

# Personality Itemmetrics: What Differentiates Among the Elite Items of the NEO-FFI?

Ronald R. Holden<sup>1</sup>, John A. Johnson<sup>2</sup>, & Greg Murray<sup>3</sup>

<sup>1</sup> Queen's University, Canada, <sup>2</sup> Pennsylvania State University, U.S.A., <sup>3</sup> Swinburne University of Technology, Australia



## ABSTRACT

Using Canadian ( $n = 240$ ) and Australian ( $n = 530$ ) samples, this research evaluated NEO Five-Factor (NEO-FFI) item quality in terms of: (a) item-total correlations; (b) item-criterion (peer-report) correlations; (c) item principal component loadings; (d) item 6-month test-retest correlations; and (e) a composite quality index. Across the 60 items, item-total correlations ( $M = .47$ ), item-criterion correlations ( $M = .31$ ), item component loadings ( $M = .54$ ), and retest correlations ( $M = .59$ ) confirmed the superior nature of the NEO-FFI. For example, the Openness item "Sometimes when I am reading poetry or looking at a work of art, I feel a chill or wave of excitement" had the greatest composite quality index with values of .64, .44, .70, and .68 for item-total correlation, item-criterion correlation, item component loading, and retest correlation, respectively. Across items, quality was then related to other properties. Quality for this restricted item range was associated with: (1) item content higher in rated affect; and (2) homogeneity of respondents' perception of item dimensionality – that is, higher quality items had greater respondent agreement concerning the dimension that the item was related to.

## INTRODUCTION

Units of analysis consisted of the 60 NEO-FFI items (Costa & McCrae, 1991), 12 from each of the Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness scales. Item characteristics, obtained from a variety of sources, are outlined below.

## METHOD

### Item Quality Measures

**Item-Total and Item-Criterion Correlations.** These were derived from Holden and Starzyk (2005). Based on the responses of 420 Canadian university students (210 roommate pairs), item-total correlations were calculated for self-report. Item-criterion correlations were calculated between self-report item responses and roommate-reported NEO-FFI scale scores (observer form).

**Item Component Loadings and Retest Correlations.** From Murray et al. (2003) for 527 respondents randomly selected from a Melbourne, Australia electoral roll, initial testing varimax-rotated component loadings for a relevant component were used. For this sample, 6-month item retest correlations were also calculated.

**Composite Quality Index.** For each NEO-FFI item, scores on the first principal component associated with item-total correlations, item-criterion correlations, item component loadings, and 6-month retest correlations were calculated.

### Item Predictor Measures

**Item Length.** Number of words in item.

**Negations.** Number of negations in item.

**Content Analysis Ratings.** Expert ratings (Pytlik Zillig et al., 2002) of NEO-FFI items regarding the presence of affective, behavioral, or cognitive processes.

**Response Strategy Ratings.** Ratings by 110 introductory psychology students on the degree that responding to an item involved: a specific behavior/experience; multiple behaviors/experiences; a general characteristic/trait; multiple characteristics/traits; a comparison to a specific person; comparisons to several people; referring to what another person had said about the respondent; referring to what several people had said about the respondent (Gordon & Holden, 1996).

**Item Homogeneity.** Holden et al. (2006) had 114 introductory psychology students undertake a free sort of NEO-FFI items. Based on the NEO-FFI item co-occurrence matrix, items were scaled into multidimensional space. For the present study, a measure of item homogeneity was generated by calculating the sum of the squared multidimensional space coordinates.

## TABLES

Table 1. Descriptive Statistics for Item Quality Measures ( $N = 60$  items)

Item Quality Measure	Mean	SD	Measure Intercorrelations			
			ITC	ICC	ICL	RC
Item-Total Correlation (ITC)	.47	.14				
Item-Criterion Correlation (ICC)	.31	.07	.71			
Item Component Loading (ICL)	.54	.12	.87	.69		
Retest Correlation (RC)	.59	.08	.10	.43	.17	
Composite Quality Index (CQI)	.00	1.00	.91	.89	.91	.39

