

# Future directions in personality, occupational and medical selection: myths, misunderstandings, measurement, and suggestions

Eamonn Ferguson<sup>1</sup> · Filip Lievens<sup>2</sup>

Received: 8 July 2016 / Accepted: 29 December 2016 / Published online: 20 February 2017  
© The Author(s) 2017. This article is published with open access at Springerlink.com

**Abstract** This paper has two objectives: (1) presenting recent advances in personality theory whereby personality traits are conceptualized within a framework that focuses on the dynamic interactions of behaviour, biology, context, and states, and (2) discussing the implications of these developments for measurement and medical selection. We start by presenting evidence that traits are no longer regarded as stable deterministic predictors of behaviour. Instead, traits are found to change across generations, the life span, and in response to environmental contingencies. Thus, there is an urgent need to explore how traits change as function of medical education. Second, drawing on recent theory and research (behavioural reaction norms and the density distribution model) we highlight evidence to show how the expression of trait relevant behaviour is dependent on context, and is distributed with an average (typical behaviour or personality) and a variance (plasticity or adaptability), with traditional personality measure associated with typical responding. Third, we demystify that some traits are better than others showing that so-called “good” traits have a dark-side. Fourth, we show how these developments impact on how personality might be assessed, thereby presenting recent evidence on the use of contextualized personality measures, situational judgment tests, other reports, and implicit measures. Throughout the paper, we outline the key implications of these developments for medical selection practices.

**Keywords** Personality · Personality change · Medical selection · Health · Five Factor Model · Behavioural reaction norms · Situational judgement tests · Contextualization · Implicit measures

---

✉ Eamonn Ferguson  
eamonn.ferguson@nottingham.ac.uk

Filip Lievens  
fillip.lievens@ugent.be

<sup>1</sup> School of Psychology, University of Nottingham, Nottingham NG7 2RD, UK

<sup>2</sup> Department of Personnel Management and Work and Organizational Psychology, Ghent University, Ghent, Belgium

## Introduction

Gough (1953) conducted probably one of the earliest investigations of the role of personality in medical selection and education. Undoubtedly, his aim, to “contribute to a broader understanding of some of the non-intellective factors relating to academic achievement” (p. 361), still reflects the traditional approach to personality in medical selection. That is: (1) to identify traits based on their ability to predict a future outcome and (2) to select people on these traits, usually selecting higher scores (Ferguson et al. 2002). However, recent advances in personality theory question the validity of this tradition with clear implications for medical selection practices. We address some of the myths/misunderstandings about personality that form the basis of the traditional model and suggest ways forward.

## Developments in the conceptualization of personality

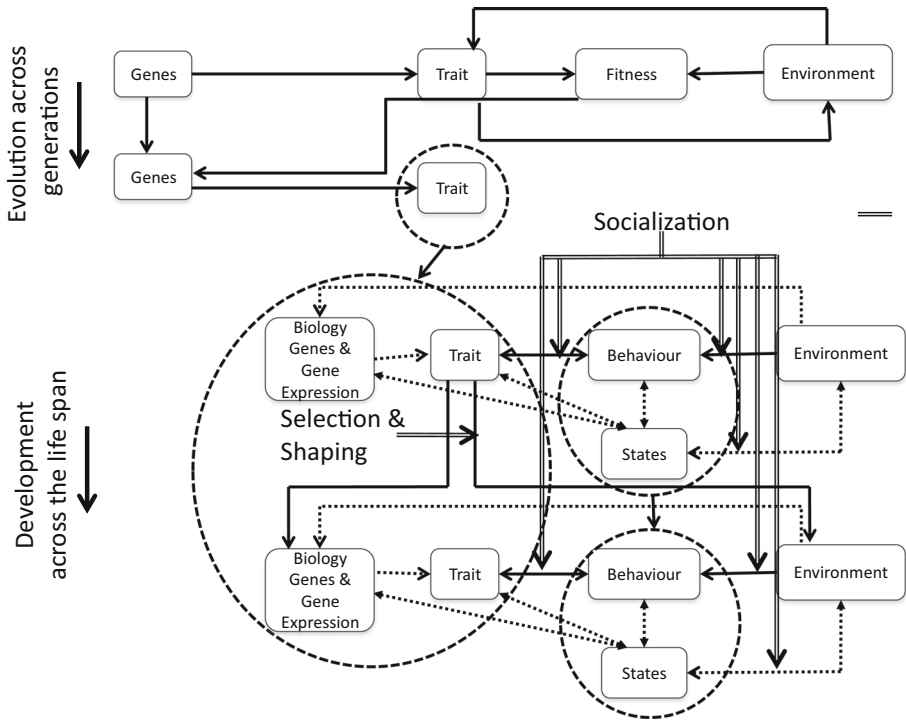
### Myth/misunderstanding 1: traits as stable deterministic predictors

Whereas it is a widely held belief that traits are stable deterministic predictors of behaviour (see Roberts 2009, for an overview), current thinking from ecology (Dingemans et al. 2010), economics (Ferguson et al. 2011) and personality psychology (Roberts and Jackson 2008) suggests the opposite. Traits are conceived as dynamically linked to context and biology, allowing for traits not only to change but also to influence the expression of behaviour (trait expression) across contexts.

