Beyond the Big Five factors: using facets and nuances for enhanced prediction in life outcomes

Maiken Due Nielsen[1] · Petri Kajonius[1]

Abstract
Research on personality traits predicting life outcomes has typically been investigated using the Big Five factors and only occasionally their facets. However, recent research suggests that the use of items (reflecting personality nuances) can account for more predictive variance. The aim of the present study was to examine the predictive validity for various life outcomes comparing the hierarchical levels (factors, facets, and nuances) of the personality trait structure. These were measured using one of the publicly available instruments, IPIP-NEO-120, in a Swedish sample ($N = 440$). Confirmatory Factor Analyses (CFA) were performed to confirm the structures of the Big Five levels, and we used Elastic Net Regressions (ENR; with 10-fold cross-validation and shrinkage parameter), trained and applied for prediction in two separate samples. The results showed that nuances (item-level models) on average provided greater explained variance (34%) than facets (22.5%) and factors (12%) for all six life outcome predictions. Findings suggest that there may be psychometric value to using the lowest item-level of personality trait measurements. Implications, limitations, and directions for future research are discussed.

Keywords Personality traits · IPIP-NEO · Nuances · Items · Life outcomes · Facets

The field of personality psychology explores how individual differences relate to personal life outcomes. Obtaining a comprehensive understanding of how such individual differences impact the trajectory of life is paramount in gaining insights into human behaviour. While research on the five broad personality factors, also commonly referred to as domains, (Goldberg, 1990) is ubiquitous and replicable (Soto, 2019), there is a clear-cut case for the increased use of narrower traits (i.e., facets and items) (Seeboth & Mõttus, 2018) for improved predictive power and, subsequently, a more detailed understanding of personality (Stewart et al., 2022). As an example, individuals high in the personality factor conscientiousness report greater work satisfaction, whereas individuals high in extraversion and agreeableness report greater social satisfaction (Olaru et al., 2023). However, little is understood as to why these broad factors are associated with work and social satisfaction or how broad traits relate to life outcomes in general. It is possible that there are more informed relationships underlying such associations, visible only in the measurement of the narrower traits (i.e., facets and items), which makes up the Big Five domains. The present study set out to compare explained variance between factors, facets, and items (markers for nuances), predicting various life outcomes with a long, publicly available instrument (IPIP-NEO-120; Johnson, 2014) in a non-English sample.

Levels of personality (factors, facets and nuances)

Most commonly, personality-outcome associations are studied using the Five Factor Model (FFM; McCrae & John, 1992) or the Big Five (Goldberg, 1990), which consists of five broad traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism. These have been associated with a variety of life outcomes, such as divorce, career success, or even mortality (Roberts et al., 2007; Soto, 2019). Personality plays a role in advanced professions (Raoust et al., 2023). Personality factors are highly heritable (Polderman et al., 2015), stable across adulthood (Mõttus et al., 2017), and predict various life outcomes, often better than other factors such as socio-economic status (Kajonius & Carlander, 2017).

✉ Maiken Due Nielsen
Maikendue@hotmail.com

1 Department of Psychology, Lund University, Lund, Sweden

The Big Five factors, which are considered the highest level of the personality hierarchy, are composed of narrower traits known as facets (Johnson, 2014). For instance, neuroticism comprises the facet traits anger, self-consciousness, immoderation, and vulnerability (IPIP-NEO-120; Johnson, 2014). Similar to the Big Five traits, facets have been demonstrated to be heritable and stable (Briley & Tucker-Drob, 2012) while also capturing additional variance beyond factors (Mõttus et al., 2014). This implies that facets represent enduring traits with capacity to capture important variation in personality, which might go unnoticed when aggregating to the domain level (Briley & Tucker-Drob, 2012). As such, people can obtain identical scores for broad factors, while their facet scores could be drastically different from one another. For example, the meta-analysis by Vainik et al. (2019) reported on the relationship between personality and Body-Mass Index (BMI) as a measure of obesity. While BMI was positively linked to neuroticism and negatively to conscientiousness, the analysis revealed relationships between BMI and 15 facets across all five factors. Moreover, it was found that facets explained four times more variance than factors for obesity. Interestingly, the factor neuroticism was associated with higher BMI, which could suggest a role of increased anxious behavior. However, when inspecting BMI-facet associations, BMI was associated foremost with higher impulsiveness, anger, and hostility. Notably, while BMI was also positively associated with facets of extraversion, such as warmth and assertiveness, the effect of factor-level extraversion on BMI was inconsistent. Similar studies have found that facets provide a more precise perspective on the relationship between personality and life outcomes (Espinoza et al., 2023; Schwaba et al., 2019). By utilizing narrower traits, Vainik et al. (2019) facilitated a more comprehensive understanding of the relationship between personality and obesity, demonstrating the ability of narrow traits to reveal more meaningful relationships.