Table 2. NEO-FFI Items Strongest and Weakest in Composite Quality

### Strongest

- Sometimes when I am reading poetry or looking at a work of art, I feel a chill or wave of excitement. ITC = .64; ICC = .44; ICL = .70  
RC = .68; CQI = 1.65
- I am a cheerful high-spirited person. ITC = .62; ICC = .43; ICL = .71  
RC = .67; CQI = 1.56
- I'm pretty good about pacing myself so as to get things done on time. ITC = .64; ICC = .42; ICL = .67  
RC = .65; CQI = 1.40

### Weakest

- I generally try to be thoughtful and considerate. ITC = .38; ICC = .15; ICL = .46  
RC = .44; CQI = -1.51
- Most people I know like me. ITC = .20; ICC = .14; ICL = .25  
RC = .61; CQI = -2.32
- Once I find the right way to do something, I stick to it. ITC = .04; ICC = .08; ICL = .19  
RC = .54; CQI = -3.30

## REFERENCES

- Costa, P. T., Jr., & McCrae, R. R. (1991). *NEO Five-Factor Inventory (Form S)*. Odessa, FL: PAR.
- Gordon, E. D., & Holden, R. R. (1996). Use of item ratings to examine personality test item cognitive response processes. *Personality and Individual Differences*, 21, 897-905.
- Holden, R. R., & Starzyk, K. B. (2005, August). *Social desirability moderates the validity of the NEO Five-Factor Inventory*. Paper presented at the American Psychological Association Annual Convention, Washington.
- Holden, R. R., Wasylykiw, L., Starzyk, K. B., Book, A. S., & Edwards, M. J. (2006). Inferential structure of the NEO Five-Factor Inventory: Construct validity of the Big Four personality clusters. *Canadian Journal of Behavioural Science*, 38, 24-40.
- Murray, G., Rawlings, D., Allen, N. B., & Trinder, J. (2003). NEO Five-Factor Inventory scores: Psychometric properties in a community sample. *Measurement and Evaluation in Counseling and Development*, 36, 140-149.
- Pytlik Zillig, L. M., Hemenover, S. H., & Dienstbier, R. A. (2002). What do we assess when we assess a Big 5 trait? A content analysis of the affective, behavioral, and cognitive processes represented in Big 5 personality inventories. *Personality and Social Psychology Bulletin*, 28, 847-858.

Table 3. Correlations between Item Quality and Predictor Measures ( $N = 60$  items)

	Item Quality Measure				
	ITC	ICC	ICL	RC	CQI
Item Length	.07	.02	.07	.06	.07
Negations	.04	.04	-.02	-.46**	-.05
<b>EXPERT RATINGS</b>					
Affective Rating	.35**	.11	.33*	-.09	.26*
Behavioral Rating	-.21	-.06	-.20	.15	-.14
Cognitive Rating	-.19	-.08	-.18	-.09	-.17
<b>RESPONSE STRATEGY</b>					
Specific Behavior	.08	-.04	.04	-.02	.03
Multiple Behaviors	.09	.02	.05	-.00	.05
General Trait	.07	.14	.11	.00	.11
Multiple Traits	.14	.22	.13	.01	.17
Compared to 1 Other	.18	.15	.17	-.01	.17
Compared to > 1 Other	-.13	-.05	-.09	.02	-.09
What 1 Other Said	.09	.08	.15	-.02	.11
What > 1 Other Said	-.12	-.15	-.15	.01	-.15
Item Homogeneity	.14	.34**	.17	.36**	.28*

\* $p < .05$ ; \*\*  $p < .01$

Note: Correlations of .10, .30, and .50 correspond to small, medium, and large effect sizes, respectively.

## SUMMARY

1. NEO-FFI items are a highly refined, elite collection possessing a highly restricted range of strong quality.
2. Within this highly restricted range of item quality, higher item quality is associated with:
  - a. Avoidance of negations in item wording.
  - b. Item content that is affective in nature.
  - c. Item content that is perceived homogeneously by respondents.

## Correspondence

Ronald R. Holden, Department of Psychology, Queen's University, Kingston, ON, K7L 3N6, Canada. Telephone: 613-533-2879, Email: holdenr@post.queensu.ca