Derailers vs. Personality Disorders: What are the Differences?

Heather Hayes, Camber Corporation

Jeff Foster, Hogan Assessment Systems

Blaine Gaddis, Hogan Assessment Systems

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

Hogan Research Division

Abstract

There remains little consensus regarding the structure and meaning of personality derailers. The current research aims to fill this gap by comparing items from the HDS and the PID-5. Results support the conception of derailers as personality constructs that align with disorders but are not clinically debilitating.

Introduction

Interest in personality derailers continues to grow, as indicated by a number of recent publications such as reviews (e.g., Spain, Harms, & LeBreton, 2013), special issues (e.g., Harms & Spain, 2015), and focal articles (e.g., Guenole, 2014). Sometimes called dark-side or maladaptive personality, derailer scales generally measure characteristics that can negatively affect job performance and may be disastrous for one's career. However, derailment measurement currently suffers from substantial inconsistencies in terminology, measurement, structural models, and nomenclatures (Foster & Gaddis, 2014). Although other models exist, such as the Dark Triad (Furnham, Richards, & Paulhus, 2013; Wu & Le Bretton, 2011), one of the most common approaches to derailer measurement is to create scales that align with personality disorders from the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR; American Psychiatric Association, 2000). Unfortunately, this process results in confusion between derailers and clinical assessments, particularly with respect to how they differ in score interpretation.

We seek to alleviate this confusion by examining differences in the psychometric properties of items from one of the most widely used personality derailer instruments in the world, the Hogan Development Survey (HDS; Hogan & Hogan, 2009), and the Personality Inventory for DSM-5 (PID-5; Krueger, Derringer, Markon, Watson, & Skodol, 2012), which was designed to assess personality disorders based on the most recent version of the DSM (DSM-V, American Psychiatric Association, 2013).

Derailment versus Clinical Assessments

It is not surprising that derailment scales are often confused with clinical assessments. Researchers often describe scales like those on the HDS using a variety of clinically-related terms such as sub-clinical (e.g., Harms, Spain, & Hannah, 2011), symptoms (e.g., Douglas, Bore, & Munro, 2012), or even go so far as calling them measures of personality disorders (e.g., Furnham, 2006; Furnham & Trickey, 2011). For most derailment instruments, however, these descriptions are inaccurate and misleading. This is the case when derailer instruments are intended for use with normal adult working populations and have no scoring options for making clinical diagnoses, such as with instruments like the HDS (Hogan & Hogan, 2009).

However, simply stating that an instrument isn't intended for use with clinical populations or does not including scoring that could lead to a diagnosis doesn't necessarily mean that clinicians could not use it for such purposes. Instead, the critical difference between measures like the HDS and clinical diagnostic assessments is that the HDS was developed for use with normal working populations to predict behaviors that may hinder job

performance, but are not so extreme as to likely result in an inability to perform normal daily functions. In other words, the HDS assesses individual characteristics that can still be functional, if not beneficial, within certain situations (Hogan & Hogan, 2009). Therefore, although these behaviors might ultimately alienate supervisors and coworkers, they do not reach such extremes as to represent maladaptive behaviors that can hinder ones' abilities to properly function within society.

Therefore, we propose that items on a derailer instrument should align with items written to measure similar albeit abnormal personality derailers, but will, on average, represent less extreme measures within these aligned constructs. In other words, when using a statistical approach such as Item Response Theory (IRT; Embretson & Reise, 2000; Hambleton & Swaminathan, 1985) to indicate item location and discrimination parameters for items measuring the same construct, items from derailer measures will be significantly lower on the latent trait continuum than items for clinical measures while also discriminating more effectively among individuals who lie on the lower end of this continuum. Therefore, we propose the following two hypotheses:

> *Hypothesis 1: Scales on the HDS will align with scales on the PID-5 developed to measure associated personality disorders.*
>
> *Hypothesis 2a: HDS items will be more informative and discriminate among individuals at the lower end of the latent trait continuum in comparison with items from the PID-5.*
>
> *Hypothesis 2b: Item location parameters for HDS items will be, on average, significantly lower than item location parameters from the PID-5.*

