Original Article

The Berlin Multi-Facet Personality Inventory

An IPIP-Based Measure of Big Five Personality Facets

Victor Rouco[1], Anja Cengia[2], Richard Roberts[3], Christoph Kemper[4], and Matthias Ziegler[2]

[1]Department of Developmental, Personality and Social Psychology, Ghent University, Belgium
[2]Department of Psychology, Humboldt Universität zu Berlin, Germany
[3]Research and Assessment: Design Science Solution (RAD Science), Morrisville, PA, USA
[4]Psychology Department, HSD-Hochschule Döpfer, Cologne, Germany

Abstract. A novel personality inventory is presented in this article, named the Berlin Multi-Facet Personality Inventory. This new instrument is an adaptation of items from the International Personality Item Pool (Goldberg, 2006) aimed at a more comprehensive set of Big Five facets. This tool has been developed to comprise a large number of nonredundant facets below each of the Big Five domains. Two language versions of the same inventory have been developed (English and German) and tested for measurement invariance in order to facilitate international usability. In addition to the construction of the inventory, this work presents first evidence for the psychometric quality of its scores in two different populations across two different studies. The inventory is freely available online.

Keywords: personality, Big Five, facet models

Over the last decades, the Five Factor Model (Costa et al., 1992) as well as the Big Five (Goldberg, 1990) have become widely accepted models for describing general attributes of personality. Often the terms are used synonymously, which is why we will refer to the Big Five from here on. In both cases, personality is conceived as a hierarchical model which describes individual differences in personality at the dispositional level: one of the most basic, universal, biologically influenced, and stable layers of inter-individual differences in behavior, cognition, and feeling (McAdams et al., 2006). This hierarchical conception is relevant to acknowledge behavior from the most specific (nuances) to the more general (domains) differences in personality, through a varying number of mid-level characteristics (facets). Most of the research concerning criterion validity of scores from Big Five inventories has focused on the covariation between Big Five domain scores and relevant external outcomes. However, specific dispositional characteristics captured on the facet level might be of utility to provide more complex descriptions of individuality and to predict life outcomes to a major extent (Kretzschmar et al., 2018; Lounsbury et al., 2002; Paunonen et al., 2001; Ziegler, 2014; Ziegler et al., 2010). Unfortunately, the number and nature of facets below the Big Five domains is far from being consensual. In fact, many different sets of facets have been proposed. One potential reason for this proliferation could be that many facet-level models were developed as an elaboration or extension to an existing domain level measure. This ad-hoc inception has the disadvantage of potentially limiting the search space for possible facets. The current research project was conducted to overcome such limitations and to develop a facet set spanning an extensive behavior space.

Different Facet Models

As outlined above, there are several models that include a facet structure below the five broad domains. Among them, probably the most widely known is the one proposed by Costa and McCrae (1995), the NEO-PI-R model, which defines six facets per domain. Other equally popular models include the Big-Five-Inventory 2 (Soto et al., 2017), the International Personality Item Pool (IPIP; Goldberg et al., 2006), and the HEXACO model (Lee et al., 2016; which assumes six broad domains). A table is available in the electronic supplementary material (ESM 1), providing an overview of these different elaborations, and listing psychometric information such as internal consistency estimates and correlations with external constructs (see

Table E1 in ESM 1). Other models have also been developed, such as the Faceted Inventory of the Five-Factor Model by Watson et al. (2017), or the Big Five Aspect Scales by DeYoung et al. (2007).

The abundance of proposed facets related to each of the Big Five domains evidences the reality of a topic that is not consensual and that gives rise to many different elaborations. Despite these differences, however, there is still some degree of overlap. Soto et al. (2009) inspected the convergence between the NEO-PI-R and the first version of the BFI, suggesting that two constructs per domain were measured at the facet level by both inventories. The constructs defined by Soto et al. (2009) were: *Altruism* and *Compliance* for agreeableness; *Anxiety* and *Depression* for neuroticism; *Order* and *Self-Discipline* for conscientiousness; *Assertiveness* and *Activity* for extraversion; and *Aesthetics* and *Ideas* for openness. Likewise, the existence of such "core" constructs was also suggested by DeYoung et al. (2007), in what they termed aspects. Even though both contributions' labels vary, they have a substantial degree of similarity in terms of content. Furthermore, these core constructs are present not only in the models which Soto et al. (2009) analyzed, but also in all models listed in Table E1 in ESM 1. Some of these constructs are explicitly covered (e.g., *Anxiety*), while others are implicitly encompassed, which becomes obvious when given attention to the items' content (e.g., *Liveliness* in HEXACO resembles the "core" construct *Activity*, present in all other instruments).

While most models based on the Big Five include these "core" constructs, there is still an abundance of other constructs which could be termed "peripheral," whose inclusion is more variable. Soto and John (2017, page 118) referred to this diversity of constructs by saying that the Big Five dimensions "can be conceptualized and assessed more broadly or more narrowly," either by focusing on central or on peripheral facets, depending on the interest of research. It is due to this multiplicity of peripheral constructs that an effort to convey different models would be beneficial to obtain a comprehensive inventory which subsumes a maximum of these peripheral facets, as well as including the essential core facets.

An important step toward building such an inventory is to ensure that the proposed set of facets predict consequential outcomes. One of the most classical approaches to relate the measured constructs with external constructs is to define a nomological network between the personality traits and external outcomes (Cronbach & Meehl, 1955). Relying on dimensions to describe behavior and to predict external outcomes can benefit from ease of interpretability. However, the domain level is sometimes too distal to depict behavioral mechanisms underlying personality-to-outcome associations thoroughly. Using nuances to predict behavior might yield a stronger predictive power (Seeboth et al., 2018), as specificity to situations and contexts is enhanced (Ziegler et al., 2016). Nonetheless, using nuances in the prediction of external outcomes can have the disadvantage of dealing with extreme complexity. Facets are, by definition, in a middle ground between nuances and dimensions, representing a compromise between specificity and sensitivity in the bandwidth-fidelity dilemma.