#### *Change and stability*

Evidence demonstrates that traits change across generations (Twenge 2000, 2009), and the life course (Roberts et al. 2006; Caspi et al. 2005; Roberts and Mroczek 2008 for reviews) as well as a function of proximal events: (1) environmental exposure such as university, work, unemployment (Ludtke et al. 2011; Robins et al. 2005; Roberts et al. 2013; Boyce et al. 2015), (2) training (Jackson et al. 2012), and (3) therapeutic and psychological interventions to promote change (Tang et al. 2009; Hudson and Fraley 2015). The size of these changes can be substantial. It is of the same magnitude of economic markers of change like income and wealth (Boyce et al. 2013). The degree of personality change is itself subject to individual differences (Ludtke et al. 2011) and influenced by beliefs (e.g., about the changeability of personality, Robins et al. 2005).

Models such as the *socio-genomic model* of human personality aim to explain personality change within generations by highlighting the dynamic interaction between biology, environment, states (emotions, cognition, and beliefs), and traits (Roberts and Jackson 2008). According to this model, personality change is brought about by the influence of both (1) the environment on the organism’s biology as well as (2) states acting to mediate the link between biology and trait change. Similarly, both evo-devo models (a synthesis of evolutionary and developmental biology, whereby in some cases developmental ‘biases’ the adaptive quality of a phenotype: see Laland et al. 2011, p. 1514 and Toth and Robinson 2007) and niche construction theories (whereby organisms act as ‘eco-engineers’ and adapt and change their environments in such a way that the evolutionary gains are passed on: see Odling-Smee et al. 2013) indicate that traits change both within



**Fig. 1** Hybrid socio-genomic model of personality and evo-devo/niche construction model adapted and developed from Laland et al. (2011) and Roberts and Jackson (2008). The dotted lines represent the sociogenomic model (some of which is already in the evo-devo models). Development/change of a trait is influenced directly via biology (including gene expression). States (cognitions, emotions beliefs) also have an indirect influence on trait change via biology and behaviour and reciprocal links to the trait itself. The environment influence the type of behaviour expressed with the trait influencing this directly and via states. The trait can influence the selection and shaping of future environments. Socialization will then influence trait behavioural expression and the trait itself via environmental constraints and state processes that will influence the interpretation of behaviour in a given context. These developmental changes in personality may be selected for (if sexually desirable, or functionally adaptive for survival) and survive to the next generation

and across generations (Laland et al. 2011). These models suggest that traits lead organisms to both select environments to operate in as well as to change/shape these environments. All of which leads to trait change. Indeed selection, shaping, and socialization (traits change as a function of experiencing environments) have been observed for human personality (Ludtke et al. 2011; Wrzus et al. 2016; Caspi et al. 2005). There is also evidence for the “corresponsive principle”, whereby the trait that drives selection into an environment is the trait that changes the most from the selected environment (Roberts et al. 2013). As Caspi et al. (2005) succinctly put it “... the most likely effect of life experience on personality development is to deepen the characteristics that lead people to those experiences in the first place ... For example, if people assume more leadership positions because they are more dominant, then they will become more dominant through their experience as leaders.” (p. 470).

A hybrid model, combining these approaches (Fig. 1), indicates that traits, as well as being influenced by the environment via biology, also influence environments across

developmental time—we choose the environments we operate in, change them and our traits consequently change as a function of these experiences. Certain trait changes that increase fitness may be selected for and survive to the next generation.

### *Context specific prediction*

Whereas the developments above deal with changes in people's personality traits across the life span, within generations and to context, other developments emphasize how the *expression of behaviours associated with traits* (trait expression) changes within individuals across situations. Roberts (2009) suggests that traits represent a function that determines the probability that a person will act in a particular way given a particular context. This resonates with behavioural reaction norms (BRNs) used to conceptualize personality in ecology (Dingemanse et al. 2010). A BRN assesses trait relevant behaviours at multiple time points across a context that varies along some graded parameter (e.g., stress: contextual gradient). The mean value across context indicates how the organism *typically* behaves (personality) and variance/covariance across context denotes how the organism adapts and changes (plasticity).<sup>1</sup> Fleeson et al. (see Fleeson 2004; Fleeson and Law 2015) applied a similar approach to human personality: the *density-distribution* approach. Here multiple assessments of trait relevant behaviours (e.g., for conscientiousness: 'During the last half hour, how hardworking have you been?') are collected across contexts. Both personality and plasticity were observed (see Fig. 2). Furthermore, typical behaviour on 1 day predicted typical behaviour the next day (stability). In relation to assessment with traditional personality scales, Fleeson et al. showed that the distribution shifts to the right for those who score high (+1 SD), compared to low (−1 SD), on a standard personality scale (see Fig. 3). So, for example, high conscientious people will still express some low conscientiousness at times. Finally, scores on traditional personality scales were positively correlated with the mean and extreme of the distribution. Thus, individuals have a stable mean (typical behaviour), that is correlated with their score on a personality test, based on their own distribution of trait expression. In sum, personality reflects both typical behaviour, stability and plasticity across contexts, and traditional personality inventories pick up only typical behavioural tendencies but not plasticity.