Method

Sample

Our sample consisted of 326 individuals who participated in a series of online assessments on MTurk during the summer of 2016. Participant age ranged between 18 and 75 (M = 35; SD = 11.43). This sample included 47% male respondents and 53% female respondents. Among participants who reported race/ethnicity, 76.4% were White, with Black/African-American (7.9%), Hispanic/Latino (4.6%), Asian (5.6%), American Indian/Alaska Native (1.0%), Two or More Races (2.2%), Native Hawaiian/Other Pacific Islander (0.3%), 0.1% representing other reported racial/ethnic groups. With respect to job characteristics, 50.8% of the sample reported being employed full-time, 15.4% part-time, 16.9% self-employed, 4.3% full-time students, 2.0% retired, and 9.5% unemployed (1.1% did not respond). Moreover, 34.1% of the sample reported their employer as being public, 55.5% private, 7.4% non-profit, and 3.0% did not respond. In terms of organizational size, 25.8% of the participants reported working with fewer than 10 people, 9.8% 10-24 people, 8.3% 25-49, 4.8% 50-74, 10.1% 75-199, 8.0% 200-499, and 30.3% 500 or more.

Measures

The HDS measures 11 dysfunctional dispositions that negatively influence the ability to get along with others and get ahead in careers (Horney, 1950). Although intended for use in normal (i.e., non-clinical) populations, these dispositions theoretically resemble less extreme levels of various personality disorders from the DSM-II (R. Hogan & J. Hogan, 2009). For example, the Excitable scale of the HDS is theoretically similar to Borderline Personality Disorder, and Skeptical is somewhat akin to Paranoid Personality Disorder. Table 1 presents descriptions of each HDS scale. The response format for each item is True/False.

The PID-5 measures 25 maladaptive traits that are based on the DSM-5 trait model of personality pathology (Krueger, Derringer, Markon, Watson, & Skodol, 2012). Empirical research supports an alignment between personality disorders from the DSM-5 and particular configurations of PID-5 scales (Strickland, 2014). For example, Narcissistic Personality Disorder is a composite of the following two PID-5 scales: Grandiosity and Attention-Seeking. The response scale is as follows: "Very Untrue of Me," Somewhat Untrue of Me," "Somewhat True of Me," and "Very True of Me."

Scale Matching Procedure

We matched items from the PID-5 and HDS in three stages. First, we identified the 11 DSM-5 personality disorders theoretically associated with each of the 11 HDS scale (see Table 2). All personality disorders from the DSM-5 are associated with multiple PID-5 scales. Consequently, these disorders are multidimensional in terms of latent construct representation. A prerequisite for the current IRT analysis is unidimensionality because the matched items will be compared on the same latent trait continuum. Thus, it is inappropriate to collapse multiple PID-5 scales into one scale. Consequently, we ultimately chose only one PID-5 scale for comparison with a given HDS scale. Given that our primary research question deals with where items aligning on the same continuum fall in terms of difficulty parameters (see below), we felt it is e adequate to compare single PID-5 scales to individual HDS scales because of the final number of comparisons and comprehensiveness of all scales included in our analyses.

Second, subject matter experts (SMEs; three Ph.D. researchers with a combined experience of 45 years in personality assessment and psychometrics) independently examined the item content of the PID-5 and HDS scales and matched those most closely related to one another. Specifically, for each personality disorder, SMEs ranked the PID-5 scales on similarity in content with the corresponding HDS scale. We aggregated the results, and SMEs met to resolve conflicts and reach a consensus.

Third, we computed correlations between each HDS scale and all potentially-matched PID-5 scales from Table 2. We retained the PID-5 scale with the highest correlation with the relevant HDS scale. SMEs met to reconcile any inconsistencies between previously established consensus and the HDS-PID-5 correlations patterns in order to determine the final HDS-PID-5 scale matchings.

The aforementioned matching procedure resulted in the following HDS-PID-5 parings: 1) Excitable (HDS) and Hostility (PID-5); 2) Skeptical (HDS) and Suspiciousness (PID-5); 3) Cautious (HDS) and Anxiousness (PID-5); 4) Reserved (HDS) and Withdrawal (PID-5); 5) Bold (HDS) and Grandiosity (PID-5); 6) Colorful (HDS) and Attention-Seeking (PID-5); 7) Imaginative (HDS) and Unusual Beliefs and Experiences (PID-5); 8) Diligent (HDS) and Rigid Perfectionism (PID-5); and 9) Dutiful (HDS) and Submissiveness (PID-5). See Table 3 for the correlations between these matched scales.