Facets Associated With Consequential Outcomes

The question of whether narrow measures have a superior predictive power over broad measures has enjoyed a lively debate. Importantly, the idea here is not that a single facet score outperforms a domain score. Rather than that, based on the finding that facet scores can have differential criterion correlations, sometimes even differing in direction (e.g., Ziegler et al., 2010), we assumed that a linear combination of facet scores could explain more criterion variance compared to the domain score. The idea here is that each facet comprises some specific variance which might improve the amount of explained variance (Ziegler & Bäckström, 2016). To test this, we compared the multiple R^2 of the facet scores predicting criteria with the bivariate r^2 for the domain score. Research summarized below suggests an advantage for scores derived from more narrow measures. However, there is also research pointing in the opposite direction (Chen, 2012; Salgado, 2017). The following section provides an overview of evidence for relations between domains, facets, and three consequential outcomes that are of interest to researchers.

Satisfaction With Life

One of the outcomes that has been largely evidenced to be predicted by personality is satisfaction with life (SWL). There is a consensus on emotional stability and extraversion as the most important domains when predicting SWL (Heller et al., 2004; Steel et al., 2008). At the facet level, Steel et al. (2008) and Schimmack et al. (2002) identified the NEO facets of *Depression* and *Positive Emotions* as the most important predictors, with medium sized regression coefficients. In line with these findings, we hypothesized that the set of facets which measure emotional stability and extraversion in our inventory should significantly correlate with SWL, with a moderate to large effect size.

Academic Performance

Another relevant outcome predicted by personality scores is academic achievement. Conscientiousness has been labeled the strongest predictor (Poropat, 2009; 2014). At the facet level, De Fruyt et al. (1996) hypothesized that facets of conscientiousness related to volition would predict academic achievement more strongly. There is a collection of research that is consistent with this idea, linking academic performance with facets such as *Achievement-Striving* (Gray et al., 2002; O'Connor et al., 2007) or *Work Drive* (Lounsbury et al., 2002). Nonetheless, other conscientiousness facets more related to duty or moral behavior have been found to predict grade point average (GPA): for instance, *Self-Discipline* (Gray et al., 2002; O'Connor et al., 2007), or *Dutifulness* (O'Connor et al., 2007). The relation of academic achievement with openness at the domain level has been more variant, in part because the facets of openness can be related in opposite directions with this outcome. Paunonen et al. (2001) found that the openness facet *Understanding* correlates with academic achievement. Noftle et al. (2007) identified a set of openness facets which predicted academic achievement (the HEXACO facets of *Aesthetic*, *Inquisitiveness*, *Creativity* and *Unconventionality*, plus the NEO-PI-R facets of *Fantasy*, *Aesthetics*, *Feelings,* and *Ideas*, also see Ziegler et al., 2010). In line with these findings, we hypothesized that the facet scores which entail the conscientiousness domain in our inventory would correlate with academic performance, and that openness to experience sscores would yield a mixed pattern at the facet level.

Academic Absenteeism

Personality has also been reported as a robust predictor of absenteeism, especially with an inverse correlation between this behavior and conscientiousness scores (Judge et al., 1997; Ones et al., 2003). Some specific facets of conscientiousness have been highlighted, like *Work Drive* (Lounsbury et al., 2004) or *Need for Achievement* (Wegge et al., 1993). Chamorro-Premuzic and Furnham (2003) as well as Furnham and Medhurst (1995) found significant correlations with openness, while correlations with extraversion were reported by Judge et al. (1997). The relation of absenteeism with personality at the domain level has been more vague. We expect that this ambiguity could be resolved by switching the focus to the facet level.

This Research

Attending to the multiplicity of facet models available in the personality literature and the resurgence of narrow constructs as relevant units of analysis to describe personality and to predict important life outcomes, we strove to develop a comprehensive measure of personality facets which covers a broad content of the Big Five domains.

This novel measure is an adaptation of items from the IPIP (Goldberg, 2006) aimed at a more concise measure of personality traits with a focus on narrow facets. An antecedent to this study can be found in MacCann et al. (2009), where part of the stimuli set presented here was analyzed to design a measure of conscientiousness containing a maximum number of relevant facets. The current research extends MacCann et al. (2009) to all Big Five domains. Furthermore, the inventory presented here has been developed with the scope of being open source, by making it available to researchers and practitioners at no cost; and with the intention of facilitating internationally usage by testing its applicability in two different cultures. Furthermore, the same taxonomy has been operationalized in another study, albeit with a different set of contextualized items (Ziegler et al., 2019). In that study, the authors provide empirical support for convergent and discriminant validity of their test scores.

It must be noted that proposing such a higher number of facets requires evidence to support the usefulness of each facet (Siegling et al., 2015). At the same time, it is also clear that such an endeavor must be considered as permanent work in progress. Within this paper we will test the idea that facets outperform domain scores with regard to the amount of criterion variance explained. Moreover, we will use the approach suggested by Rosenbusch et al. (2020) and test the content similarity of our facet items with scales uploaded in a data bank. Importantly, this data bank contains the IPIP scales and keys. This way we can avoid jingle-jangle fallacies.

In order to support its applicability among cultures, this research uses two independent samples from two different countries: the USA and Germany. In the first study, using the US sample, we empirically defined a facet model, selected items, and tested the facet scores' psychometric properties with regard to factorial validity evidence, internal consistency, and test-criterion correlations. In Study 2, we replicated the findings with a German sample, and further tested the measurement invariance (MI) of the suggested models.

Study 1

Methods

Participants

This sample consisted of 722 American undergraduate students who gave voluntary acceptance to their inclusion

in this research. Their mean age was 21.60 years (*SD* = 5.90), and 59% of them were female. Students were e-mailed a link to a computerized assessment battery that included the items as well as several other cognitive tests not focused in this paper. The data set was randomly split into two equally sized subsamples. Both subsamples were matched in relation to missing values, outliers, and extreme values. Subsample 1's mean age was 21.80 years (*SD* = 6.30), subsample 2's mean age was 21.50 years (*SD* = 5.60).

Measures

Altogether, 528 items from the IPIP were used in this study, as indicators of the Big Five domains. The IPIP is an open-source database of personality items, launched in 1996 and containing over 2,000 items (Goldberg et al., 2006). Participants were asked to provide self-ratings for the items on a 5-point rating scale, ranging from 1 (*not all like me*) to 5 (*very much like me*).