While the usefulness of facets in personality-outcome prediction has been demonstrated, it has been suggested that there is yet another indispensable level of traits below facets, labeled nuances (McCrae, 2015). Due to the lack of classification, nuances are operationalized as individual items in a personality instrument. However, personality nuances and items may not be *the same*. Nuances refer to the smallest unique aspects of an individual's personality (McCrae, 2015), which can be represented by a single item as well as sets of items that capture no distinct information from one another (Stewart et al., 2022). Items are standardized statements designed to capture tiny aspects of personality and are often treated as interchangeable within item-pools (Mõttus et al., 2017). For instance, extraversion can contain items such as "I have a lot of fun" and "I look at the bright side of life." While these items capture the facet cheerfulness and tend to correlate with each other, they contain unique information on their own. "I look at the bright sight of life" may capture a nuance labeled "positive thinking," whereas "I have a lot of fun" may reflect a nuance labeled "lively" or "playful." While treated as interchangeable in personality instruments, items capture distinct nuances of personality (Speer et al., 2022). Indeed, nuances have been found to capture variance beyond that of facets and share some of the properties of higher-order personality traits; stability, heritability, and observability across raters (Mõttus et al., 2017).

In a comprehensive study on the predictive value of nuances, Seeboth and Mõttus (2018) surveyed a British sample ($N = 8719$) with 40 life outcomes and 50 items based on the International Personality Item Pool (IPIP). They found that on average, nuances explained more variance than factors, and outperformed factors in predicting 37 of the 40 outcomes. Even after dropping the 10 most predictive items, nuances outperformed factor models.

Building further on similar findings, Stewart et al. (2022) examined associations using the Big Five Inventory-2 (BFI-2; Soto & John, 2017) to predict 53 life outcomes using factor, facet, and nuance models in a large US sample ($N = 6126$). Nuances (20.9%) outperformed both facets (18%) and factors (16.6%) across outcomes. This held true even after removing factor and facet variance from items, suggesting that unique item-variance was behind associations between personality and outcomes.

The present study

Given that facets and nuances account for unique variance and have shown to outperform factors in predictions (Seeboth & Mõttus, 2018; Stewart et al., 2022), there is a want for research using various samples and other instruments to determine their potential. Extending on previous studies, using a different sample and different outcomes, is of importance, as it helps validate and strengthen the reliability of research findings, while also ensuring that previous findings are not only the result of flukes or biases (Soto, 2019). Moreover, there are ongoing calls for a bottom-up taxonomy and instruments for measuring nuances (Condon et al., 2020), which all require a critical mass of empirical studies. There is also a need for analyzing various lengths of instruments in trait-outcome association studies, as a substantial proportion of informative personality variance is left unaccounted for by short scales (Sleep et al., 2021).

The aim of the present study was to compare the predictive validity of the three hierarchical levels in personality models (factor-, facet-, and item-level) in predicting life outcomes. We posed two hypotheses: (1) facet-level models will outperform factor-level models in predictions, and (2) item-level models, nuances, will outperform facet- and factor- level models in life outcome predictions. The purpose of the study was to extend on the works by Seeboth and Mõttus

Table 1 Life outcome measures used in the present study

Outcome label	Item question
Job satisfaction	How much would you say you enjoy your current work/job?
Social satisfaction	How much would you say you are satisfied with your current social life/friends?
Empathy	To what extent do you empathize and take part in others' feelings?
Bright future	How much would you say you believe your personal future will be bright?
Intrinsic reward	To what extent is personal development your source of motivation for working?
Extrinsic reward	To what extent is reward your source of motivation for working?

(2018) and Stewart et al. (2022), using a new sample, different outcomes, and a different personality instrument, and thus contributing to a growing body of literature on the value of using all levels of personality.

Methods

Participants and procedure

The sample used in the present study is a convenience sample. It consists of individuals who voluntarily participated in the research by accessing and completing an online questionnaire for personality testing in Swedish. All gave their informed consent, and since the collected data was anonymous and voluntary, no ethical review was required. In total, $N = 568$ participants completed the survey. However, participants who skipped $\geq 10\%$ of items ($n = 52$) were excluded from analysis. Also, participants below 25 years of age ($n = 76$) were excluded from the study due to the low longitudinal stability of personality during childhood, which significantly increases as individuals transition into adulthood (Briley & Tucker-Drob, 2014), as well as the potential lack of sufficient life outcome experience. The final sample size was $N = 440$ (57% female), ranging from 25 to 65 years old ($M = 42.0$, $SD = 10.5$). See Supplemental Materials and Appendix A for available demographics.

Measurements

IPIP-NEO-120 This is a widely used self-report instrument that assesses the Big Five with 5 factors, 30 facets, and 120 items (Johnson, 2014). The instrument is one of the few extensive and publicly free instruments with very high reliabilities ($\alpha > .85$). See Kajonius and Johnson (2019) for more psychometric properties. Respondents rate items on a five-point Likert scale ranging from 1 = *very inaccurate* to 5 = *very accurate*, with balanced (+ and -) keying. Example items: "I make friends easily" (extraversion), "I love to help others" (agreeableness), "I usually leave a mess in my room" (reversed conscientiousness), "I am not bothered by difficult social situations" (reversed neuroticism), and "I have a lively imagination" (openness). Item scores are

summarized into facet traits with four items per facet. Facet traits are summed and averaged into factors, with six facets per factor. See Appendix B for a complete overview of the IPIP-NEO-120 scale.