Two HDS scales –Leisurely and Mischievous – could not be matched. For Leisurely, the most similar PID-5 scale was Hostility, although this scale was more highly correlated with Excitable (0.37 and 0.64, respectively). And, Hostility was more appropriately matched to Excitable. Multiple PID-5 scales – particularly, Manipulativeness and Impulsivity – were equivalently related to Mischievous based on both SME review and the pattern of correlations. Also, the relationship between these two PID-5 scales was too low (less than .30) to justify collapsing them into one, overall scale. Therefore, we determined there was not a sufficient match for the Leisurely HDS scale.

Analysis

To test Hypothesis 1, thereby confirming unidimensionality of the paired HDS-PID-5 scales, we submitted item responses from each of the matched HDS-PID-5 scales to a series of exploratory factor analyses (EFA) in Mplus (Muthén & Muthén, 2004). However, in order to make the data between the two assessments comparable for analysis, we dichotomized the PID-5 response scales. Previous researchers comparing normal and abnormal personality assessments via factor analysis and IRT have followed a similar approach (Samuel, Sims, Clark, Livesley, & Widiger, 2010; Samuel, Carroll, Rounsaville, & Ball, 2013).

To test Hypotheses 2a and 2b, we used the 2PL model to estimate item parameters and information functions for each HDS-PID-5 scale pairing. In other words, we jointly estimated parameters for all items within a given scale pairing. We chose a 2PL model for the current analysis because the 2 PL provided the best fit to the data when compared to the 1 PL, Rasch, and 3 PL models, and the 2 PL permits variation in discrimination among items, which is a crucial part of the current analysis. Given the small sample size of the current study, we used a Bayesian algorithm (Gibbs sampling package in the R statistical software) to estimate model parameters. The advantage of Bayesian algorithms is that, during the model parameter estimation process, observed information (from a small sample size) is augmented with prior information (i.e., theoretical distributions for each parameter) to avoid issues such as asymptotic results (Fox, van den Berg, & Veldkamp, 2015; Lee & Song, 2004).

To determine the extent to which the HDS and PID-5 items overlap in range of construct coverage along the same (unidimensional) latent trait continuum as well as the degree to which they differ in level of discrimination, we compared item information provided by HDS versus PID-5 items via a) item information curves (Hypothesis 2a) and b) t-test comparisons of mean item location parameters (Hypothesis 2b).

Results

We assessed the degree of unidimensionality attained by each set of HDS and PID-5 items based on multiple indices; namely, root mean square residuals (RMSR), root mean square error of approximation (RMSEA), and comparative fit index (CFI) values for a one-factor model, the eigenvalue ratio of the first to second factors, and the "knee" or bend in scree slopes (Thorpe & Favia, 2012). Although the RMSR was above the recommended level (0.10) for unidimensionality for some HDS-PID-5 matchings (Kline, 2005), none exceeded .17. CFI values were all within acceptable ranges (mid .80s to mid .90s), and RMSEAs were all approximately .05. As recommended by Ruscio and Roche (2012), we found evidence for unidimensionality for all HDS-PID-5 matchings based on the ratio of eigenvalues for factors one to two, all of which exceeded 2.5, and the "knee" or bend in scree slopes, which were most substantial between the first and second factor. Thus, we argue that the HDS-PID-5 matched scales are sufficiently unidimensional, thereby confirming Hypothesis 1 and justifying the subsequent IRT analysis.

For each of the nine HDS-PID-5 matchings, we graphed a mean item information curve (MIC) for HDS items and a second MIC for PID-5 items, and placed them side by side for visual comparison. The nine pairs of MICs associated with each of the nine matchings are located in Figures 1 through 9. Discrepancies in the range of latent trait coverage were most prominent for Imaginative (HDS) versus Unusual Beliefs and Experiences (PID-5), Diligent (HDS) versus Rigid Perfectionism (PID-5), and Dutiful (HDS) versus Submissiveness. Although the differences were not as noticeable for Excitable (HDS) versus Hostility (PID-5), Cautious (HDS) versus Anxiousness (PID-5), Reserved (HDS) versus Withdrawal (PID-5), and Colorful (HDS) versus Attention-Seeking, for all of these matchings, the PID-5 items had higher levels of discrimination than their HDS counterparts. Thus, the range of construct coverage for these PID-5 items, while overlapping with that of HDS items, were more pronounced at the extreme end of the latent trait continuum.