The item set used originates from a complex item reduction conducted before. Here, all IPIP items were rated by experts regarding their prototypicality for the Big Five in general and each domain specifically. Based on these ratings, the items for the current study were selected and represent the most prototypical items for each domain. More details can be found in MacCann et al. (2009).

Satisfaction With Life

SWL was measured with a 5-item composite defined in Diener et al. (1985), using a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Items included are (a) "In most ways my life is close to ideal," (b) "The conditions of my life are excellent," (c) "I am satisfied with my life," (d) "So far I have gotten the important things in my life," and (e) "If I could live my life over, I would change almost nothing." In our sample, the reliability estimate was $\alpha = .88$.

Grade Point Average

To measure academic performance, participants reported their GPA scores at the end of high school.

Absences

Participants reported an estimation of days they were absent from college without justification. This was an item extracted from a larger set of student social behavior indicators (MacCann et al., 2009). Absences were log transformed prior to analyses, as proposed by Lounsbury et al. (2004).

Statistical Analyses

Analyses in this study were conducted applying a split-sample method to separate allow an exploratory and confirmatory phase. Several exploratory factor analyses (EFA) were conducted with Subsample 1, to identify the number of facets underlying the personality items from each domain. Subsample 2 was used to estimate reliabilities by means of internal consistency, to conduct tests of unidimensionality by applying confirmatory factor analysis (CFA) to each facet, and to obtain evidence for structural validity for the combination of all facet scores by exploratory structural equation modeling (ESEM). Based on prior research (Bäckström et al., 2009; Ziegler & Bühner, 2009) we specified a bifactor ESEM where the bifactor reflects variance due to socially desirable responding. We also used an oblique target rotation to maximize loadings on the respective domains and to minimize loadings on other domains (McCrae et al., 1996). Finally, a test of relations derived from the nomological network was conducted with the complete sample to maximize parameter precision and power. We elaborated the following hypotheses regarding our nomological network.

Hypothesis 1: SWL will be predicted by facet scores of emotional stability mimicking NEO-PI-R *depression*, and facet scores of extraversion covering *positive emotions*, in line with Schimmack et al. (2002). Emotional stability and extraversion scores will be the strongest predictors in the personality-SWL association at the domain level.

Hypothesis 2: The conscientiousness score will be associated with academic achievement. Openness will entail facet scores with positive effects and facet scores with negative effects on GPA scores.

Hypothesis 3: The conscientiousness score will yield the strongest associations with absenteeism at the domain level, as compared with the other four domains, and facet scores tapping volitional components such as *goal orientation* or *wish to work* will reveal such relations on facet level.

A more detailed description of the statistical analysis can be found in ESM 1.

Results

Exploratory Factor Analyses

Exploratory analyses revealed that the domains could be structured into 8 to 11 facets. Model fit information for the EFA procedure, as well as Eigenvalues, and results from the MAP tests and parallel analyses, are presented in ESM 1 (Table E2). To ensure the homogeneity of the facets and to reduce the risk of cross domain loadings, items with factor loadings of less than .30 which in addition had non-

central content to the domain in question were excluded (Ziegler, 2014).

Eight facets were retained for the domain agreeableness, after two were eliminated due to weak loadings and clusters whose content was difficult to interpret. These facets were named *Appreciation*, *Integrity*, *Low Competitiveness*, *Readiness to Give Feedback*, *Search for Support*, *Trust*, *Genuineness,* and *Altruism*. Items corresponding to these and the following facets can be found in ESM 1.

Conscientiousness was defined by nine facets after one facet with factor loadings below .30 was excluded, these were: *Dominance*, *Persistence*, *Self-Discipline*, *Planfulness*, *Goal Orientation*, *Carefulness*, *Orderliness*, *Wish to Work* (to capacity), and *Initiative*.

Extraversion was formed by nine facets. A new facet (*Energy*) was added in order to tap the physical component of extraversion, which was missing in the eight-facet solution the EFA suggested. These facets were labeled *Sociability*, *Readiness to Take Risks*, *Wish for Affiliation*, *Cheerfulness*, *Assertiveness*, *Communicativeness*, *Humor*, *Gregariousness,* and *Energy*.

Neuroticism (interpreted here as emotional stability) consisted of seven facets. One facet was dropped due to poor interpretability and was therefore not included in the subsequent analyses. The final set of facets were named *Patience*, *Confidence*, *Carefreeness*, *Toughness*, *Drive*, *Emotional Robustness,* and *Public Self-Consciousness*.

Openness to experience could be split into nine facets. One facet was identified as a method factor and eliminated, as it solely contained negatively formulated items and no coherent underlying trait could be identified. Furthermore, an extra facet was added (*Intellect*), as it was not present in the EFA solution and represents a core construct in other important facet models. The final set of facets of the openness domain were named *Creativity*, *Adventurousness*, *Open-Mindedness*, *Interest in Reading*, *Culture*, *Curiosity*, *Willingness to Learn*, *Empathy,* and *Intellect*.

Reliability

Reliability estimates for each of the facets and all domains were obtained using Cronbach's α (Cronbach, 1951) and McDonald's ω (McDonald, 1999). 95% CI estimates of McDonald's ω for the domain scores were: agreeableness ranged from .85 to .91, conscientiousness ranged from .83 to .88, openness ranged from .91 to .94, emotional stability ranged from .90 to .93, extraversion ranged from .89 to .92. All in all, reliability coefficients were at least good ($\omega > .8$) for all the domains, and at least acceptable ($\omega > .7$) for the majority of the facets (60%). Only one facet had poor internal consistency (*Altruism*, $\omega = .52$. All internal consistency estimates for the facets can be found in Table 1.

Confirmatory Factor Analysis

CFA was applied to each of the facets identified in the previous step, using Subsample 2. All measurement models fitted well according to goodness-of-fit indices. The fit information of three facets was not available as these models were reflected by only three indicators and therefore just identified (*Energy*, *Public Self-Consciousness*, and *Intellect*). Goodness of fit estimations for each facet are available in Table 1.

