



Improving Prediction Through Personality and Criterion ABC Alignment

Michael Tapia, Hogan Assessment Systems

Matthew Mol, The University of Tulsa

Kimberly Nei, Hogan Assessment Systems

This paper was accepted as a SIOP poster for the 2017 conference.

Abstract

We hypothesize that by aligning the affective, behavioral, and cognitive content of personality and workplace criteria, prediction will be improved. This hypothesis is tested in two datasets that have both personality and performance data. The results generally support the hypothesis; there was better prediction of performance on average.

Introduction

Personality is defined as a relatively enduring pattern of affect, behavior, and cognition (Roberts, Walton, & Viechtbauer, 2006). These three sub-traits are collectively referred to as the “ABC” components of personality. Personality test items can be reliably categorized into these ABC components. All of these are present in varying frequencies for each of the Big Five factors, with slight variation across the scales used (Pytlik Zillig, Hemenover, & Dienstbier, 2002; Werner & Pervin, 1986; Wilt & Revelle, 2015).

Personality Components

Affect. Affective items are described as “internal, motivational, and evaluative, valenced states, including patterns of feeling, emotions, ‘feeling-like’ states, and preferences.” (Pytlik Zillig et al., 2002, p. 850). This definition encapsulates what are commonly thought of as drivers and feelings. This means items such as, “*I enjoy parties*” and “*I get nervous at large gatherings*” would both fall under the category of affective personality items, even though the former is measuring a driver, while the latter is measuring a feeling. What is common amongst all of the constructs subsumed in the category of affect is that they refer to how we feel (Wilt & Revelle, 2015).

Behavior. Items that encompass behavioral component of personality are defined as the “overt and directly observable actions [taken by individuals], including both active (e.g., bike-riding) and passive (e.g., watching television) behaviors” (Pytlik Zillig et al., 2002, p. 850). An example of a behavioral personality item would be, “*I go to parties frequently.*” A commonality amongst many behavioral items is they tend to refer to the frequency with which a behavior is engaged in. Behavior is the most proximal of the three components to real world behavior, while affect and cognition serve more to drive and guide eventual behavior (Hogan, Hogan, & Roberts, 1996).

Cognition. Items within the cognition component of personality are defined as “thoughts, beliefs, patterns, or modes of thinking” (Pytlik Zillig et al., 2002, p. 850). An example of a cognitive personality item is, “*Parties are useful tools for meeting new people.*” Cognitive items tend to start with phrases like “I believe,” or “I think,” although the example item shows that this is not always true.

Categorization of Personality Items

Previous research (Pytlik Zillig et al., 2002; Werner & Pervin, 1986; Wilt & Revelle, 2015) has categorized a variety of personality scales into their components. Studies have previously sorted items based on content categories such as: behavior and cognition (Pytlik Zillig et al., 2002); affect, behavior, cognition, and desires (Wilt & Revelle, 2015); and

cognitive-self, cognitive-beliefs, values and opinions, affective preferences, affective feelings, and behaviors (Werner & Pervin, 1986). What all of these categorization systems have in common are the ABCs. Scales that have been subjected to this categorization range from the MMPI (Hathaway & McKinley, 1951) to the PRF (Jackson, 1974), but the Hogan Personality Inventory (HPI; Hogan, Hogan, & Warrenfeltz, 2007) has yet to be subjected to the process. This study aimed to examine the relative distribution of ABC components for each of the HPI scales.

Research Question: What is the ABC distribution for the Hogan Personality Inventory?

Personality and Prediction

Beyond the ABC portion of the definition, another key component of the definition of personality is the relative temporal stability. This temporal stability allows for the prediction of future behavior and mental states from current scores on personality constructs. Personality has been shown to predict a variety of outcomes useful in the workplace, including job performance (Tett, Jackson, & Rothstein, 1991) and job attitudes (Judge, Heller, & Mount, 2002). As a result of this predictive ability, personality testing for the selection of new employees had already grown into a \$400 million industry in the United States by 2004 (Hsu, 2004) and has only continued to grow.

Beyond temporal stability, two key concepts allow for prediction: construct overlap and bandwidth fidelity. Construct overlap refers to the degree to which two constructs are similar. From this perspective it is not surprising that personality is predictive of a variety of outcomes. Facets of personality overlap with performance behaviors across a variety of jobs. For instance, Conscientiousness, characterized by detail and achievement orientation, is essential for detail-oriented positions, such as accountants. It stands to reason that somebody who is highly detail-oriented, by default, would excel in a position that requires attention to detail. Bandwidth fidelity refers to the degree to which the breadth or narrowness of the predictor and the criterion are similar (Ones & Viswesveran, 1996). The greater the similarity of breadth, the higher the correlation between the two constructs.

