An SJT to predict expressions of subclinical personality disorders at work[1]

Sheila K. Keener
Old Dominion University

Michael A. McDaniel
Work Skills First

Sven Kepes
Virginia Commonwealth University

Although cognitive ability is generally considered the best predictor of job performance (e.g., Schmidt & Hunter, 2004), some individuals who perform well on cognitive ability assessments may not ultimately have high job performance. For instance, there is evidence to suggest that expressions of certain personality disorders have a negative effect on job performance (e.g., Moscoso & Salagado, 2004). Yet, many psychologists now consider personality to exist on a continuum, such that individuals may have some symptoms associated with personality disorders (and decreased job performance), but not a clinically diagnosable disorder (e.g., De Fruyt & Salagado, 2003; Trull & Durrett, 2005). Therefore, as part of a selection process, it may be beneficial to identify individuals that are likely to exhibit behaviors associated with subclinical levels of personality disorders in the workplace.

Situational judgement tests (SJTs) are particularly well suited to assessing subclinical levels of personality as part of an employment selection process. First, SJTs are low fidelity simulations (e.g., Motowidlo, Dunnette, & Carter, 1990) that use workplace-specific scenarios and response options. Second, SJTs with knowledge instructions (e.g., how *effective* is each behavior likely to be) are less susceptible to faking than typical personality inventories (Nguyen,

[1] Paper accepted at the 34th Annual Conference of the Society for Industrial and Organizational Psychology. Austin, TX. Conference cancelled (Covid-19).

Biderman, & McDaniel, 2005) and may, therefore, provide more accurate assessments than traditional personality inventories. Third, evidence suggests that SJTs provide incremental validity over cognitive ability assessments (McDaniel, Hartman, Whetzel, & Grubb, 2007). Thus, including them in test batteries is like to result in better overall prediction. Therefore, in this paper, we develop an SJT whose scenarios and responses sought to assess tendencies toward expressions of personality disorders in the workplace. To assess the construct validity of the SJT, we also administered a personality disorder inventory and examined the correlations between predictor and criteria scales.

Method

Sample

The sample consisted of 1,651 respondents, some of whom were undergraduates in a southeastern university and who completed the measures as part of a psychology department course requirement. The remaining respondents were drawn from adult U.S. residents through the Amazon Mechanical Turk Platform and were paid for their participation.

Measures

A 23-scenario SJT with 162 scorable response options was developed. The response options included behaviors that would be attractive to those with subclinical levels of personality disorders. The effectiveness of each of the 162 items were rated by respondents on a 6-point Likert scale. The SJT was scored with two consensus approaches (raw consensus and standardized consensus; McDaniel, Psotka, Legree, Yost, & Weekley, 2011) as well as elastic net regression (Friedman, Hastie, & Tibshirani, 2008) and extreme gradient boosting (Chen & Guestrin, 2006). Respondents were also administered the PDQ-4 (Hyler, 1994), which assesses standing on 12 personality disorders dimensions. Due to IRB restrictions, only 11 of the

personality disorder dimensions were assessed. The PDQ-4 was scored consistent with the instructions provided by the test publisher. Scale scores on the PDQ-4 were aggregated into higher-order factors and also summed to yield a personality disturbance scale.

Analyses

Higher-order factors of the PDQ-4 were derived based on the three cluster-model of personality disorders (American Psychiatric Association, 2000, 2013) and confirmatory factor analyses (CFA). CFA models were developed in one random sample and evaluated in a confirmation sample. Analyses involving elastic net regression and extreme gradient boosting require a training sample to develop the prediction model and a test sample to evaluate the prediction model. Each PDQ-4 scale has its own elastic net model and its own extreme gradient boosting model. The training sample was 1,000 randomly drawn observations. The remaining 651 observations were the test sample. The raw and consensus scale development did not rely on knowledge of the PDQ-4 derived scales and did not require a training and test sample. However, for consistency across analyses, we developed the consensus scales in the training sample and reported their validity in the test sample. Following McDaniel et al. (2011), the two consensus scoring scales were calculated using all 162 items and again without the items whose means were in the mid-range of the Likert scale. All SJT scale correlations with the PDQ-4 derived scales were calculated on the test sample. The SJT scales and the PDQ-4 were scored such that high scores indicated possible subclinical personality disorders.

