Sherman, R. A., Rauthmann, J. F., Brown, N. A., Serfass, D. G., & Jones, A. B. (2015). The independent effects of personality and situations on real-time expressions of behavior and emotion. *Journal of Personality and Social Psychology, 109*, 872.

- Shoda, Y., LeeTiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and distinctiveness from intra- and interpersonal interactions. *Personality and Social Psychology Review*, 6, 316–325.
- Smith, R.E., Shoda, Y., Cumming, S.P., & Smoll, F.L. (2009). Behavioral signatures at the ballpark: Intraindividual consistency of adults' situation-behavior patterns and their interpersonal consequences. *Journal of Research in Personality*, 43, 187–195.
- Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, 88, 500–517.
- Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. *Journal of Research in Personality*, 34, 397–423.
- Thornton, G. C., III, & Mueller-Hanson, R. A. (2004). *Developing organizational simulations: A guide for practitioners and students*. Mahwah, NJ: Erlbaum.

Table 1. A matrix of dynamic methods in selection procedures.

		Selection procedures		
		Personality questionnaires	Situational Judgement Tests	Assessment Centers
Dynamic concepts of personality	Within-person personality variability (<i>i.e., density distribution, traitedness, personality strength</i>)	<ul style="list-style-type: none"> Frequency-based response personality inventories (Edwards & Woehr, 2007) 	<ul style="list-style-type: none"> Construct-driven SJTs SJTs based on situational taxonomies to maximize situational variance per dimension 	<ul style="list-style-type: none"> Multiple speed assessment (Brannick, 2008; Herde & Lievens, 2016) Assessor ratings of candidates' consistency (Gibbons & Rupp 2009)
	Person – situation interactions (<i>i.e., CAPS model, trait activation theory</i>)	<ul style="list-style-type: none"> Work-related frame-of-reference inventories (Schmit, Ryan, Stierwalt & Powell, 1995) 	<ul style="list-style-type: none"> Verbal protocol analysis (Ployhart, 2006) Big Five of Personality in Occupational Situations (Ziegler, 2017) 	<ul style="list-style-type: none"> Assessor/coder ratings of situational factors Virtual adaptive simulations and games (Rayburn, 2007)