This is the logic behind the contextualization of personality scales (Roberts, 2007). By narrowing the assessment of personality to just personality “at work,” rather than the broad construct of personality across all situations, the prediction of job performance is improved. The improvements in prediction are not only due to increased construct overlap, via eliminating parts of the predictor that are not relevant at work, but also through improved bandwidth fidelity. By eliminating areas of the predictor that are not similar to the criteria, prediction could presumably be improved.

Integrating Components of Personality with Principles of Prediction

Personality items are not the only thing that can be categorized into the ABC components. Many, if not all, criteria can be categorized as affective, behavioral, cognitive, or some combination of the three. For example, performance is the expected organizational value of a series of behaviors and behavioral episodes (Motowidlo, 2003). In contrast, job attitudes are affective and cognitive evaluations of the workplace (Brummel & Bowling, 2013). When taking into account the principles of construct similarity and bandwidth fidelity, it stands to

reason that scales composed of behavioral items, for example, would predict behavioral outcomes better than scales that are composed of a combination of personality items from each of the ABC components.

Rauthmann and Denissen (2011) explored a similar concept. They examined whether an extraversion scale on a behavioral frequency scale, or an extraversion scale on a valence scale would do a better job of predicting behaviors. They found that across a variety of behavioral outcomes, the behavioral frequency scale out-predicted the valence scale. Another interesting piece of empirical support comes through the combination of the results from Tett et al. (1991), and Pytlik Zillig et al. (2002). Tett et al.'s meta-analysis shows that for many jobs, conscientiousness is an important predictor of performance. Combine that finding with Pytlik Zillig et al.'s (2002) finding that conscientiousness has the highest proportion of behavioral items, and one might conclude that perhaps one of the reasons conscientiousness is a consistent predictor of performance is because these scales typically have a large proportion of behavioral items.

The bandwidth fidelity and construct similarity rationale combined with the empirical evidence supports the hypothesis that using pure subscales composed of only affective, behavioral, or cognitive items of current personality scales may yield better prediction of a variety of matched outcomes. Specifically, when personality is used to predict performance in work scenarios, we believe that personality constructs broken down into ABC subscales will derive different levels of prediction of performance.

Hypothesis 1: Behavioral subscales of the HPI will outperform affective, cognitive, and composite scales of personality in the prediction of performance.

Method

Measures

Hogan Personality Inventory (HPI). The HPI is a measure of day-to-day personality. It consists of seven scales that are conceptually mapped to the Five Factor Model (FFM). HPI Adjustment aligns with FFM Emotional Stability and measures the degree to which a person appears calm and self-accepting. HPI Ambition aligns with FFM Extraversion, in part, and measures the degree to which a person seems socially self-confident, leaderlike, competitive, and energetic. HPI Sociability aligns with FFM Extraversion, in part, and measures the degree to which a person seems to need and/or enjoy interactions with others. HPI Interpersonal Sensitivity aligns with FFM Agreeableness and measures the degree to which a person seems perceptive, tactful, and socially sensitive. HPI Prudence aligns with FFM Conscientiousness and measures the degree to which a person seems conscientious, conforming, and dependable. HPI Inquisitive aligns with FFM Openness to Experience, in part, and measures the degree to which a person is perceived as bright, creative, and interested in intellectual matters. HPI Learning Approach measures the degree to which a person seems to enjoy academic activities and value educational achievement (Hogan, Hogan, & Warrenfeltz, 2007). The assessment consists of 206 items and takes approximately 15-20 minutes to complete.

Performance. To collect performance data, participants were rated by supervisors using a Performance Rating Form (PRF) to solicit overall performance ratings and ratings on competencies identified as critical by subject matter experts during a job analysis. The PRF's differed by study to account for the differences in job requirements.

Participants

Data. The datasets used for this study were provided by Hogan Assessment Systems. Data were obtained through two collaborations with organizations that sought to improve their selection procedures for one organization's salesforce and for the other's leadership positions. In the sales dataset, a total of 223 sales personnel (159 males and 64 females) completed the HPI and were rated on their performance. Ages of sales personnel ranged from 26 to 70 with an average age of 41. In the leadership dataset, a total of 319 potential leaders (172 males, 146 females, and 1 unknown) completed the HPI and were rated on their performance. Ages of potential leaders ranged from 25 to 66 with an average age of 42.