Life outcomes Personal life outcomes were measured by six self-rated single items. These were chosen as a battery of questions concerning having a balanced life, in terms of one's job, relationships, having space to empathize with others, having a positive outlook on future endeavors, and the incentives of working. Such were considered tapping into healthy life outcomes. See Table 1 for questions and labels used in the present study. Questions were answered using a seven-point Likert scale ranging from 1 = *not at all* to 7 = *completely*.

Statistical analyses

All statistical analyses were performed using R (R Core Team, 2021). The packages Psych (Revelle, 2021), dplyr (Wickham et al., 2023a), tidyr (Wickham et al., 2023b), lavaan (Rosseel, 2012), semPlot (Epskamp, 2022), ggplot2 (Wickham, 2016), reshape2 (Wickham, 2007), caTools (Tuszynski, 2021), tidyverse (Wickham et al., 2019), glmnet (Friedman et al., 2010), lme4 (Bates et al., 2015), and caret (Kuhn, 2022) were utilized for analyses and plots. Prior to all analyses, missing variables were dealt with using the R package mice (van Buuren & Groothuis-Oudshoorn, 2011). In this study, the mean value of each column was computed and inserted in the place of missing variables.

First, descriptive statistics on factor-, facet-, and item-level were analyzed. Second, we analyzed the structural validity of each of the Big Five hierarchical structures, using second-order Confirmatory Factor Analysis (CFA) in a SEM framework. Five models were tested, with respective Big Five factor at the top, loaded by respective six facets which in turn loaded by 24 items (4 items for each facet). See Appendix E for the visual structure of the models. No modification indices were used. Model fits were reported using point estimate values of the Standardized Root Mean Residuals (SRMR), robust Root Mean Squared Error of Approximation (RMSEA), and Comparative Fit Index (CFI) (Cheung & Rensvold, 2002). While conventional

model fit criteria often constitute at least the following: CFI ≥ .90, RMSEA ≥ .06 and SRMR ≥ .06 (Brown, 2014), low model fit in the context of personality measures is not necessarily unacceptable (Hopwood & Donnellan, 2010). The multidimensional and multifaceted nature of personality makes it challenging to precisely define the relationships among traits and items. Specifically, the complexity of using CFA for personality measures arises from challenges related to specifying accurate models and dealing with correlated residuals. Consequently, the conventional criteria for model fit are seen as guiding principles, rather than rigid cut-off criteria. Third, bivariate zero-order correlational analyses between Big Five factors and six life outcomes were performed. Fourth and last, Elastic Net Regression (ENR) was conducted to compare the predictive validities of factor, facet, and nuance models on each of the six life outcomes (Table 1). As our study aimed to extend on Stewart et al.'s (2022) and Seeboth and Mõttus's (2018) research, where they employed ENR in their analyses, we opted to adopt the same approach for the sake of comparative analysis. ENR is a variant of classic linear regression often used in studies predicting life outcomes from personality traits (Roberts et al., 2007). ENR differs from regular linear regression by utilizing penalty terms from Lasso (L1) and ridge (L2) regression.[1] These penalty terms guard against overfitting and handles correlated features and multicollinearity in the data.

The data set was randomly split into a training (67%) and a validation (33%) sub-sample, inspired by Stewart et al. (2022). This strategy enables us to train the model of one subset and evaluate its performance on an independent dataset. It is common practice to reserve a bigger portion of the sample for training the model, as more information produces a more accurate model. The training sample was used to set up ENR models for each of the six outcomes, with the Big Five factors, facets, or nuances as predictors. To further enhance model performance, we employed a 10-fold cross-validation and shrinkage parameters to obtain the optimal parameter lambda (λ), which minimizes cross-validation

error and prevents overfitting the data to the model.[2] Next, we evaluated the trained models by applying them to the validation subsample to predict each outcome. This allowed us to assess the model's predictive accuracy and its ability to generalize to unseen data. Finally, we quantified the model's performance by calculating the correlation between the predicted outcomes and the actual observed values. Each correlation was squared to show the percentage of explained variance. These steps were performed using by implementing a data analysis pipeline in R.

Results

A descriptive analysis was performed for the Big Five personality structure. Table 2 displays mean value, standard deviation, skewness, and kurtosis, along with Cronbach alpha values for facets and factors. The mean reliability for facets was α = .68, with 16 facets > .70. This can be compared with a meta-analysis which reported that the mean Cronbach's alpha (k = 4286) was α = .77, and that 25% of personality measurements report below α = .70 (Peterson, 1994). The factors and facets were overall acceptably symmetrical and normally distributed. Agreeableness and altruism showed notably high kurtosis. Furthermore, descriptive statistics on items are reported in Appendix C, together with a visual correlational heatmap for the 30 facets (Appendix D).