Results

Results of the CFAs found that three of the 11 personality disorders loaded on multiple factors (i.e., narcissistic, histrionic, and obsessive-compulsive personality disorders). Therefore, these were excluded from analyses at the cluster-level (see Table 3). Results for all examined

models, including the final model, are presented in Table 1. Table 2 shows the 15 personality disorder scales (and their definitions) that the SJT scales sought to predict. Table 3 presents the correlations of six SJT scales with 15 personality disorder scales.

Both raw and standardized consensus scales performed poorly relative to the elastic net regression and extreme gradient boosting scales. The SJT was best at predicting the Cluster A ("odd or eccentric") scale and the Paranoid, Schizoid, Schizotypal, Histrionic, and Narcissistic personality disorder subscales, as well as the overall personality disturbance scale.

Discussion

The purpose of this study was to create an SJT to assess subclinical levels of several personality disorders, develop several potential scoring methods for the SJT, and collect construct validity evidence regarding how correlated the SJT scales are with the PDQ-4. Results indicated that the SJT predicted overall personality disturbance, Cluster A personality disorders, the overall personality disturbance scale, and some specific personality disorders well. Therefore, this study shows that SJTs can be an effective way to identify individuals who likely to display behaviors associated with subclinical levels of personality disorders at work. Results of this study also indicated that the two most common methods of scoring SJTs (the raw consensus and standardized consensus methods) did not perform as well as elastic net regression and extreme gradient boosting. Thus, other SJTs may benefit from the use of these methods. Future research should build on this study by examining the predictive validity of this SJT, as well as its incremental validity over cognitive ability and traditional measures of the Big 5 personality traits.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders,* Fourth Edition, Text Revision. DSM-IV-TR™. Washington, DC.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining (pp. 785-794). ACM.

De Fruyt, F., & Salgado, J. (2003). Applied personality psychology: Lessons learned from the IWO field. *European Journal of Personality*, *17*(S1), S123-S131.

Friedman, J., Hastie, T. and Tibshirani, R. (2008). Regularization paths for generalized linear models via coordinate descent, *Journal of Statistical Software, 33*, 1-22.

Hyler, S. E. (1994). *Personality Diagnostic Questionnaire-4 (PDQ-4)*. New York: New York State Psychiatric Institute.

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L., III (2007). Situational judgment tests, response instructions and validity: A meta-analysis. *Personnel Psychology*, 60, 63−91.

McDaniel, M.A., Psotka, J., Legree, P.J., Yost, A.P., & Weekley, J.A. (2011). Toward an understanding of situational judgment item validity and group differences. *Journal of Applied Psychology, 96, 327-336.*

Moscoso, S., & Salgado, J. (2004). “Dark side” personality styles as predictors of task, contextual, and job performance. *International Journal of Selection and Assessment, 12,* 356-362.

Motowidlo, S. J., Dunnette, M. D., & Carter, G.W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology, 75*, 640−647.

Nguyen, N. T., Biderman, M. D., & McDaniel, M. A. (2005). Effects of response instructions on faking a situational judgment test. *International Journal of Selection and Assessment, 13*, 250−260.

Schmidt, F. L., & Hunter, J. E. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 86, 162–173.

Trull, T.J., & Durrett, C.A. (2005). Categorical and dimensional models of personality disorder. *Annual Review of Clinical Psychology*, 1, 355–380.

Table 1. Confirmatory factor analysis of eleven personality disorder scales

Model	χ^2	*df*	CFI	RMSEA	SRMR
Model 1	460.152***	41	.876	.110 (.101, .120)	.062
Model 2	292.404***	32	.911	.099 (.088, .109)	.051
Model 3	177.323***	24	.939	.087 (.076, .100)	.043
Model 4 – Sample 1	109.008***	17	.959	.080 (.066, .095)	.032
Model 4 – Sample 2	134.804***	17	.950	.092 (.078, .107)	.041

Note. *** indicates $p < .001$; Model 1: Cluster A (Paranoid, Schizoid, and Schizotypal personality disorders), Cluster B (Negativistic, Borderline, Histrionic, and Narcissistic personality disorders), Cluster C (Avoidant, Depressive, Dependent, and Obsessive-Compulsive personality disorders); Model 2: Cluster A (Paranoid, Schizoid, and Schizotypal personality disorders), Cluster B (Negativistic, Borderline, and Histrionic personality disorders), Cluster C (Avoidant, Depressive, Dependent, and Obsessive-Compulsive personality disorders); Model 3: Cluster A (Paranoid, Schizoid, and Schizotypal personality disorders), Cluster B (Negativistic and Borderline personality disorders), Cluster C (Avoidant, Depressive, Dependent, and Obsessive-Compulsive personality disorders); Model 4: Cluster A (Paranoid, Schizoid, and Schizotypal personality disorders), Cluster B (Negativistic and Borderline personality disorders), Cluster C (Avoidant, Depressive, and Dependent personality disorders)