Content Coding. To create the ABC component subscales, five coders (I/O graduate students and professionals) categorized items from the HPI.

Procedure & Analyses

HPI data and criterion performance ratings were obtained in late 2014 through a criterion study looking at the predictive ability of a personality profile using scales from the HPI and other core Hogan assessments. Coders categorized HPI items during the Summer of 2016 and were instructed to categorize each HPI item as affective, behavioral, or cognitive, based on definitions from Pytlik et al. (2002). Rater agreement was measured using Fleiss' Kappa based on categorical variables and multiple raters.

After items were categorized into affective, behavioral, or cognitive components, ABC component subscales for each of the HPI scales were created for a total of 21 subscales. HPI total scale and ABC subscale scores were correlated with performance variables from two of Hogan's archived criterion studies to look at the generalizability of the findings across performance domains. To examine the strength of the correlations, absolute values of the correlations were also examined. Supervisor ratings of job performance lack reliability which affects the relationship between predictors and measures of job performance. To account for this, a correction for unreliability of the criterion was applied using a .52 reliability coefficient proposed by Viswesvaran, Ones, and Schmidt (1996). We chose not to correct for correlation coefficients to estimate validity at the construct level, following protocol outlined in Gaddis and Foster (2010).

Results

Rater Agreement and ABC Subscale Reliabilities

182 of the 206 HPI items were subject to categorization based on ABC content definitions. A total of five raters completed the coding and Fleiss' Kappa was .49. Based on the interpretive ranges presented by Landis and Koch (1977), this can be interpreted as a

moderate degree of agreement. Items categorized in the same ABC component by three or more coders (94%) were used for creating the ABC subscale. Items that did not have substantial agreement (6%) were categorized by the authors in a final expert content analysis and added to the subscales. Percentages of scale items in each ABC component subscale are presented in Table 1. Reliability alphas for each scale and subscale are presented in Table 2. Alphas were higher for the total scale, and we saw substandard alphas for some of the ABC subscales.

Through the content coding process, we found that the HPI scales did not have equivalent item representation within each of the ABC components. Behavioral items were the least represented across the HPI scales. One scale, interpersonal sensitivity, had zero behavioral items. Affect and cognition were roughly equally represented for the HPI in general but differed by scale. Inquisitive and Learning Approach, two constructs that are subsumed within the facet of openness to experience, showed opposite representation of affect and cognition. This strong representation of affect items, suggests that the Inquisitive scale measures more of the desire for information, while the Learning Approach scale is built more on a system of beliefs regarding information and learning.

Correlations with Performance

Table 2 provides average correlations from the two criterion studies. Further, we present averages of the strengths of the correlations. These were computed as an average of the absolute value of the correlations. These correlations are presented with the correction for unreliability of the criterion (.52). Correlations are presented in this manner for each of the HPI scales and each respective ABC component subscale. We see that the behavioral component subscales predict the same as, or better than the total scale in 10 of the 12 comparisons (excluding Interpersonal Sensitivity). Further, we see that the behavioral component subscales predict the same as, or better than the affective and cognitive component scales for 9 of the 12 comparisons (excluding Interpersonal Sensitivity).

Discussion

The primary goal of this study was to explore the hypothesis that behavioral items would be better suited to predict job performance due to the behavioral nature of job performance. To examine this hypothesis, we used two large validation studies ($N = 234$, $N = 319$) with available personality data and performance information. Based on the average strength of these correlations, we saw that the behavioral component subscales tended to correlate on par, or above the total scales and affective or cognitive component subscales. However, there was not always consistency across the two studies. For example, the Ambition behavioral component subscale was the strongest predictor of sales performance compared to the Ambition total scale score, and affective and cognitive component subscales. However, the total scale, affective and cognitive component subscales were the best predictors of leadership performance. This may be due to the difference in the job requirements and performance appraisal processes for each job. For instance, Ambition may be driven more by affective and cognitive components in a leadership position, where sales roles are more visibly Ambitious, aligning with the behavioral component. Because performance is ultimately a behavior, it does make sense that even though the behavioral subscales were not always stronger predictors than the other two scales, they did predict

equally or better than the total score in two-thirds of the correlations for leaders. In the sales representative study, the behavioral items performed as predicted, it was equal to or better than all of the other scales except for the Inquisitive scale. This could be due to the lack of representation of behavioral items within the Inquisitive scale (8%) leading to range restriction. Ultimately, the hypothesis that creating alignment of predictors and criterion variables based on ABC content to increase prediction was tentatively supported.