To test the structural validity of the Big Five, second order CFAs were conducted (S-CFA). Table 3 reports fit indices ($\chi2$(df), RMSEA, SRMR, and CFI), and standardized item loadings in rows for each respective facet. Neuroticism was the best-fitting model with an RMSEA of .05 and a CFI of .93. The remaining models had acceptable fits, with RMSEAs ranging from .06 to .07. CFIs ranged lower (from .81 to .93). All the models had SRMR values of .08 or less. The CFI value, although slightly below the conventional criterion of ≥ .90, is still close to it, suggesting a reasonable model fit. However, as stated earlier, achieving a perfect model fit in personality research is often unattainable due to the intricate nature of studying personality. The standardized factor loadings were on average of λ = .61 for extraversion, λ = .55 for neuroticism, λ = .55 for agreeableness, λ = .46 for conscientiousness and λ = .57 for openness. The full range was λ = .16–.95. Furthermore, see Appendix E for visual

---

[1] The Lasso regression adds a penalty term (L1) proportional to the absolute value of the coefficients, encouraging the model to use the most important features, and reducing the coefficients of less important features to zero. The Ridge regression adds a penalty term (L2) proportional to the square of the coefficients, encouraging the model to use all features, but reducing their magnitudes (shrinking them towards zero). Lasso regression produces modest models protected against overfitting but is sensitive to outliers and non-normality. Ridge regression handles correlated features and multicollinearity in the data but is prone to overfitting the model to the training data (Zou & Hastie, 2005).

[2] 10-fold cross-validation means that the training data is split into ten random sub-samples of equal size, called folds. Nine of those sub-samples are used for training and one for testing. This process is then repeated ten times, each time reserving a different fold for testing. The result of each run is averaged to produce an estimate of the model's performance.

Table 2 Descriptives for facets and factors

Factor and facet traits	α	M	SD	Skewness	Kurtosis
Extraversion	.88	14.94	1.86	−0.57	0.97
E1_Friendliness	.73	16.70	2.55	−0.84	0.46
E2_Gregariousness	.70	14.64	2.93	−0.42	−0.26
E3_Assertiveness	.75	14.61	2.50	−0.30	0.37
E4_Activity	.53	13.55	2.51	−0.20	0.04
E5_Excitement	.77	13.64	3.04	−0.07	−0.29
E6_Cheerfulness	.79	16.56	2.35	−0.93	1.91
Agreeableness	.83	18.49	1.75	−1.19	3.55
A1_Trust	.83	15.79	2.82	−0.82	1.02
A2_Morality	.68	17.81	2.41	−1.32	1.61
A3_Altruism	.61	17.24	2.36	−1.38	3.41
A4_Cooperation	.44	16.10	2.15	−0.71	0.86
A5_Modesty	.70	11.37	2.88	0.18	−0.21
A6_Sympathy	.76	16.76	2.85	−0.95	−0.98
Neuroticism	.87	7.74	1.81	0.78	0.96
N1_Anxiety	.80	8.00	2.82	0.85	1.04
N2_Anger	.73	7.59	2.70	0.83	1.05
N3_Depression	.80	6.29	2.37	1.45	2.50
N4_Self-conscious	.51	8.51	2.74	0.38	−0.36
N5_Immoderation	.51	9.00	2.46	0.32	0.08
N6_Vulnerability	.61	7.05	2.11	0.76	0.59
Openness	.81	14.32	1.77	−0.03	−0.27
O1_Imagination	.81	12.51	3.51	0.04	−0.51
O2_Artistic	.67	15.13	3.16	−0.39	−0.44
O3_Emotionality	.53	16.22	2.24	−0.33	−0.32
O4_Adventurous	.71	14.03	2.82	−0.07	−0.32
O5_Intellect	.67	15.61	2.95	−0.29	−0.54
O6_Liberalism	.48	12.42	2.61	−0.22	0.27
Conscientiousness	.86	17.58	1.63	−0.68	0.58
C1_Self-efficacy	.81	17.56	2.16	−1.18	3.48
C2_Orderliness	.81	16.23	3.16	−0.86	0.49
C3_Dutifulness	.69	18.70	1.76	−1.81	3.62
C4_Achievement	.61	16.98	2.16	−0.51	−0.10
C5_Self-discipline	.73	16.53	2.46	−0.76	−0.60
C6_Cautiousness	.60	14.26	2.46	−0.07	0.16

Overview of the facet and factor scores (N = 440). α = Cronbach alpha; M = Mean; SD = Standard deviation