Exploratory Structural Equation Modeling

The final ESEM model was constructed after removing four facets that were not significantly loaded by their intended domain: *Empathy* (expected to load on openness), *Search for Support* and *Readiness to Give Feedback* (expected to load on agreeableness), and *readiness to take risks* (expected to load on extraversion). Furthermore, four residuals were allowed to be correlated: *Planfulness* and *Carefulness*, *Goal-Orientation* and *Wish to Work*, *Confidence* and *Toughness*, *Confidence* and *Goal-Orientation*. The addition of these correlated residuals was consistent with the facet's content and revealed that a significant amount of specific variance was still present in the facets. The resulting model included 38 facet scores, all of them with significant loadings in their intended domains. It yielded a model fit of χ^2 (*df*) = 1,287.09 (486), CFI = .88, RMSEA = .065, SRMR = .038. We consider this model fit as sufficient to approximate our data. The standardized loadings can be found in Table 1. As it is usual in ESEM procedures, some facets had cross-loadings from other domains all of which were below $\lambda = .36$, though. Factor intercorrelations were also all smaller than $r = .34$ (Openness and Extraversion).

Nomological Network

Assumptions from a nomological network with our proposed set of facets and three external criteria were tested to provide evidence of criterion validity. Table E3 in ESM 1 summarizes these findings and highlights that, overall, the hypotheses outlined in the methods section were met. H1 stated that emotional stability and extraversion should be the domains which had stronger associations with SWL. Indeed, both domain scores showed the highest correlations and their facets accounted for the most variance explained. At the facet level, *Confidence* and *Cheerfulness* (both $\beta > .5$) were strongly associated with the criterion, also in line with what was hypothesized in H1. H2 stated that conscientiousness would be associated with GPA with a medium to small effect size and that openness would yield a mixed pattern of association at the facet level. Conscientiousness was, in line with previous research, the domain with higher associations with SWL, with a correlation of $r = .3$. Three openness facets were significantly

Table 1. Internal consistency, CFA model fit, and ESEM standardized loadings in the intended domain

	IC		CFA				ESEM
	α	ω	χ^2 (*df*)	*p*	CFI	RMSEA	λ std*
Agreeableness	—	.86	—	—	—	—	—
Appreciation	.71	.73	15.02 (5)	.01	.99	.08	0.38
Integrity	.72	.74	3.17 (5)	.67	1	0	0.64
Low Competitiveness	.72	.72	1.99 (5)	.85	1	0	0.76
Good Faith	.65	.69	33.59 (5)	< .001	.97	.13	0.23
Genuineness	.65	.68	5.5 (5)	.36	1	.02	0.64
Altruism	.52	.56	0.37 (2)	.83	1	0	0.35
Conscientiousness	—	.88	—	—	—	—	—
Dominance	.71	.73	38.45 (5)	< .001	.93	.14	0.27
Persistence	.57	.62	19.72 (5)	< .001	.98	.09	0.32
Self-Discipline	.68	.69	13.62 (5)	.02	.98	.07	0.30
Task Planning	.81	.81	5.66 (5)	.34	1	.02	0.82
Goal Orientation	.77	.77	13.6 (5)	.02	.99	.07	0.68
Carefulness	.68	.69	12.94 (5)	.02	.98	.07	0.58
Orderliness	.82	.83	25.64 (5)	< .001	.99	.11	0.46
Wish to Work to Capacity	.63	.67	10.41 (5)	.06	.99	.06	0.35
Productivity	.68	.69	12.17 (5)	.03	.98	.06	0.40
Extraversion	—	.90	—	—	—	—	—
Sociability	.66	.68	13.27 (5)	.02	.99	.07	0.75
Wish for Affiliation	.65	.68	16.52 (5)	.01	.98	.08	0.69
Positive Attitude	.82	.83	1.75 (5)	.88	1	0	0.55
Forcefulness	.68	.70	20.94 (5)	< .001	.97	.09	0.20
Communicativeness	.75	.75	18.27 (5)	< .001	.98	.09	0.70
Humor	.79	.79	18.77 (5)	< .001	.99	.09	0.29
Conviviality	.69	.71	14.89 (5)	.01	.98	.07	0.74
Energy	.71	.74	0 (0)	< .001	1	0	0.49
Emotional Stability	—	.90	—	—	—	—	—
Equanimity	.74	.75	9.38 (5)	.09	1	.05	0.39
Mental Balance	.86	.86	10.02 (5)	.07	.99	.05	0.54
Carefreeness	.77	.77	8.46 (5)	.13	1	.04	0.76
Confidence	.70	.71	8.2 (5)	.15	1	.04	0.41
Drive	.62	.64	13.21 (5)	.02	.98	.07	0.59
Emotional Robustness	.75	.76	13.6 (5)	.02	.99	.07	0.73
Self-Attention	.60	.63	0 (0)	< .001	1	0	0.63
Openness	—	.92	—	—	—	—	—
Creativity	.68	.68	17.19 (5)	< .001	.98	.08	0.81
Wish for Variety	.70	.72	9.96 (5)	.08	1	.05	0.42
Open-Mindedness	.66	.67	19.17 (5)	< .001	.98	.09	0.77
Interest in Reading	.85	.86	5.79 (5)	.33	1	.02	0.54
Artistic Interests	.81	.82	18.32 (5)	< .001	.99	.09	0.59
Wish to Analyze	.78	.79	11.04 (5)	.05	.99	.06	0.78
Willingness to Learn	.81	.82	8.03 (5)	.15	1	.04	0.71
Intellect	.80	.81	0 (0)	< .001	1	0	0.62

Note. IC = Internal Consistency. * = all factor loadings are significant with $p < .05$.

associated with the outcome: *Creativity* was inversely associated (β = −.14), while *Interest in Reading* and *Intellect* had a positive association (β = .13 and .18). Thus, our nomological network showed the expected properties regarding H2. In addition, two facets of agreeableness were linked with high school GPA, they were *Low Competitiveness* (β = −.13) and *Genuineness* (β = .14), as well as emotional stability's *Patience* (β = .13). H3 stated that conscientiousness should be the domain yielding highest associations with absenteeism, and that conscientiousness' facets related to volitional aspects would highlight this association. Indeed, conscientiousness had the strongest associations with absenteeism (in an inverse relation, $r = -.28$), and two facets related to volition, *Planfulness* (β = −.14) and *Initiative* (β = −.13), were significantly associated with this outcome. Furthermore, H3 stated that the relation of absenteeism with other domains will be clearer at the facet level. Here, we only found modest associations at the domain level, but some facets like *Genuineness* (β = −.15), *Readiness to Take Risks* (β = .18), *Energy* (β = −.18), *Willingness to Learn* (β = −.15), or *Drive* (β = −.14) were significantly associated with the outcome, thereby confirming H3. Importantly, the variance explained by the linear combination of the facet scores mostly outweighed the variance explained by the single domain score, strongly supporting the usefulness of the facet scores.