### *Implications for medical selection research*

At present, there are no data on how personality changes as a function of medical training. Such an examination is important for predicting adaptability of medical students and physicians to a career in medicine. If medical training changes levels of conscientiousness for example, does it make the trainee/physician more or less conscientious? Changes to lower conscientiousness have been linked to poorer health (Human et al. 2012; Jokela et al. 2014) which will ultimately affect practitioner performance and the predictive validity of conscientiousness. The traditional approach does not consider personality change (Ferguson 2013), but it is clear that personality change needs to be part of any predictive model.

Trait expression and the context specific prediction from traits also implies that - depending on context—people high on a trait, such as conscientiousness, may be more or less likely to express that behaviour. For example, conscientiousness may be a good

---

<sup>1</sup> BRNs are formally described using multi-level models where the intercept gives the mean and slope the plasticity coefficient.

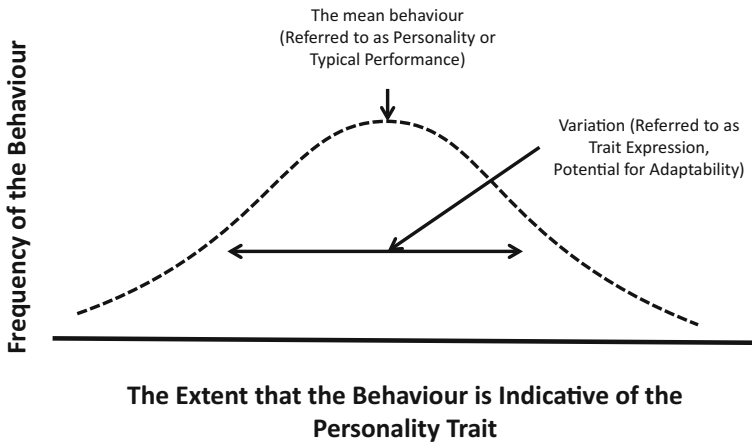


Fig. 2 Phenotype and phenotypic expression

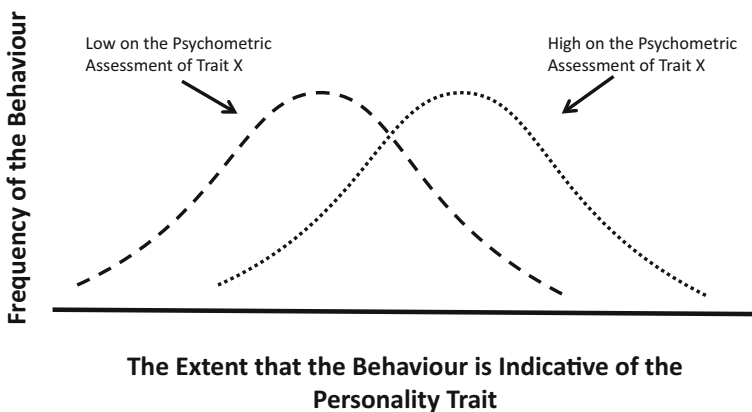


Fig. 3 Phenotype and phenotypic expression as a function of high and low trait scores

predictor in a context where high conscientiousness is expressible (e.g., MCQ exam, OSCE) but not so much in context where it is not (e.g., patient interaction). Thus, the predictive validity of the trait is context dependent and we need to not only know about differential prediction across the medical career but also how this influences and is influenced by personality change.

**Myth/misunderstanding 2: the dark and bright sides**

There is an assumption that some traits are “good” (conscientiousness) and others are “bad” (neuroticism). However, recent findings show that “good” traits have a “dark-side” and “bad” traits a “bright-side” (Boyce et al. 2010; Ferguson et al. 2014). For example, neuroticism has anxiety as a one of its facets, which brings benefits in terms of increased vigilance to danger, but carries a cost: negative reactions to stress (see Nettle 2006 for an evolutionary explanation). Developing this idea further, Widiger and Mullins-Sweatt (2008) suggest the Big 5 traits may be related to behaviours that are associated with being

either maladaptive or normal, with these further divided into high and low. Thus, maladaptive high conscientiousness (ridged, single minded), maladaptive low conscientiousness (careless, sloppy) and normal low conscientiousness (disorganized) would all hinder good performance, whereas normal high conscientiousness (organized, resourceful) would enhance it. This suggests that a curvilinear function links the trait to performance. For example, too little conscientiousness and too much conscientiousness may be problematic, whereas just enough conscientiousness (the “Goldilocks’ hypothesis”) is the optimal solution (Martin and Keyes 2015).