Table 3 S-CFA fit indices and standardized item loadings

| Factors and facets	Item a	Item b	Item c	Item d
Extraversion $\chi^2(246) = 798.82$; RMSEA = .07; SRMR = .07, CFI = .85				
Friendliness	.57	.47	.56	.59
Gregariousness	.77	.56	.47	.66
Assertiveness	.63	.67	.55	.34
Activity level	.26	.70	.58	.36
Excitement-seeking	.74	.77	.50	.64
Cheerfulness	.54	.51	.51	.54
Neuroticism $\chi^2(246) = 494.22$; RMSEA = .05; SRMR = .06, CFI = .93				
Anxiety	.73	.66	.63	.52
Anger	.58	.69	.62	.43
Depression	.69	.38	.66	.41
Self-consciousness	.64	.65	.45	.29
Immoderation	.51	.22	.51	.56
Vulnerability	.41	.35	.46	.41
Agreeableness $\chi^2(246) = 762.11$; RMSEA = .07; SRMR = .08, CFI = .84				
Trust	.74	.70	.68	.47
Morality	.75	.55	.47	.23
Altruism	.50	.54	.40	.42
Corporation	.17	.44	.49	.36
Modesty	.27	.95	.82	.45
Sympathy	.74	.71	.48	.59
Conscientiousness $\chi^2(246) = 762.74$; RMSEA = .07; SRMR = .08, CFI = .85				
Self-efficacy	.46	.46	.52	.50
Orderliness	.35	.80	.83	.90
Dutifulness	.42	.32	.34	.38
Achievement striving	.47	.38	.39	.46
Self-discipline	.42	.43	.63	.63
Cautiousness	.53	.16	.28	.71
Openness $\chi^2(246) = 691.85$; RMSEA = .06; SRMR = .07, CFI = .81				
Imagination	.81	.71	.80	.82
Artistic interests	.80	.59	.61	.60
Emotionality	.42	.39	.42	.42
Adventurousness	.58	.70	.52	.56
Intellect	.69	.74	.53	.43
Liberalism	.58	.36	.67	.24

χ^2 = chi-square value; *RMSEA* Root Mean Square Error of Approximation, *SRMR* Standardized Root Mean Residuals, *CFI* Comparative Fit Index, *α* Cronbach alpha. Item a, Item b, Item c, and Item d denote the four items making up each facet trait

CFA figures for each Big Five structure, showing the facet loadings. Only two facets, cooperation and modesty, failed to show appreciable loading estimates. Overall, the structural validity of the Big Five hierarchies reported acceptable psychometric properties.

Factors, facets and nuances in life outcomes

Before running the models comparing the predictive validities between the levels in the personality structure, a bivariate correlation analysis between the Big Five factors and outcome measures were performed. Table 4 displays correlations between the Big Five factors and the six life outcomes in the present study. Extraversion was positively associated with all six life outcomes, whereas associations with neuroticism were mostly negative (with empathy as the exception). Conscientiousness was mostly positively related, while openness was found to be positively related

Table 4 Correlations between big five factors and outcome measures

	1	2	3	4	5	6	7	8	9	10	11
1 Agreeableness											
2 Extraversion	−.01										
3 Neuroticism	−.13	**−.57**									
4 Conscientiousness	−.09	**.46**	**−.61**								
5 Openness	.00	.28	−.20	−.04							
6 Job satisfaction	.04	.21	−.16	.06	.10						
7 Social satisfaction	−.03	**.45**	**−.44**	**.37**	.01	.21					
8 Empathy	−.09	.10	.04	.04	.06	.14	.17				
9 Bright future	.06	**.33**	**−.31**	.20	.16	.08	.15	−.03			
10 Intrinsic reward	−.04	**.31**	−.29	.25	−.21	.09	.04	.05	**.34**		
11 Extrinsic reward	−.01	.12	−.06	.12	−.13	−.06	.19	−.03	.05	.21	

($N=440$). Significances are as follows: $r \geq .09$, $p < .05$, $r \geq .13$, $p < .01$, and $r \geq .16$, $p < .001$. Boldface indicates strong effect sizes as recommended by Gignac and Szodorai (2016)
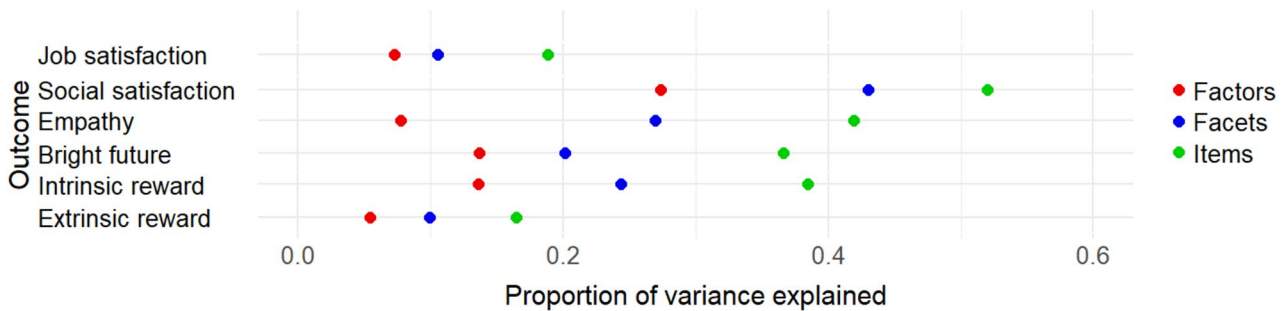
Fig. 1 Variance accounted for in each life outcome. *Note.* Results of the Elastic Net Regression models, with six life outcomes at three levels of personality prediction – factors (red dots), facets (blue dots), and nuances (green dots). The scores reflect the average score after 100 permutations

to positive beliefs about the future and negatively linked to both extrinsic and intrinsic reward. No relationship was observed between agreeableness and the other outcomes.