Finally, we examined the mean differences in item location between each scale of the two assessments via a series of one-tailed t-tests. We chose to use the one-tailed test because we are specifically interested in whether or not HDS items are located significantly lower than PID-5 items on the latent trait continuum. The mean item location parameters as well as results of the t-tests for all nine HDS-PID-5 matchings are located in Table 5. The HDS item location parameters were significantly lower on the latent trait continuum than those of their PID-5 counterparts with one exception: Cautious and Anxiousness. Although Anxiousness items were more extreme in value compared to Cautious items, this difference was not statistically significant. Results were significant, however, for the remaining 8 pairings, thereby largely confirming Hypotheses 2a and 2b.

Discussion

Our primary object was to determine if items on the HDS are significantly different from items on the PID-5. Specifically, although they are expected to lie on the same latent trait continuum, to the extent do they provide information at different points this continuum, where the lower end represents maladaptive or derailer tendencies and the higher end represents dysfunctional and clinical personality disorders. Results indicate that (a) HDS

items do align with similar constructs from the PID-5 and (b) HDS and PID-5 items represent different score ranges within the constructs being measured. In other words, derailment scales align with personality disorder scales but measure significantly less extreme characteristics associated with those constructs. Both this alignment and distinction are critical to understanding how to properly measure and define personality derailers. The distinction also has potentially important legal implications in that it helps clarify differences between derailer and clinical measures.

These results also have several important implications for research on personality derailers. First, they show that derailers can be described as characteristics that align with personality disorder measures but are significantly less extreme on those dimensions in terms of the behaviors they represent. Second, they help fill a gap by identifying a consistent set of derailers that align with personality disorder models. Finally, they clearly distinguish derailer measurements from measures of personality disorders. Not only do items occupy different spaces along the same continuums, but at least with the HDS, derailer items are not so extreme to be useful as diagnostic tools.

However, several limitations of the current project should be noted. First, the results of this study may be specific to just the two instruments under investigation, the HDS and PID-5. Future research should examine results from additional derailer and clinical measures. Second, although MTurk samples continue to be more prominent in our field (Cheung, Burns, Sinclair, & Sliter, 2016), research should replicate results with additional working samples. Finally, although our primary hypothesis was largely supported, results were not statistically significant for one scale. Future replications with larger and more diverse samples are needed to determine if results replicate across all potential derailer scales.

Also, while our results demonstrate statistically significant differences across items, they cannot identify an ideal range of difficulty for derailer or clinical assessments. Largely a theoretical question, future work could further distinguish personality derailers from clinical assessments by identifying specific points or thresholds where characteristics turn from potentially problematic on the job to generally debilitating in life. This seems particularly useful given recent attempts to upgrade personality disorders in the DSM-V and the current existence of two different models for measurement and diagnosis. The current study is an important step in this direction and provides important and useful information concerning the nature of and distinction between personality derailers and personality disorders.

References

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

American Psychiatric Association. (n.d.). Retrieved September 6, 2016, from http://www.dsm5.org/Pages/Default.aspx

Cheung, J. H., Burns, D. K., Sinclair, R. R., & Sliter, M. (2016). Amazon Mechanical Turk in organizational psychology: An evaluation and practical recommendations. *Journal of Business and Psychology.* Advance online publication.

Douglas, H., Bore, M., & Munro, D. (2012). Distinguishing the dark triad: Evidence from the five-factor model and the hogan development survey. Psychology, 3, 237–242.

Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. Mahwah, NJ: Lawrence Erlbaum.

Foster, J.L., & Gaddis, B. H. (2014). Personality derailers: Where do we go from here? Industrial and Organizational Psychology: Perspectives on Science and Practice, 7, 403-408.

Fox, J.-P., van den Berg, S. M., & Veldkamp, B. P. (2015). Bayesian psychometric methods. In P. Irwing, T. Booth, & D. Hughes (Eds.) *Handbook of Psychometric Testing.* Wiley-Blackwell: West Sussex, UK.

Furnham, A. (2006). Personality disorders and intelligence. Journal of Individual Differences, 27, 42–46.

Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The Dark Triad of personality: A 10-year review. Social and Personality Psychology Compass, 7, 199-216.

Furnham, A., & Trickey, G. (2011). Sex differences in the dark side traits. Personality and Individual Differences, 50, 517–522.

Guenole, N. (2014). Maladaptive personality at work: Exploring the darkness. Industrial and Organizational Psychology: Perspectives on Science and Practice, 7, 85–97.

Hambleton, R. K., & Swaminathan, H. (1985). Item response theory: Principles and applications. Norwell, MA: Kluwer Academic Publishers.