### *Implications for medical selection research*

The implication is that “good” traits that are typically seen as key to select in medicine, like conscientiousness, may carry a dark-side. Indeed, there is evidence for this in medicine (Ferguson et al. 2003, 2014). Ferguson et al. (2014) showed that conscientiousness was a negative predictor of clinical knowledge, but that a little anxiety aided clinical skills development. This idea can be extended to a trait like empathy, which is seen as a good trait to select on in medicine because it indicates a positive value and virtue required in medical staff. However, there are costs associated with empathy in terms of susceptibility to psychological illness, reduced pain thresholds, and psychopathology (see Ferguson 2016). Thus, while undoubtedly virtue and values, are key to a caring medical system, without considering the potential dark-side of empathy, and testing for it, selecting such a ‘good’ trait may be too simplistic and even counterproductive. Indeed, recent evidence shows that empathy in carers may carry costs (see Manczak et al. 2016). Furthermore, the Goldilocks’ hypothesis suggests that selection should focus on the optimal level of trait. The issue is how to identify that optimal level.<sup>2</sup>

### **Developments in personality measurement: some solutions for medical selection practice**

Our review above suggests that personality assessment should allow at the very least for some assessment of ‘contextual sensitivities’ or ‘trait expression’. In addition, there should be attention to both bright and dark sides of personality traits. Do new developments in personality assessments offer some immediate solution to these conceptual issues?

### **Contextualized measures of personality**

Contextualized personality inventories typically add a tag to existing generic items. Examples of such tags are “at-work” or “at-school” (Lievens et al. 2008). A recent meta-analysis in the employment domain (Shaffer and Postlethwaite 2012) revealed that adding such simple tags to a generic personality scale substantially raised the validity of the personality scores for predicting job performance.

Another approach is to focus on context specific motivations linked to a trait. This approach to contextualization is exemplified by the Five Individual Reaction Norms Inventory (FIRNI) “...which conceptualizes the Five-Factor Model dimensions as stable individual differences in people’s motivational reactions to circumscribed classes of environmental stimuli” (Denissen and Penke 2008: p. 1297). For example the FIRNI item “When I am acting on a plan I do not easily let myself be distracted by short-term needs”

<sup>2</sup> Trait change would also still need to be considered.

could be re-written as “When dealing with patients I would not easily let myself be distracted by short-term needs.”

### *Implications for medical selection research*

It may seem straightforward to add contextualized tags such as “when dealing with patients” to traits like conscientiousness (Jackson et al. 2010). However, these tags need to be carefully considered. For example, what contexts should be tagged? Also applicants (unless having had extensive work experience) would not be able to say how they would typically react in such a context (“When dealing with patients I am a very competent person”). Last, it should be noted that the specificity of the tags used should probably mirror the outcome measures that one wants to predict. That is, general outcome measures (life satisfaction, well-being) are best predicted with traditional generic personality inventories, whereas the opposite is true for more narrow and fine-grained outcome measures.

### **Trait expression**

Situational judgment tests (SJTs) offer even more possibilities for measuring personality in relation to situations. Situational judgment tests (SJTs) present realistic, job-related situations and ask participants to indicate what should be done to handle each situation effectively (McDaniel et al. 2007). Due to their predictive validity and diversity benefits SJTs have made inroads in medical education and selection (e.g., Libbrecht et al. 2014; Lievens and Patterson 2011).

A new development consists of using SJTs designed to infer people’s standing on personality (Motowidlo et al. 2006; Lievens and Motowidlo 2015). Such SJTs present a situation that activates a specific personality trait (e.g., agreeableness) and then lists response options that differ in terms of their level of agreeableness. The underlying logic is that people who score high on agreeableness will be better able to discriminate between the different options because they possess more accurate beliefs (referred to as implicit trait policies, Motowidlo et al. 2006), of what an effective agreeable reaction is in that given situation. Thus, implicit trait policies capture procedural knowledge about the behaviour linked to the trait that would be effective in that context.

### *Implications for medical selection research*

So far, SJTs assessing implicit trait policies have not been adopted in medical education (but see Ghosh et al. 2015). Such contextualized personality approaches, however, are promising ones that deserve more recognition and applications in medical education.

### **Other-reports and implicit personality measures**

Whereas the two aforementioned assessment approaches link to some of the key theoretical advancements highlighted in this paper, two other methods (other-reports and implicit personality measures) could complement these. Over the last decade, these two measures have received substantial research attention in the personality and personnel selection domains but not so in the medical education domain.