The main aim of the study was to investigate the predictive validity of factors, facets, and nuances for various life outcomes. The hypotheses were that facets would outperform factors, and that nuances would outperform both facets and factors. Figure 1 demonstrates the results of the analysis, with variance explained by each of the models for all six life outcomes.[3] As illustrated by Fig. 1, clear support was found for both hypotheses. The factors (red dots) were all far to the left of the facets (blue dots) and items (green dots) in the diagram, indicates lesser explained variances. While the life outcomes varied in the degree to which the personality models predicted them, facets consistently explained more variance than factors for all outcomes, and nuances explained

more variance than facets. On average, factors, facets, and nuances accounted for 12%, 22.5% and 34% respectively of explained variance across all outcomes (see Table 5).

As further shown in Table 5, social satisfaction was the strongest personality-outcome association, with nuances explaining 52% of variance. Extrinsic reward was the weakest outcome-association, with nuances explaining 16% of

Table 5 Proportion of variance explained for each outcome measure

Outcome	Factor	Facet	Nuance
Job satisfaction	.07	.11	.19
Social satisfaction	.27	.43	.52
Empathy	.07	.27	.42
Bright future	.14	.20	.37
Intrinsic reward	.14	.24	.38
Extrinsic reward	.05	.10	.16

Explained variance for six life outcomes based on the Elastic Net Regression analysis ($N=440$). The percentage of explained variance were calculated by squaring each correlation

[3] A multiple linear regression (MLR) was performed as well to allow for easier interpretation of results across studies and to facilitate inter-methods comparisons. See Appendix F for the results of the MLR.

variance. The amount of variance explained also varied across outcomes. For instance, nuances explained 12% additional variance than factors for job satisfaction, while nuances explained an additional 35% compared to factors for empathy. Overall, facets tended to explain 1.5 to 2 times more than factors, and nuances tended to explain 2 to 3 times more in these life outcomes, confirming both our hypotheses.

Discussion

The present study investigated the predictive validity of the hierarchical structure of personality traits comparing factor-, facet-, and item-level models. Results supported both hypotheses: First, facets provided better predictive validity than factors. Secondly, nuances provided better predictive validity than facets and factors. These results were largely consistent with recent research (Kajonius & Johnson, 2019; Speer et al., 2022; Stewart et al., 2022).

Particularly, inspired by and contrasted to Stewart et al. (2022), we applied the long personality instrument IPIP-NEO-120 in Swedish. Interestingly, the present study reported even greater differences between factors-, facets-, and item-models in predicting outcomes. On average, nuances explained 13% more variance than facets, while Stewart et al. (2022) reported a difference of 3%. Such differences could depend on the number of items in the IPIP-NEO (120 items) compared to the BFI-2 (60 items). Nevertheless, reliability (Cronbach's alphas) was similar between studies and yet differences were larger than expected. It could be that our present study made use of only single-item outcomes which could have decreased some error variance, making item-level models artificially superior. Furthermore, the self-reported outcomes were quite like some personality traits (e.g., satisfaction, optimism, or motivation), and occasional items might have fitted the outcome prediction, thus making item-level unrepresentatively better than factor-level.

Value of nuances in personality

Item-level nuances accounted for the most variance in the social satisfaction outcome (see Table 1 for outcome descriptions), explaining 52%, which is unusually strong in personality research (equivalent to a multiple $r = .72$). This was a 25% additional variance than was explained by the factor-level model. Even in the least predictive outcome, extrinsic reward, accounting for 16% variance, nuances outperformed factors by a factor of 3. It seems that the aggregation of items into broader facets and then factors loses nuanced information in how personality traits relate to personal outlooks on life. The largest discrepancy in predictive validity between models was observed for the empathy outcome, with nuances explaining a 35% additional variance compared to

factors. This could exemplify the occasional biased advantage of item-level use, seeing that IPIP-NEO-120 contains the facet sympathy, which includes items such as "I feel sympathy with homeless" and "I feel sympathy with those with problems". Moreover, it contains the facet emotionality, including items such as "I feel others' emotions", which is akin to empathy. Considering the overlap between such items and the empathy outcome, a post-hoc correlational analysis was conducted, between the IPIP-NEO items and the life outcome variables (See Appendix G). However, no significantly strong correlation was found between the empathy life outcome and items akin to empathy. As such, it can be argued that the use of items more accurately reflects nuances of personality than the use of the facet sympathy or the factor agreeableness.

Implications

Generalized predictions in personality research are often derived from broad, parsimonious Big Five factors, practical and easy to communicate. Often item selection is standardized and items making up facets are viewed as interchangeable (see the large item-pool in IPIP) despite items capturing distinct and unique variance on their own (Mõttus et al., 2014). While broad factors are associated with both positive and negative life outcomes, narrow facet or nuance models may be more appropriate if precision and prediction are the desired goals.