Content Analysis

The linguistic similarity check (available in ESM 1) revealed that some of the facets we propose, are already suggested in the IPIP keys. At the same time, there are also several facets that are not in those keys. In order to avoid jingle-jangle-fallacies, we included the results of our similarity check in ESM 2.

Study 2

Procedure and Participants

Study 2 was conducted with an independent sample of 387 German speakers (49.10% male) with a mean age of 45.60 years (SD = 17.50). Data collection was done by a private company in their test facilities. Participants were paid for their participation. Participants were targeted to achieve representativeness with regard to age, gender, and education level.

Measures

For the German version of the presented tool, the IPIP items selected in Study 1 were translated and back-translated by bilingual speakers. Nonmatching back-translations were flagged as inadequate and were further adapted by the same experts.

Data Analysis

Similarly as in Study 1, internal consistency and structural validity by means of one CFA model per facet and one bifactor ESEM model for the full inventory was estimated with the German sample. In addition, we present in this section a MI procedure used to test the equivalency of the measurement models in the two countries. To this end, the German sample was combined with the US Subsample 2, used for confirmatory analyses in Study 1.

Measurement Invariance

Following tests for structural validity mirroring the procedures from Study 1, MI tests were conducted for each facet using the German and the US data. Three levels of MI were analyzed here: configural, metric, and scalar invariance. Model comparisons were based on suggestions by Chen (2007). Metric invariance was accepted whenever Δ CFI < .01, Δ RMSEA < .015 or Δ SRMR < .03; and scalar invariance whenever Δ CFI < .01, Δ RMSEA < .015 or Δ SRMR < .01. MI for the full model was tested using ESEM.

In addition to full invariance tests, partial invariance was also tested at the facet level. Partial invariance was investigated by allowing a maximum of two factor loadings (for metric invariance) or intercepts (for scalar invariance) to differ between language versions. The robust maximum likelihood estimator was used.

Results

MI at the Facet Level

Configural invariance was found for all facet models. This was the highest degree of invariance obtained for one facet, *Readiness to Give Feedback*, a facet of agreeableness. A relatively high number of facets reached partial scalar invariance (48.9%) after freeing a maximum of two intercepts in their respective models. One facet reached full scalar invariance: *Sociability*, a facet of extraversion. Furthermore, the vast majority of facets showed at least partial metric invariance between both countries (97.6%; see Table E5 in ESM 1).

As stated above, the partially invariant solutions were tested after freeing a maximum of two parameters between groups in each MI stage (see Table E7 in ESM 1). As a general trend, noninvariant factor loadings of extraversion are larger in the US sample than in the German

sample, as well as positive factor loadings of openness. Conversely, negative factor loadings of openness are bigger in the German sample. Some freed factor loadings seem to deviate due to age differences between samples, as "I resist authority," an indicator of *Dominance*, and "I am easily discouraged," an indicator of *Persistence*, both yielding higher factor loadings in the younger, German sample. At the scalar level, most noninvariant parameters of conscientiousness' facets had higher intercepts for Germans, except for those of *Persistence*, which were higher in the American sample. Nonetheless, *Persistence* indicators may be higher in the American sample due to age differences ("I like to take my time" and "I never give up" are examples of items belonging to this facet). In fact, many of the non-scalar invariant indicators may be best understood by the effect of age; for instance: "I act impulsively when something is bothering me," "I am easily talked into doing silly things," "I get overwhelmed by emotions," all of these yielding higher intercepts in the younger group.

MI of the Full Model

The bifactor ESEM model with the German sample showed similar fit as with the American sample (χ^2 (*df*) = 1,218.84 (486), CFI = .90, RMSEA = 0.065, SRMR = 0.032). Importantly, all facets loaded significantly on their intended domains. The MI approach revealed that configural invariance was tenable in the integrated model (χ^2 *(df)* = 2,505.93 (972), CFI = 0.892, RMSEA = 0.065, SRMR = 0.035). Importantly, while factor congruence was at least moderate for the Big five domain loadings (all larger than .5), it was negative for the socially desirable responding (SDR) factor. Thus, we assumed that full metric invariance will not be achievable. Considering that partial invariance for ESEM is currently not available, we decided to go ahead with the invariance tests. The results show that the CFI decreased as expected for the metric model (CFI = .861) but increased for the scalar model (.863) which supports the tentative conclusion that MI might be present at the structural level, when SDR loadings are excluded. Also, there were only negligible and nonsignificant latent mean differences (all < |.01|).

Discussion

The personality test presented here, named Berlin Multi-Facet-Personality Inventory (BMF-PI), was developed to cover the need for a tool maximizing the coverage of facets within the Big Five framework. Starting from a large item pool, we have developed a questionnaire which assesses 38 facets with 202 items. The selected facets cover central constructs which are present in most Big Five models that include facet levels, as well as more peripheral constructs which could help to describe individual differences in a more specific manner. First evidence for reliability, construct and test-criterion validity of the facet scores is promising. In addition, the BMF-PI has been developed to enable cross-cultural research and to align with the principles of open accessibility, ensuring that researchers worldwide can benefit from this tool. These analyses also revealed interesting patterns of noninvariance, potentially informing cross-cultural research. For instance, on facet level, higher factor loadings were found in the American sample for positively keyed items in the openness domain, whereas inversely keyed items loaded highly in German's openness. This suggests that the indicators which reflect a low level of openness have a relatively higher influence in how this domain is defined by Germans. Also, noninvariant intercepts in conscientiousness were consistently higher in the German sample, indicating a higher baseline level in this domain for this group. On a structural level, the results suggest partial scalar invariance. Here, the SDR factor seems to differ between cultures, already on the metric invariance level.