Other-reports of someone's personality, by a person well acquainted to the target across a diversity of situations, sheds light on someone's reputation (Hogan and Shelton 1998). Other-reports are typically obtained with the same inventories as self-reports and two meta-analyses showed that other-reports of personality are valid and substantially add incremental predictive validity over self-reports (Connelly and Ones 2010; Oh et al. 2011). The evidence reviewed above shows that self-reported personality can change in response to environmental contingencies and events. If an individual's self-reported personality changes with respect to environmental factors (e.g., life events, exam stress), the question is whether their well acquainted peers are able to pick this up and indicate similar changes in the targets personality, or is the change not necessarily apparent to observers. It may be that some people are able to manage their expression of personality relevant behaviour change so acquaintances do not pick it up whereas others do not. Difference between peer and self-reports of personality, with respect to change, may provide an index of reputation and/or impression management. At present, there is limited data on how changes in other-reports link to changes in self-reported personality (Jackson et al. 2009).

The key difference between implicit and explicit measures of personality is that the latter inventories (e.g., NEO-PI) ask someone directly to describe their personality, whereas the former infer someone's personality. For example, people are asked to make associations, complete sentences or arrange pictures and their responses are used to infer their personality (Uhlmann et al. 2012). Within a dual-system framework, implicit measures tap intuitive and fast processing, whereas traditional explicit measures assess slow and deliberative processing (Strack and Deutsch 2004). As such, implicit and explicit personality measures are tapping different processes. Indeed, the association between implicit and explicit measures of the Big 5 are low, but generally strongest for neuroticism, extraversion, and conscientiousness (see Grimm and von Collani 2007; Schmukle et al. 2008; Steffens and Konig 2006; Vianello et al. 2013). Implicit personality measurement has the potential to add extra information over and above explicit personality measures (Back et al. 2007; Lang et al. 2012; Uhlmann et al. 2012), predict spontaneous behaviours (Steffens and Konig 2006), and are less susceptible to faking good (Vecchione et al. 2014). We need to know the extent to which such implicit measures also show changes across and within generation and are influenced by context.

### *Implications for medical selection research*

In medical selection, other-reports based on Five-Factor Model inventories have not been used. Yet, they could be operationalized through references and letters of recommendation that are structured based on the Five-Factor Model of personality (for an example in personnel selection, see Taylor et al. 2004).

With respect to implicit measures, the main concern is often their practicability in actual high-stakes selection, as they involve lengthy assessments of computer based reaction times. However, one specific type of implicit measures, namely conditional reasoning tests, might offer a solution (Berry et al. 2010). Here test-takers are presented with situations and possible responses, which are interwoven with traditional reasoning items. They are asked to select the response that follows most logically from the situation. In a conditional reasoning test of aggression, for instance, the four responses refer to different justification mechanisms, with some of them indicating more aggressive tendencies of the endorser, with repeated endorsements of such response options indicative of aggressiveness. Reviews show that the predictive validity of conditional reasoning tests of aggression is similar to the one of self-report personality measures (Berry et al. 2010) and that they are



not susceptible to faking good unless people are told that the test measures aggressiveness instead of conditional reasoning (LeBreton et al. 2007). So, under these circumstances, conditional reasoning tests might be additions to the existing tests used in medical admissions.

Another way of assessing implicit personality is to develop a technique introduced by Quirin and colleagues (see Quirin and Bode 2014; Quirin et al. 2009). In this procedure participants indicate the extent to which words from an artificial language (e.g., Tunba) sound like a series of mood adjectives (happy, energetic, helpless, etc.). This could be easily developed via Goldberg’s adjective markers for the Big 5 (Goldberg 1992). So far, however, we do not know of any colleges using these tests in medical selection.

### Predictive validity issues and trait change

First and foremost the above shows that stable trait expression, as indexed by the score on a personality scale, can change as function of context. At present, there are no data on how personality changes as a function of medical education. This is a fundamental gap in our knowledge as the degree and extent of behaviour change has profound implications for predictive validity. If the level of a trait changes across medical training and career, as function of that training, then at the very least multiple assessments of the trait are needed to help to untangle cause and effect as training performance ( $TR$ ), at any one point, will be a function of the initial trait score  $t_n$  plus the trait assessment proximal to the performance estimate ( $t_{n+1}$ ), and the previous training environment ( $TR_{n-1}$ ). Previous training is used to model the ‘learning back-bone’ along which trait will have their effect (Ferguson et al. 2014). Furthermore, the level of the trait at ( $t_{n+1}$ ) will depend not only on the initial trait level ( $t_n$ ) but also the effect of any previous training environment ( $TR_{n-1}$ ). This results in the simple model in Fig. 4. Thus, predictive the validity coefficients can be expressed as ( $t_n$ ) and ( $t_{n+1}$ ) (these can be used to control for initial levels and residual change) controlling for previous training. Thus we can ask if (1) trait scores at selection ( $t_n$ ) predict across the training backbone equally (generalized predictive validity), or if the strength or direction changes, (2) if proximal assessments ( $t_{n+1}$ ) have stronger predictive power than distal ones or (3) if it is the degree of trait change that matters (if so it becomes important to identify predictors of change).