Personality-outcome connections frequently serve as the foundation for interventions. As such, we must weigh the benefits of general face validity against the loss of precision and prediction. For instance, researchers may be interested in increasing the well-being of the general population. Considering the consistent negative correlation between neuroticism and well-being, as well as the positive association between conscientiousness and well-being (Gilberto et al., 2020) interventions may focus on strategies aimed at either reducing neuroticism or enhancing conscientiousness. However, this is a very complex task. If, instead, it turns out that only a select few specific traits, discernible through facets or nuanced measures, underlie these associations, it suggests that interventions can be more targeted and actionable. In such a scenario, should the Big Five factors continue as the prevailing level of analysis for personality-outcome research when a potentially informative item-level alternative is available? The study's results indicate that achieving greater success in personality-based interventions may be attainable by targeting precise behavioral, cognitive, affective, and motivational tendencies, captured by narrower traits. Furthermore, researchers should aim not only to establish connections between traits and outcomes but also to offer

explanations that pave the way for actionable insights. The findings of this study suggest that employing nuance models may facilitate better insight into personality outcome relationships.

Limitations and conclusion

While the study has several strengths such as a novel testing of a long personality instrument in a sample and language other than English, it also has limitations. The sample size can be considered meager for the current purpose (Cui & Gong, 2018). A larger sample is not only more representative, increasing the generalizability, but is also more robust to random fluctuations. Correspondingly, very little is known about the present sample having been collected through an internet webpage, making it difficult to gauge the generalizability of the sample. Very likely, there is an overrepresentation of high openness and agreeableness (cf. Kajonius & Johnson, 2019).

Furthermore, the life outcomes assessed in the present study were represented by singular items framed in a positive manner, potentially leading to increased social desirability tendencies among the participants. Research has demonstrated that items perceived as highly desirable have been shown to artificially produce a general factor of personality (Bäckström et al., 2009). This could imply that some of items (e.g., "to what extent is reward your source of motivation for working?") could have overestimated predictive validities. Future studies should take into account whether their life outcome measures are framed in a neutral manner.

Another limitation worth mentioning is the low CFA fit indices that suggest a lack of internal structure. Notably, personality structure are renowned for not showing great model fits. Presumably this could be due to the error-proneness both in response styles as well as individuals differing slightly in internal factor-structures. This is considered a problem within the field of personality research, and some degree of model misfit is to be expected when working with personality assessments (Hopwood & Donnellan, 2010).

Also, both a strength and a limitation is the use of an instrument with a large item-pool, such as IPIP-NEO. This instrument is intended to maximize common variance and scope in facets and factors. Item and residual correlations make it almost impossible to gauge how many nuances are captured by the instrument. It is most likely less than 120, as the items overlap in content, and some are close to reverse-keyed duplicates.

The present study was able to conclude that personality measured at item-level clearly outperformed both facets and factors. Additionally, the predictive advantage was partly contingent on the specific outcome. Future research should encompass a wider range of life outcomes to explore which specific results are most accurately predicted by narrow traits. The Big Five factor models have traditionally produced countless trait-outcome associations and are liked for their practicality and theoretical simplicity. However, this may be the time to put forward the recommendation of making more use of narrower traits, and nuances, especially when enhanced statistical prediction is desired.

Declarations

Ethics Since the collected data was anonymous and voluntary, no ethical review was required.

Conflict of interest The authors report no conflicts of interest.

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*(3), 335–344. https://doi.org/10.1016/j.jrp.2008.12.013

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Briley, D. A., & Tucker-Drob, E. M. (2012). Broad bandwidth or high fidelity? Evidence from the structure of genetic and environmental effects on the facets of the five factor model. *Behavior Genetics, 42*(5), 743–763. https://doi.org/10.1007/s10519-012-9548-8

Briley, D. A., & Tucker-Drob, E. M. (2014). Genetic and environmental continuity in personality development: A meta-analysis. *Psychological Bulletin, 140*(5), 1303–1331. https://doi.org/10.1037/a0037091

Brown, T. A. (2014). *Confirmatory factor analysis for applied research*. Guilford Press.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Condon, D. M., Wood, D., Mõttus, R., Booth, T., Costantini, G., Greiff, S., Johnson, W., Lukaszewski, A., Murray, A. L., Revelle, W., Wright, A. G. C., Ziegler, M., & Zimmermann, J. (2020). Bottom up construction of a personality taxonomy. *European Journal of Psychological Assessment, 36*(6), 923–934. https://doi.org/10.1027/1015-5759/a000626

Cui, Z., & Gong, G. (2018). The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *NeuroImage, 178*, 622–637. https://doi.org/10.1016/j.neuroimage.2018.06.001

Epskamp, S. (2022). semPlot: Path diagrams and visual analysis of various SEM packages' output. R package version 1.1.6. https://CRAN.R-project.org/package=semPlot

Espinoza, J. A., O'Neill, T. A., & Donia, M. B. (2023). Big five factor and facet personality determinants of conflict management styles. *Personality and Individual Differences*, 203, 112029.https://doi.org/10.1016/j.paid.2022.112029

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Glmnet: Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22. https://www.jstatsoft.org/v33/i01/

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Gilberto, J. M., Davenport, M. K., & Beier, M. E. (2020). Personality, health, wealth, and subjective well-being: Testing a integrative model with retired and working older adults. *Journal of Research in Personality*, 87, 103959.https://doi.org/10.1016/j.jrp.2020.103959

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

Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review, 14*(3), 332–346. https://doi.org/10.1177/1088868310361240

Johnson, J. A. (2014). Measuring thirty facets of the five factor model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality, 51*, 78–89. https://doi.org/10.1016/j.jrp.2014.05.003

Kajonius, P. J., & Carlander, A. (2017). Who gets ahead in life? Personality traits and childhood background in economic success. *Journal of Economic Psychology, 59*, 164–170.