Table 2. Personality Disorder Scale Definitions[a]

Personality disorders grouped into three clusters based on descriptive similarities	
Cluster A	Includes Paranoid, Schizoid, and Schizotypal personality disorders. Individuals with these disorders often appear odd or eccentric.
Cluster B	Includes Antisocial[b], Borderline, Histrionic[c] and Narcissistic[c] personality disorders. Individuals with these disorders often appear dramatic, emotional or erratic.
Cluster C	Includes Avoidant, Dependent, and Obsessive Compulsive[d] personality disorders. Individuals with these disorders often appear anxious or fearful.
Primary Scales	
Paranoid	Is a pattern of distrust and suspiciousness such that others' motives are interpreted as malevolent
Schizoid	Is a pattern of detachment from social relationships and a restricted
Schizotypal	Is a pattern of acute discomfort in close relationships, cognitive or perceptual distortions, and eccentricities of behavior
Borderline	Is a pattern of instability in interpersonal relationships, self-image, and affects, and marked impulsivity
Histrionic	Is a pattern of excessive emotionality, and attention seeking
Narcissistic	Is a pattern of grandiosity, need for admiration, and lack of empathy
Avoidant	Is a pattern of social inhibition, feelings of inadequacy, and hypersensitivity to negative evaluation
Dependent	Is a pattern of submissive and clinging behavior related to an excessive need to be taken care of
Obsessive-Compulsive	Is a pattern of preoccupation with orderliness, perfectionism, and control
Summed Scale	
Personality Disturbance	The sum of all primary scales.

Notes. [1]All definitions in the second column of this table, excepting the Personality Disturbance definition, are drawn verbatim from American Psychiatric Association (2000), page 685; [b] The antisocial personality scale was not used in this research because some items reference behavior that could be criminal. This decision was made as part of communication with a university human subjects committee; [c] In our analyses, histrionic and narcissistic were not included in Cluster B analyses because their inclusion harmed confirmatory factor model. The American Psychiatric Association (2000, p. 686) noted that the cluster taxonomy has not been validated consistently; [d] In our analyses, Obsessive-Compulsive was not included in Cluster C analyses because its inclusion harmed the confirmatory factor model. The American Psychiatric Association (2000, p. 686) noted that the cluster taxonomy has not been validated consistently.

Table 3. Correlations of test sample SJT scales (N = 651) with personality disorder constructs by type of SJT scoring method

Constructs	Raw Consensus		Standardized Consensus		Elastic Net	Extreme Gradient Boosting
	All items	Midrange items dropped	All items	Midrange items dropped	All items	All items
Cluster A "odd or eccentric"	.05	.12	.18	.20	.30	.30
Cluster A subscales						
Paranoid	.03	.08	.14	.14	.27	.25
Schizoid	.09	.11	.18	.15	.24	.23
Schizotypal	.00	.09	.11	.16	.29	.27
Cluster B "dramatic, emotional, or erratic"	-.08	-.01	.02	.08	.22	.19
Cluster B subscales						
Negativistic	-.06	.00	.03	.07	.16	.13
Borderline	-.07	-.01	.02	.07	.19	.16
Cluster C "anxious or fearful"	-.08	-.05	.00	.03	.13	.19
Custer C subscales						
Avoidant	-.08	-.07	-.02	.00	.15	.18
Depressive	-.08	-.08	-.01	-.02	.11	.14
Dependent	-.03	.04	.05	.10	.17	.17
Scales not included in clusters						
Histrionic[a]	.03	.11	.05	.11	.25	.27
Narcissistic[b]	.03	.15	.16	.23	.35	.34
Obsessive-Compulsive[c]	-.02	-.01	-.02	.00	.14	.10
Personality Disturbance[d]	-.03	.05	.09	.14	.25	.25

Notes. [a]Histrionic is typically in Cluster B but was removed to provide an acceptable CFA fit; [b]Narcissistic is typically in Cluster B but was removed to provide an acceptable CFA fit; [c]Obsessive-Compulsive is typically in Cluster C but was removed to provide and acceptable CFA fit; [d]Personality Disturbance is the sum of all PDQ items.