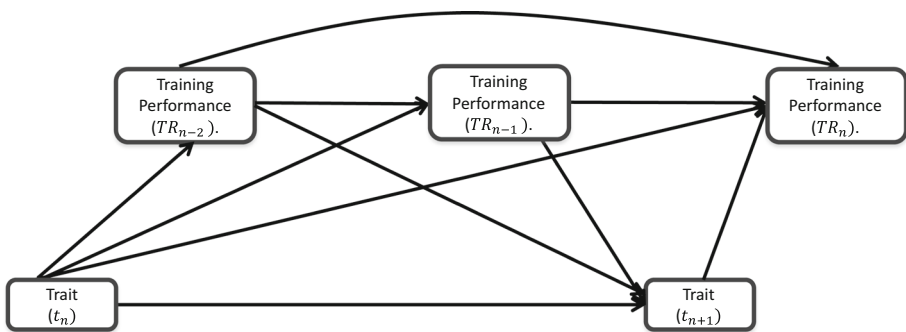


Fig. 4 Structural model for incorporating personality change into predictive validity

## Epilogue

There are three big gaps at this time in the literature of medical selection with respect to personality: (1) lack of data on personality change and its implications for healthcare professional's health and performance, (2) the lack of assessment of trait expression and context sensitivity, and (3) no real recognition of the 'bright' and 'dark' sides of personality and again its implications for healthcare professionals' performance. The intensity and nature of medical training is very likely to result in personality change (Ludtke et al. 2011), also the changing medical training context is going to mean that 'dark-side' and 'bright-side' aspects of traits will have particular roles to play with respect to specific aspects of training (Ferguson et al. 2014).

The overall implication is that the traditional model of "selecting on a trait that has overall positive predictive validity" is itself of limited validity in the face of dynamic trait change and context specificity of trait expression. At a minimum, if the traditional approach is used, then analyses should be conducted to explore if the trait has any 'dark-sides' in the context of medical selection. We also need to know how traits change in this context. If change is not factored into the predictive models, then these models may well artificially under- or over-estimate validity coefficients. Also, by knowing which traits change and for whom we will be in a position to develop much stronger and more precise predictive models. For example, we might select the level of the trait that shows the least change but the highest overall predictive value. Or we might identify those for whom greatest trait change is likely. If this change is likely to be detrimental interventions to help can be put into place (e.g., treating personality change and management as part of medical education as well as selection, Jackson et al. 2012; Hudson and Fraley 2015).

We also make a few recommendations concerning how current available personality testing approaches may allow, to a certain extent, for some of these conceptual issues to be incorporated. Our recommendations focus in particular around context specificity (SJT, FIRNI, etc.) and following them up will enhance the medical selection process. However, that said, the big questions of trait change, expression, and the dark-side still need to be explored.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Back, M. D., Schmukle, S. C., & Egloff, B. (2007). Predicting actual behavior from explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology, 97*, 533–548.
- Berry, C. M., Sackett, P. R., & Tobares, V. (2010). A meta-analysis of conditional reasoning tests of aggression. *Personnel Psychology, 63*, 361–384.
- Boyce, C. J., Wood, A. M., & Brown, G. D. A. (2010). The dark side of conscientiousness: Conscientious people experience greater drops in life satisfaction following unemployment. *Journal Research in Personality, 44*, 535–539.
- Boyce, C. J., Wood, A. M., Daly, M., & Sedikides, C. (2015). Personality change following unemployment. *Journal of Applied Psychology, 100*, 991–1011.
- Boyce, C. J., Wood, A. M., & Powdthavee, N. (2013). Is personality fixed? Personality changes as much as "variable" economic factors and more strongly predicts changes in life satisfaction. *Social Indicators Research, 111*, 287–305.

- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, *56*, 453–484.
- Connelly, B. S., & Ones, D. S. (2010). Another perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, *136*, 1092–1122.
- Denissen, J. J. A., & Penke, L. (2008). Motivational individual reaction norms underlying the Five-Factor Model of personality: First steps towards a theory-based conceptual framework. *Journal of Research in Personality*, *42*, 1285–1302.
- Dingemans, N. J., Kazem, A. J. N., Reale, D., & Wright, J. (2010). Behavioral reaction norms: Animal personality meets individual plasticity. *Trends in Ecology and Evolution*, *25*, 81–89.
- Ferguson, E. (2013). Personality is of central concern to understand health: Towards a theoretical model for health psychology. *Health Psychology Review*, *7*, S32–S70.
- Ferguson, E. (2016). Empathy: The good, the bad and the ugly. In A. Wood & J. Johnson (Eds.), *Positive clinical psychology: An integrative approach to studying and improving well-being* (pp. 103–125). London: Wiley.
- Ferguson, E., Heckman, J. J., & Corr, P. J. (2011). Personality and economics: Overview and proposed framework. *Personality and Individual Differences*, *51*, 201–209.
- Ferguson, E., James, D., & Madeley, L. (2002). Factors associated with success in medical school: Systematic review of the literature. *British Medical Journal*, *324*, 952–957.
- Ferguson, E., James, D., O'Hehir, F., & Sanders, A. (2003). A pilot study of the roles of personality, references and personal statements in relation to performance over the 5 years of a medical degree. *British Medical Journal*, *326*, 429–431.
- Ferguson, E., Semper, H., Yates, J., Fitzgerald, J. E., Skatova, A., & James, D. (2014). The 'dark side' and 'bright side' of personality: When too much conscientiousness and too little anxiety are detrimental to the acquisition of medical knowledge and skill. *PLoS ONE*, *9*(2), e8860. doi:[10.1371/journal.pone.0088606](https://doi.org/10.1371/journal.pone.0088606).
- Fleeson, W. (2004). Moving personality beyond the person-situation debate: The challenge and the opportunity of within-person variability. *Current Directions in Psychological Science*, *13*, 83–87.
- Fleeson, W., & Law, 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*, 1090–1104.
- Ghosh, K., Motowidlo, S. J., & Nath, S. (2015). Technical knowledge, prosocial knowledge, and clinical performance of Indian medical students. *International Journal of Selection and Assessment*, *23*, 59–70.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*, 26–42.
- Gough, H. G. (1953). The construction of a personality scale to predict scholastic achievement. *Journal of Applied Psychology*, *37*, 47–52.
- Grimm, M., & von Collani, G. (2007). Measuring Big-Five personality dimensions with the implicit association test—implicit personality traits or self-esteem. *Personality and Individual Differences*, *43*, 2205–2217.
- Hogan, R., & Shelton, D. (1998). A socioanalytic perspective on job performance. *Human Performance*, *11*, 129–144.
- Hudson, N. W., & Fraley, C. (2015). Volitional personality trait change: Can people choose to change their personality traits. *Journal of Personality and Social Psychology*, *109*, 490–507.
- Human, L. J., Beisanz, J. C., Miller, G. E., Chen, E., Lachman, M. E., & Seeman, T. E. (2012). Is change bad? Personality change is associated with poor psychological health and greater metabolic syndrome in midlife. *Journal of Personality*, *81*, 249–269.
- Jackson, J. J., Bogg, T., Walton, K. E., Wood, D., Harms, P. D., Lodi-Smith, J., et al. (2009). Not all conscientiousness scales change alike: a multimethod, multi-sample study of age differences in the facets of conscientiousness. *Journal of Personality and Social Psychology*, *96*, 446–459.
- Jackson, J. J., Thoenes, F., Jonkmann, K., Ludtke, O., & Trautwein, U. (2012). Military training and personality trait development: Does the military make the man, or does the man make the military? *Psychological Science*, *23*, 270–277.
- Jackson, J. J., Wood, D., Bogg, T., Walton, K. E., Harms, P. D., & Roberts, B. W. (2010). What do conscientious people do? Development and validation of the behavioral indicators of conscientiousness (BIC). *Journal of Research in Personality*, *44*, 501–511.
- Jokela, M., Hakulinen, C., Singh-Manoux, A., & Kivimaki, M. (2014). Personality change associated with chronic disease: Pooled analysis of four prospective cohort studies. *Psychological Medicine*, *44*, 2629–2640.
- Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W., & Uller, T. (2011). Cause and effect in biology revisited: Is Mayr's proximate-ultimate dichotomy still useful. *Science*, *334*, 1512–1516.