Kajonius, P. J., & Johnson, J. A. (2019). Assessing the structure of the five factor model of personality (IPIP-NEO-120) in the public domain. *Europe's Journal of Psychology, 15*(2), 260–275. https://doi.org/10.5964/ejop.v15i2.1671

Kuhn, M. (2022). Caret: Classification and regression training. R package version 6.0–93. https://CRAN.R-project.org/package=caret

McCrae, R. R. (2015). A more nuanced view of reliability. *Personality and Social Psychology Review, 19*(2), 97–112. https://doi.org/10.1177/1088868314541857

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

Mõttus, R., Mccrae, R., Allik, J., & Realo, A. (2014). Cross-rater agreement on common and specific variance of personality scales and items. *Journal of Research in Personality, 52*, 47–54. https://doi.org/10.1016/j.jrp.2014.07.005

Mõttus, R., Kandler, C., Bleidorn, W., Riemann, R., & McCrae, R. R. (2017). Personality traits below facets: The consensual validity, longitudinal stability, heritability, and utility of personality nuances. *Journal of Personality and Social Psychology, 112*(3), 474–490. https://doi.org/10.1037/pspp0000100

Olaru, G., van Scheppingen, M. A., Bleidorn, W., & Denissen, J. J. (2023). The link between personality, global, and domain-specific satisfaction across the adult lifespan. *Journal of Personality and Social Psychology, 125*(3), 590–606. https://doi.org/10.1037/pspp0000461

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of Consumer Research, 21*(2), 381–391.

Polderman, T. J., Benyamin, B., De Leeuw, C. A., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics, 47*(7), 702–709.

R Core Team (2021). R: A language and environment for statistical computing. R foundation for statistical computing. https://www.r-project.org/

Raoust, G., Kajonius, P., & Hansson, S. (2023). Personality traits and decision-making styles among obstetricians and gynecologists managing childbirth emergencies. *Scientific Reports, 13*(1), 5607.

Revelle, W. (2021) psych: Procedures for psychological, psychometric, and personality research, R package version 1.7.8. https://CRAN.R-project.org/package=psych

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*(4), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Schwaba, T., Robins, R. W., Grijalva, E., & Bleidorn, W. (2019). Does openness to experience matter in love and work? Domain, facet, and developmental evidence from a 24-year longitudinal study. *Journal of Personality, 87*(5), 1074–1092. https://doi.org/10.1111/jopy.12458

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

Sleep, C. E., Lynam, D. R., & Miller, J. D. (2021). A comparison of the validity of very brief measures of the big five/five-factor model of personality. *Assessment, 28*(3), 739–758. https://doi.org/10.1177/1073191120939160

Soto, C. J. (2019). How replicable are links between personality traits and consequential life outcomes? The life outcomes of personality replication project. *Psychological Science, 30*(5), 711–727. https://doi.org/10.1177/0956797619831612

Soto, C. J., & John, O. (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, 117–143.https://doi.org/10.1037/pspp0000096

Speer, A. B., Christiansen, N. D., Robie, C., & Jacobs, R. R. (2022). Measurement specificity with modern methods: Using dimensions, facets, and items from personality assessments to predict performance. *Journal of Applied Psychology, 107*(8), 1428–1439. https://doi.org/10.1037/apl0000618

Stewart, R. D., Mõttus, R., Soto, C. J., Seeboth, A., & Johnson, W. (2022). The finer details? The predictability of life outcomes

from big five domains, facets, and nuances. *Journal of Personality, 19*(2), 167–182. https://doi.org/10.1111/jopy.12660

Tuszynski, J. (2021). caTools: Moving window statistics. R package version 1.18.2. https://CRAN.R-project.org/package=caTools

Vainik, U., Dagher, A., Realo, A., Colodro-Conde, L., Mortensen, E. L., Jang, K., Juko, A., Kandler, C., Sørensen, T. I. A., & Mõttus, R. (2019). Personality-obesity associations are driven by narrow traits: A meta-analysis. *Obesity Reviews, 20*(8), 1121–1131. https://doi.org/10.1111/obr.12856

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Wickham, H (2007). reshape2: Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12), 1–20. https://www.jstatsoft.org/v21/i12/

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., & Yutani, H. (2019). Tidyverse: Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023a). dplyr: A grammar of data manipulation. R package version 1.1.0. https://CRAN.R-project.org/package=dplyr

Wickham, H., Vaughan, D., & Girlich, M. (2023b). tidyr: Tidy messy data. R package version 1.3.0. https://CRAN.R-project.org/package=tidyr

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67*, 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.