Facet Structure

The instrument presented in this work covers the "core" facets proposed by Soto and John (2009, 2017). In some cases, these core constructs have been labeled similarly to the proposal of the cited authors, that is the case for *Energy*, *Altruism*, *Orderliness*, and *Self-Discipline*. In some other cases, our proposed labels were different as these were defined even more narrowly than in Soto and John (2009). That is the case for *Toughness* (instead of *Anxiety*), *Emotional Robustness* (*Depression*), or *Culture (Aesthetics)*. In the remainder of cases, the core constructs were represented by more than a single facet, to account for nuances in facets that we believe are more heterogeneous. This is the case of *Low Competitiveness* and *Integrity*, which both can be thought as related to *Assertiveness*; or *Open-Mindedness* and *Curiosity*, which are tapping the *Ideas* component of the openness domain.

Using a faceted instrument requires specific information within the facet scores (Ziegler & Bäckström, 2016). Here we provide initial evidence that the facet scores we propose are relevant with regard to test-criterion correlations. Moreover, based on linguistic similarity we could show that several of the facets we propose reflect facets already specified in the IPIP framework. This was expected and should be seen as support of the newly proposed taxonomy as it reveals overlap with existing other taxonomies. However, the similarity check also underscores the potential of the newly proposed facets revealing low similarity with other established personality measures in many

cases. Some of the facets may be highly specific, for example, interest in reading. Prior research has shown, though, that this facet plays an important role in the relation between Openness and cognitive ability (Trapp & Ziegler, 2019). Despite these encouraging findings it is mandatory that future research further investigates the structural robustness as well as the potential for test criterion correlations of all proposed facets.

Psychometric Properties

After defining the facet structure with an independent sample, we tested psychometric properties in terms of internal consistency and structural validity. We found good internal consistencies on facet score level, with 67% greater than .70 and 95% $\omega > .60$. The domain scores were also reliably measured, with ω ranging from .83 to .92.

Structural validity was assessed by fitting a CFA to each of the facets. Goodness of fit measures signaled that the data is consistent with the facet models, with 88% of the chi-square tests yielding nonsignificant results and all facets at least approximately fitting the data according to goodness-of-fit indices. These results suggest that the facets included in the BMF-PI can be used independently, in case that researchers and practitioners are more interested in a specific set of facets rather than in the full Big Five picture.

The higher order structure was tested using bifactor ESEM, a method which allows to overcome the constraint of independent clusters solution usually imposed in a CFA. Although some researchers could argue that such constraints are beneficial to ensure a high degree of discriminant validity, the independent cluster solution may be too strict to model constructs that are highly interrelated, as it often occurs in personality traits, at least on item or behavior level (Horstmann & Ziegler, 2020). A control procedure to deduce a reasonable degree of convergent and discriminant validity in the ESEM solution is to verify that the primary factor loadings are distinct from zero, and that the proportion of significant cross-loadings is not excessive. All of this was the case for the current data in both languages.

Association With External Constructs

We have tested different hypotheses which aimed to replicate previous findings on the interplay between personality and SWL, academic performance, and school absenteeism. Overall, the scores derived here have shown similar test-criterion-correlations as scores from other personality instruments, at least regarding the external outcomes that were included in our measures. Confirming our first hypothesis we found that *Cheerfulness* and *Toughness* (akin the inverse of *Depression*) were associated with the outcome with a medium effect size ($\beta = .51$ and $\beta = .59$, respectively). Emotional stability and extraversion were, as expected, the domains which showed the strongest links to SWL. The other domains did not account for a big piece of variance of SWL (R^2 ranging from .08 and .13), although we did find some interesting associations at the facet level. For instance, *Integrity*, *Good Faith*, *Persistence*, *Confidence*, *Open-Mindedness*, and *Intellect* were significantly linked with SWL, highlighting the usefulness of a rich set of facets when inspecting associations with external outcomes.

Our second hypothesis stated that conscientiousness would be associated with academic achievement with a small to moderate effect size, and that openness would result in a heterogeneous structure of direct and inverse effects at the facet level. Our results were in line with the hypothesis. Conscientiousness' sum score was correlated with high school GPA, while at the facet level *Goal Orientation* was the only construct predicting the outcome. Openness was also related to GPA, and as hypothesized, some facets were positively related to academic achievement, as *Interest in Reading* and *Intellect*; while others were negatively related, like *Creativity*. Looking at the nomological network from a facet perspective would give us a more complete picture of the associations with academic achievement by identifying important constructs such as *Low Competitiveness* (inversely related), or *Genuineness* and *Patience* (directly related).

The third hypothesis stated domains would not be directly linked with absenteeism in high school, but that some facets would picture this association more clearly. In line with this hypothesis, all R^2's were modest (R^2 ranging from .03 to .1), but some specific facets were significantly associated with the outcome, such as agreeableness' *Genuineness*, conscientiousness' *Planfulness* and *Initiative*, extraversion's *Energy* and *Humor*, emotional stability's *Drive*, and openness' *Willingness to Learn*. These results highlight that, in order to better predict educational absenteeism, researchers should focus on narrow constructs instead of on broader domains.

All in all, our nomological network suggest a reasonable degree of construct validity for the BMF-PI. Notwithstanding, future studies combining data from the BMF-PI with other instruments in the field should focus more strongly on a more comprehensive test of the nomological network than was feasible here. For example, a comparison with other facetted measures is necessary to avoid a false increase of facets. For instance, although support for the test criterion validity evidence of several facet scores related to agreeableness and openness has been provided here, more information on the predictive validity of these

domain scores with other outcomes is needed to further support the application of the BMF-PI. Furthermore, our results suggest that switching the focus from a domain perspective toward a facet perspective may increase the strength of associations between personality scores and external criteria.