- Lang, J. W. B., Zettler, I., Ewen, C., & Hulsheger, U. R. (2012). Implicit motives, explicit traits, and task and contextual performance at work. *Journal of Applied Psychology, 97*, 1201–1217.
- LeBreton, J. M., Barksdale, C. D., Robin, J., & James, L. R. (2007). Measurement issues associated with conditional reasoning tests: Indirect measurement and test faking. *Journal of Applied Psychology, 92*, 1–16.
- Libbrecht, N., Lievens, F., Carette, B., & Côté, S. C. (2014). Emotional intelligence predicts success in medical school. *Emotion, 14*, 64–73.
- 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., & Motowidlo, S. J. (2015). Situational judgment tests: From measures of situational judgment to measures of general domain knowledge. *Industrial and Organizational Psychology—Perspectives on Science and Practice, 9*, 3–22.
- Lievens, F., & Patterson, F. (2011). The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced level high-stakes selection. *Journal of Applied Psychology, 96*, 927–940.
- Ludtke, O., Roberts, B. W., Trautwein, U., & Nagy, C. (2011). A random walk down university avenue: Life paths, life events, and personality trait change in the transition to university life. *Journal of Personality and Social Psychology, 101*, 620–637.
- Manczak, E. M., DeLongis, A., & Hen, E. (2016). Does empathy have a cost? Diverging psychological and physiological effects within families. *Health Psychology, 35*, 211–218.
- Martin, C. C., & Keyes, C. L. M. (2015). Investigating the Goldilocks hypothesis: The non-linear impact of trait change on well-being. *PLoS ONE, 10*(7), e0131316.
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology, 60*, 63–91.
- Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006). Implicit policies about relations between personality traits and behavioral effectiveness in situational judgment items. *Journal of Applied Psychology, 91*, 749–761.
- Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist, 61*, 622–631.
- Odling-Smee, J., Erwin, D. H., Palkovacs, E. P., Feldman, M. W., & Laland, K. N. (2013). Niche construction theory: A practical guide for ecologists. *Quarterly Review of Biology, 88*, 3–28.
- Oh, I. S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology, 96*, 762–773.
- Quirin, M., & Bode, R. C. (2014). An alternative to self-reports of trait and state Affect. *European Journal of Psychological Assessment, 30*, 231–237.
- Quirin, M., Kazen, M., & Kuhl, J. (2009). When nonsense sounds happy or helpless: The implicit positive and negative affect tests (IPANAT). *Journal of Personality and Social Psychology, 97*, 500–516.
- Roberts, B. W. (2009). Back to the future Personality and assessment and personality development. *Journal of Research in Personality, 43*, 137–145.
- Roberts, B. W., Caspi, A., & Moffitt, T. E. (2013). Work experience and personality development in young adults. *Journal of Personality and Social Psychology, 84*, 582–593.
- Roberts, B. W., & Jackson, J. J. (2008). Sociogenomic personality psychology. *Journal of Personality, 76*, 1523–1544.
- Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*, 31–35.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1–25.
- Robins, R. W., Noffle, E. E., Trzesniewski, K. H., & Roberts, B. W. (2005). Do people know how their personality has changed? Correlates of perceived and actual personality change in young adulthood. *Journal of Personality, 73*, 489–521.
- Schmukle, S. C., Back, M. J., & Egloff, B. (2008). Validity of the Five Factor Model for the implicit self-concept of personality. *European Journal of Psychological Assessment, 24*, 263–272.
- Shaffer, J. A., & Postlethwaite, B. E. (2012). A matter of context: A meta-analytic investigation of the relative validity of contextualized and noncontextualized personality measures. *Personnel Psychology, 65*, 445–494.
- Steffens, M. C., & Konig, S. S. (2006). Predicting spontaneous Big Five behavior with implicit association tests. *European Journal of Psychological Assessment, 22*, 13–20.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of behavior. *Personality and Social Psychology Review, 8*, 220–247.

- Tang, T. Z., DeRuberis, R. J., Hollon, S. D., Amsterdam, J., Shelton, R., & Schalet, B. (2009). Personality change during depression treatment: A placebo-controlled trial. *Archives of General Psychiatry*, *66*, 132–1330.
- Taylor, P. J., Pajo, K., Cheung, G. W., & Stringfield, P. (2004). Dimensionality and validity of a structured telephone reference check procedure. *Personnel Psychology*, *57*, 745–772.
- Toth, A. L., & Robinson, G. G. (2007). Evo-devo and the evolution of social behavior. *Trends in Genetics*, *23*, 334–341.
- Twenge, J. M. (2000). The age of anxiety? Birth cohort change in anxiety and neuroticism, 1952–1993. *Journal of Personality and Social Psychology*, *79*, 1007–1021.
- Twenge, J. M. (2009). Generational changes and their impact in the classroom: Teaching Generation Me. *Medical Education*, *43*, 398–405.
- Uhlmann, E. L., Leavitt, K., Menges, J. I., Koopman, J., Howe, M., & Johnson, R. E. (2012). Getting explicit about the implicit: A taxonomy of implicit measures and guide for their use in organizational research. *Organizational Research Methods*, *15*, 553–601.
- Vecchione, M., Dentale, F., Alessandri, G., & Barbaranelli, C. (2014). Fakability of implicit and explicit measures of the Big Five: Research from organizational settings. *International Journal of Selection and Assessment*, *22*, 211–218.
- Vianello, V., Schnabel, K., Sriram, N., & Nosek, B. (2013). Gender differences in implicit and explicit personality traits. *Personality and Individual Differences*, *55*, 994–999.
- Widiger, T. A., & Mullins-Sweatt, S. N. (2008). Five-Factor Model of personality disorder: A proposal for DSM-V. *Annual Review of Clinical Psychology*, *5*, 197–220.
- Wrzus, C., Wagner, G. G., & Riediger, M. (2016). Personality-situation transactions from adolescence to old age. *Journal of Personality and Social Psychology*. doi:[10.1037/pspp0000054](https://doi.org/10.1037/pspp0000054). (in press).