Limitations, Implications, and Future Directions

This research was initiated as an exploratory look into the use of ABC theory to differentiate between the predictive ability of ABC scales. This was applied to the HPI, which was not developed to consider ABC subscales. From the results, we see that some scales had a small proportion of items in at least one of the three categories. In addition, a number of the scales had alpha scores that were not acceptable or anomalous. Although these scales had less than desirable alphas, this did not affect the correlations. For example, the Ambition Behavioral scale had an anomalous alpha (-.07), yet the strength of the correlations was the strongest, on average, with performance (Study 1, $r = .19$). We saw this based on a consistent endorsement of items in one direction, where individuals that did not endorse all of the items tended to have lower performance scores. The lack of variance signifies range restriction and this may need to be accounted for in further studies. Future studies may attempt to use an assessment that was developed based on ABC components to create a scale that is psychometrically sound when broken into subscales.

Further, this research only examined two criterion studies to explore the differences in prediction of ABC subscales. The correlations were not equally as powerful across studies. This may be due to the differences in job requirements and performance outcomes. Future research should explore how differing jobs result in shifts in the nature of the ABC components of performance. A more effective approach will align performance variables across multiple studies and look at the predictive ability of ABC component subscales across these studies for a more accurate conclusion on the effectiveness of the ABC components in predicting performance. However, we also point out that a rater's view of performance is almost entirely based on behaviors, therefore corresponding performance ratings should be more aligned with behavioral subscales. The relationship may also be moderated by the degree to which the performance appraisal directly references behaviors, as with a behaviorally anchored rating scale (Smith & Kendall, 1963).

Future research should also attempt to align other criteria with the ABC components. We predict that by aligning affective, behavioral, and/or cognitive outcomes with corresponding predictive ABC scales, researchers and practitioners may be better able to predict important workplace criteria. Future research should examine the prediction of job attitudes using ABC alignment. Since job attitudes are an affective and cognitive evaluation of a job aspect (Brummel & Bowling, 2013) it would make sense that their items are likely reflecting affect and cognition, and would presumably be best predicted by affective and cognitive subscales of personality.

From the perspective of the HPI and its measurement of reputation, it may be possible to consider a complex use of ABC theory to drive prediction and development. Behavioral scales could provide key insights into the selection of performance, while cognitive and

affective scales could help to understand underlying drivers of performance, and thus, influence development and other distal outcomes. Alignment of personality ABC components with organizational criteria of interest could provide researchers and practitioners alike with improved predictive validity.

References

- Brummel, B. J., & Bowling, N. A. (2013). Personality and job attitudes. In N. Christiansen & R. P. Tett (Eds.) *Handbook of Personality at Work*, 718–743. New York, NY: Routledge.
- Gaddis, B., & Foster, J. (2010, April). *Development of a Competency Model for Entry-Level Selection*. Poster session presented at the meeting of the 25th annual conference of the Society for Industrial and Organizational Psychology, Atlanta, GA.
- Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist*, 51(5), 469–477. doi:10.1037/0003-066X.51.5.469
- Hogan, R., Hogan, J., & Warrenfeltz, R. (2007). Chapter 3: Hogan Personality Inventory Scales. In R. Hogan, J. Hogan, & R. Warrenfeltz (Eds.), *The Hogan Guide: Interpretation and Use of Hogan Inventories* (pp. 29–50). Tulsa, OK: Hogan Press.
- Hsu, C. (2004). The testing of America. *U.S. News and World Report*, 137(9), 68–69.
- Jackson, D. N. (1974). *Personality Research Form Manual*. Research Psychologists Press.
- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: a meta-analysis. *Journal of Applied Psychology*, 87(3), 530.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33, 159–174.
- Hathaway, S. R., & McKinley, J. C. (1951). *Minnesota Multiphasic Personality Inventory; Manual (Revised)*.
- Motowidlo, S. J. (2003). Job Performance. In APA (Ed.) *Handbook of Psychology*. (pp.39–53).
- Ones, D. S., & Viswesvaran, C. (1996). Bandwidth–fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior*, 17(6), 609–626.
- 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(6), 847–858.
- Rauthmann, J. F., & Denissen, J. J. (2011). I often do it vs. I like doing it: Comparing a frequency-and valency-approach to extraversion. *Personality and Individual Differences*, 50(8), 1283–1288.
- Roberts, B. W. (2007). Contextualizing personality psychology. *Journal of personality*, 75(6), 1071–1082.