International Usage

This instrument aims to be usable in different countries to promote internationalization of individual differences research. To this aim, it has been tested in two different languages, with samples gathered from two countries in two different continents. We have applied MI techniques to test the extent to which both versions of the inventory are equivalent. At the facet level, all of the facet models yielded configural invariance, and the vast majority reached partial metric invariance. Scalar invariance was attained for one facet. Nonetheless, it is important to notice here that the two populations were different in terms of age (effect size of the difference $d = 1.83$). Given that personality traits' structure and mean levels change within the lifespan of individuals, this age difference might have affected our invariance tests. For instance, Roberts et al. (2006) reported in their meta-analysis that the mean levels of *Dominance* increase from late childhood into late adolescence, reaching a plateau at an approximate age of 35. Our results are in line with this finding. The facet score for *Dominance* is not scalar invariant, and only partial metric invariance could be established. Future studies should investigate normative change as a possible confounder in the results presented in this paper, as our two samples' mean age differed significantly. Despite this mean age difference, the degree of MI obtained by both versions of the inventory has been satisfying, indicating that cross-cultural research is feasible with the inventory presented here. On structural level, the bifactor ESEM results shed further light onto the findings. Here, it can be assumed that invariance might be present in the personality loadings and intercepts but not the SDR loadings. This, of course, could also explain the noninvariance on facet level. It is also plausible to assume that SDR differs between the two cultural samples. Future research should further explore the role of SDR when comparing the structure of personality measures.

Limitations

Given that this study is the first using the BMF-PI, all findings are preliminary and subject to replication. More studies are needed with heterogeneous samples to test whether the found structure can be retrieved from other populations. We should bear in mind that some facets had to be excluded from the final proposal due to deficits in their integration within the Big Five framework, this problem could arise in other populations. One important limitation was sample size with regard to ESEM models which could not be fitted using a second order structure. Instead, we were forced to use the facet scores as indicators of the domains. By doing so, specific item variance might be insufficiently reflected in the ESEM models. Currently, the BMF-PI is only available in two languages. We hope to see a bigger dissemination through different cultures to further extent its international usability. Although personality traits seem to be universal, a certain degree of cultural variation does exist. Further adaptations of this inventory will reveal the extent to which the chosen indicators relate to the proposed set of facets in countries distinct from the USA or Germany. Furthermore, the demographics of the two samples differed considerably. Thus, the MI findings should be considered as very promising because they occurred despite differences in other variables which might impact the structure. Also, forthcoming studies using data collected from independent sampling would add a necessary layer of validity to our conclusions, as the split-sample method used here yields samples sharing demographic characteristics and are therefore quite homogenic. Thus, further independent replication is necessary.

Finally, several psychometric properties of the scores have not yet been tested, such as temporal stability or convergent and discriminant validity with scores from other Big Five instruments. Still, the semantic check that we have conducted in this study can be regarded as a hint supporting the notion of convergent and discriminant validity. Also, a previous study has examined these with the same taxonomy, although using a different set of items (Ziegler et al., 2019). Future research should also examine a more extensive nomological network, especially attending to the domains agreeableness and openness.

Despite these problems, we are confident that the current results have the potential to inform future research and thereby improve our understanding of the important facet structure beneath the Big Five.

Electronic Supplementary Material

The electronic supplementary material is available with the online version of the article at https://doi.org/10.1027/2698-1866/a000021

ESM 1. Big Five models, Study 1 data analysis procedure, results, nomological network, Study 2 results
ESM 2. Scale redundancy overview

References

Bäckström, M., Björklund, F., & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality*, *43*, 335–344.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, *37*(4), 319–338. https://doi.org/10.1016/S0092-6566(02)00578-0

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J.-P., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality*, *80*(1), 219–251. https://doi.org/10.1111/j.1467-6494.2011.00739.x

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences*, *13*(6), 653–665. https://doi.org/10.1016/0191-8869(92)90236-i

Costa Jr., P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the revised NEO personality inventory. *Journal of Personality Assessment*, *64*(1), 21–50. https://doi.org/10.1207/s15327752jpa6401_2

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/bf02310555

Cronbach, L., J., & Meehl, P., E. (1995). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

De Fruyt, F., & Mervielde, I. (1996). Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality*, *10*(5), 405–425. https://doi.org/10.1002/(sici)1099-0984(199612)10:5<405::aid-per255>3.0.co;2-m

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*(5), 880–896. https://doi.org/10.1037/0022-3514.93.5.880

Furnham, A., & Medhurst, S. (1995). Personality correlates of academic seminar behaviour: A study of four instruments. *Personality and Individual Differences*, *19*(2), 197–208. https://doi.org/10.1016/0191-8869(95)00026-3

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84–96. https://doi.org/10.1016/j.jrp.2005.08.007

Gray, E., & Watson, D. (2002). General and specific traits of personality and their relation to sleep and academic performance. *Journal of Personality*, *70*(2), 177–206. https://doi.org/10.1111/1467-6494.05002

Heller, D., Watson, D., & Ilies, R. (2004). The role of person versus situation in life satisfaction: A critical examination. *Psychological Bulletin*, *130*(4), 574–600. https://doi.org/10.1037/0033-2909.130.4.574

Horstmann, K. T., & Ziegler, M. (2020). Assessing personality states: What to consider when constructing personality state measures. *European Journal of Personality*, *34*(6), 1037–1059. https://doi.org/10.1002/per.2266

Judge, T. A., Martocchio, J. J., & Thoresen, C. J. (1997). Five-factor model of personality and employee absence. *Journal of Applied Psychology*, *82*(5), 11. https://doi.org/10.1037/0021-9010.82.5.745

Kretzschmar, A., Spengler, M., Schubert, A.-L., Steinmayr, R., & Ziegler, M. (2018). The relation of personality and intelligence—what can the Brunswik symmetry principle tell us? *Journal of Intelligence*, *6*(3), 30. https://doi.org/10.3390/jintelligence6030030

Lee, K., & Ashton, M. C. (2016). The HEXACO model of personality structure. In K. Lee, & M. C. Ashton, Eds, *Psychometric properties of the HEXACO-100. Assessment* (pp. 1–6). doi: https://doi.org/10.1007/978-3-319-28099-8_1227-1

Lounsbury, J. W., Sundstrom, E., Loveland, J. L., & Gibson, L. W. (2002). Broad versus narrow personality traits in predicting academic performance of adolescents. *Learning and Individual Differences*, *14*(1), 67–77.

Lounsbury, J. W., Steel, R. P., Loveland, J. M., & Gibson, L. W. (2004). An investigation of personality traits in relation to adolescent school absenteeism. *Journal of Youth and Adolescence*, *33*(5), 457–466. https://doi.org/10.1023/b:joyo.0000037637.20329.97

MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences*, *19*(4), 451–458. https://doi.org/10.1016/j.lindif.2009.03.007.

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: Fundamental principles for an integrative science of personality. *American Psychologist*, *61*(3), 204–217. https://doi.org/10.1037/0003-066x.61.3.204.

McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO personality inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, *70*(3), 552–566. https://doi.org/10.1037/0022-3514.70.3.552

McDonald, R. P. (1999). *Test theory: A unified treatment*. Psychology Press.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, *93*(1), 116–130. https://doi.org/10.1037/0022-3514.93.1.116

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, *43*(5), 971–990. https://doi.org/10.1016/j.paid.2007.03.017

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (2003). Personality and absenteeism: A meta-analysis of integrity tests. *European Journal of Personality*, *17*(S1), S19–S38. https://doi.org/10.1002/per.487

Paunonen, S. V., & Ashton, M. C. (2001). Big Five predictors of academic achievement. *Journal of Research in Personality*, *35*(1), 78–90. https://doi.org/10.1006/jrpe.2000.2309

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338. https://doi.org/10.1037/a0014996

Poropat, A. E. (2014). A meta-analysis of adult-rated child personality and academic performance in primary education. *British Journal of Educational Psychology*, *84*(2), 239–252. https://doi.org/10.1111/bjep.12019

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), 1–25. https://doi.org/10.1037/0033-2909.132.1.1

Rosenbusch, H., Wanders, F., & Pit, I. L. (2020). The Semantic Scale Network: An online tool to detect semantic overlap of psychological scales and prevent scale redundancies. *Psychological Methods*, *25*(3), 380–392. https://doi.org/10.1037/met0000244

Salgado, J. F. (2017). *Bandwidth-fidelity dilemma*. Encyclopedia of Personality and Individual Differences.

Schimmack, U., Diener, E., & Oishi, S. (2002). Life-satisfaction is a momentary judgment and a stable personality characteristic: The use of chronically accessible and stable sources. *Journal of Personality*, *70*(3), 345–384. https://doi.org/10.1111/1467-6494.05008

Seeboth, A., & Mõttus, R. (2018). Successful explanations start with accurate descriptions: Questionnaire items as personality markers for more accurate predictions. *Journal of Personality*, *32*(3), 186–201. https://doi.org/10.1002/per.2147

Siegling, A. B., Petrides, K. V., & Martskvishvili, K. (2015). An examination of a new psychometric method for optimizing multi–faceted assessment instruments in the context of trait emotional intelligence. *European Journal of Personality*, *29*(1), 42–54. https://doi.org/10.1002/per.1976

Soto, C. J., & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality*, *43*(1), 84–90. https://doi.org/10.1016/j.jrp.2008.10.002

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin*, *134*(1), 138–161. https://doi.org/10.1037/0033-2909.134.1.138

Trapp, S., & Ziegler, M. (2019). How openness enriches the environment: Read more. *Frontiers in Psychology*, *10*, 1123. https://doi.org/10.3389/fpsyg.2019.01123

Watson, D., Nus, E., & Wu, K. D. (2017). Development and validation of the faceted inventory of the five-factor model (FI-FFM). *Assessment*, *26*(1), 17–44. https://doi.org/10.1177/1073191117711022

Wegge, J., & Kleinbeck, U. (1993). Motivationale Faktoren betrieblicher Fehlzeiten: zum Einfluß leistungs-und anschlußthematischer Variablen auf die Abwesenheit am Arbeitsplatz [Motivational factors of company absenteeism: the influence of performance and follow-up-related variables on absence from work]. *Zeitschrift für experimentelle und angewandte Psychologie*, *40*(3), 451–486.

Ziegler, M. (2014). Comments on item selection procedures. *European Journal of Psychological Assessment*, *30*(1), 1–2. https://doi.org/10.1027/1015-5759/a000196

Ziegler, M., & Bäckström, M. (2016). 50 facets of a trait – 50 ways to mess up? *European Journal of Psychological Assessment*, *32*(2), 105–110. https://doi.org/10.1027/1015-5759/a000372

Ziegler, M., & Bühner, M. (2009). Modeling socially desirable responding and its effects. *Educational and Psychological Measurement*, 69, 548–565.

Ziegler, M., & Brunner, M. (2016). Test standards and psychometric modeling. In A. A. Lipnevich, F. Preckel, & R. D. Roberts (Eds.), *The Springer series on human exceptionality. Psychosocial skills and school systems in the 21st century: Theory, research, and practice* (pp. 29–55). Springer International Publishing. https://doi.org/10.1007/978-3-319-28606-8_2

Ziegler, M., Danay, E., Schölmerich, F., & Bühner, M. (2010). Predicting academic success with the Big 5 rated from different points of view: Self-rated, other rated and faked. *European Journal of Personality*, *24*(July), 341–355. https://doi.org/10.1002/per.753

Ziegler, M., Horstmann, K. T., & Ziegler, J. (2019). Personality in situations: Going beyond the OCEAN and introducing the situation five. *Psychological Assessment*, *31*(4), 567–580. https://doi.org/10.1037/pas0000654

History
Received February 22, 2021
Revision received December 10, 2021
Accepted December 17, 2021
Published online April 1, 2022

Section: Personality

ORCID
Victor Rouco
https://orcid.org/0000-0003-2604-3482
Matthias Ziegler
https://orcid.org/0000-0003-4994-9519

Matthias Ziegler
Department of Psychology
Humboldt University
Rudower Chaussee 18
12489 Berlin
Germany
zieglema@hu-berlin.de