- 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.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel psychology*, 44(4), 703–742.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 557–574.
- Werner, P. D., & Pervin, L. A. (1986). The content of personality inventory items. *Journal of Personality and Social Psychology*, 51(3), 622.
- Wilt, J., & Revelle, W. (2015). Affect, Behaviour, Cognition and Desire in the Big Five: An Analysis of Item Content and Structure. *European Journal of Personality*, 29(4), 478–497.

Table 1**HPI ABC Subscale Component Item Distributions**

HPI Scale	Subscale Component	Item Percentage (%)
Adjustment	Affect	35%
	Behavior	24%
	Cognition	41%
Ambition	Affect	52%
	Behavior	10%
	Cognition	38%
Sociability	Affect	67%
	Behavior	17%
	Cognition	17%
Interpersonal Sensitivity	Affect	41%
	Behavior	0%
	Cognition	59%
Prudence	Affect	32%
	Behavior	19%
	Cognition	48%
Inquisitive	Affect	64%
	Behavior	8%
	Cognition	28%
Learning Approach	Affect	14%
	Behavior	29%
	Cognition	57%

Note. Item Percentage is the percentage of total scale items categorized in each of the ABC Component subscales.

Table 2

Correlation Comparisons Between Overall HPI Scales and ABC Subscales in Two Criterion Studies

	Study 1: Sales Performance (n = 174-234)					Study 2: Leader Performance (n = 292-319)				
	Alpha	Avg r	Avg Strength r	Corrected Avg r	Corrected Avg Strength r	Alpha	Avg r	Avg Strength r	Corrected Avg r	Corrected Avg Strength r
Adjustment	.85	.00	.03	.00	.05	.80	.02	.05	.02	.07
ADJ_A	.70	.02	.04	.03	.06	.65	.03	.05	.04	.07
ADJ_B	.58	-.03	.05	-.04	.07	.60	-.04	.05	-.05	.08
ADJ_C	.67	.00	.04	.00	.05	.62	.04	.06	.05	.09
Ambition	.72	.03	.07	.05	.10	.76	.09	.10	.13	.14
AMB_A	.59	.01	.06	.01	.09	.66	.08	.09	.11	.12
AMB_B	-.07	.13	.14	.17	.19	.08	.05	.06	.07	.09
AMB_C	.49	.02	.05	.03	.07	.51	.09	.09	.12	.13
Sociability	.82	-.04	.06	-.06	.09	.79	.06	.07	.09	.10
SOC_A	.76	-.03	.05	-.04	.07	.72	.06	.07	.09	.10
SOC_B	.60	-.05	.06	-.07	.09	.33	.07	.08	.10	.10
SOC_C	.43	-.04	.06	-.06	.08	.60	.00	.03	.00	.04
Interpersonal										
Sensitivity	.58	-.08	.10	-.12	.13	.61	.03	.05	.04	.07
INT_A	.40	-.08	.10	-.11	.13	.45	.01	.05	.02	.07
INT_B	-	-	-	-	-	-	-	-	-	-
INT_C	.39	-.06	.07	-.08	.10	.45	.03	.04	.04	.05
Prudence	.69	.00	.04	.00	.06	.66	-.04	.06	-.06	.08
PRU_A	.35	-.02	.05	-.03	.06	.42	.01	.04	.01	.06
PRU_B	.38	.09	.09	.12	.12	.35	-.09	.09	-.12	.13
PRU_C	.51	.00	.04	.00	.05	.46	-.03	.04	-.04	.06
Inquisitive	.75	-.11	.11	-.15	.15	.77	-.01	.05	-.01	.07
INQ_A	.66	-.11	.11	-.15	.15	.71	-.01	.05	-.01	.06
INQ_B	.31	-.08	.08	-.11	.11	.18	-.05	.07	-.06	.09
INQ_C	.42	-.03	.05	-.05	.07	.50	.02	.04	.02	.06
Learning										
Approach	.75	-.05	.05	-.07	.07	.70	.03	.04	.03	.06
LRN_A	.66	.02	.05	.03	.07	.06	.02	.04	.03	.05
LRN_B	.31	-.06	.07	-.09	.10	.50	.05	.06	.07	.08
LRN_C	.42	-.06	.06	-.08	.08	.56	.01	.05	.01	.06

Note. Average Strength of Correlations calculated using absolute values of correlations; Correlations corrected for Unreliability of the Criterion Measures; A = Affective Subscale; B = Behavioral Subscale; C = Cognitive Subscale.