Harms, P. D., & Spain, S. M. (2015). Beyond the Bright Side: Dark Personality at Work. Applied Psychology: An International Review, 64, 15-24.

Harms, P. D., Spain, S. M., & Hannah, S. T. (2011). Leader development and the dark side of personality. The Leadership Quarterly, 22, 495–509.

Hogan, R., & Hogan, J. (2009). *Hogan Development Survey manual* (2nd Ed.). Tulsa, OK: Hogan Press.

Horney, K. (1950). *Neurosis and human growth*. New York, NY: Norton.

Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*, Thousand Oaks, CA: Sage.

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive trait model and inventory for DSM-5. *Psychological Medicine, 42*, 9, 1879-1890.

Lee, S.-Y., & Song, X.-Y. (2004). Evaluation of the Bayesian and maximum likelihood approaches in analyzing structural equation models with small sample sizes. *Multivariate Behavioral Research, 39*, 653-636.

Muthén, L. K., & Muthén, B. O. (2004). *Mplus: Statistical Analysis with Latent Variables, User's Guide (Version 3.0)*. Los Angeles, CA: Muthén & Muthén.

Ruscio, J., & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment, 24*, 282-292.

Samuel, D. B., Carroll, K. M., Rounsaville, B. J., & Ball, S. A. (2013). Personality disorders as maladaptive, extreme, variants of normal personality: Borderline personality disorder and Neuroticism in a substance using sample. *Journal of Personality Disorders, 27*, 5, 625-635.

Samuel, D. B., Simms, L. J., Clark, L. A., Livesley, W. J., Widiger, T. A. (2010). An item response theory integration of normal and abnormal personality scales. *Personality Disorders*, 5-21.

Spain, S. M., Harms, P., & LeBreton, J. M. (2013). The dark side of personality at work. Journal of Organizational Behavior, 35, 41-60.

Strickland, C. M. (2014). Validating the Personality Inventory for DSM-5: A Trait-Based Model of Personality Disorders. Master's Thesis.

Wu, J., & Le Bretton, J. M. (2011). Reconsidering the dispositional basis of counterproductive work behavior: The role of aberrant personality. Personnel Psychology, 64, 593-626.

Table 1. HDS Scales and Definitions

HDS Scale	Concerns seeming...
Excitable	moody and inconsistent, being enthusiastic about new persons or projects and then becoming disappointed with them.
Skeptical	cynical, distrustful, overly sensitive to criticism, and questioning others' true intentions.
Cautious	resistant to change and reluctant to take even reasonable chances for fear of being evaluated negatively.
Reserved	socially withdrawn and lacking interest in or awareness of the feelings of others.
Leisurely	autonomous, indifferent to other people's requests, and becoming irritable when they persist.
Bold	unusually self-confident and, as a result, unwilling to admit mistakes or listen to advice, and unable to learn from experience.
Mischievous	to enjoy taking risks and testing the limits.
Colorful	expressive, dramatic, and wanting to be noticed.
Imaginative	to act and think in creative and sometimes unusual ways.
Diligent	careful, precise, and critical of the performance of others.
Dutiful	eager to please, reliant on others for support, and reluctant to take independent action.

Table 2: Alignment of HDS and PID-5 Scales with DSM Personality Disorders

Personality Disorder	HDS scale	PID scales
Borderline	Excitable	Emotional Lability Anxiousness Separation Insecurity Hostility Depressivity Impulsivity
Paranoid	Skeptical	Hostility Suspiciousness Intimacy Avoidance Unusual Beliefs and Experiences
Avoidant	Cautious	Anxiousness Withdrawal Anhedonia
Schizoid	Reserved	Restricted Affectivity Withdrawal Intimacy Avoidance
Passive-Aggressive	Leisurely	Hostility Depressivity
Narcissistic	Bold	Grandiosity Attention-seeking
Antisocial	Mischievous	Hostility Manipulativeness Deceitfulness Callousness Irresponsibility Impulsivity Distractibility
Histrionic	Colorful	Perseveration Impulsivity Attention-Seeking Manipulativeness Eccentricity Perceptual Dysregulation
Schizotypal	Imaginative	Perceptual Dysregulation Unusual Beliefs and Experiences
Obsessive-Compulsive	Diligent	Perseveration Rigid Perfectionism
Dependent	Dutiful	Submissiveness Separation Insecurity Anxiousness

Table 3: Correlations Between the Matched HDS and PID-5 Scale Scores

HDS scale	PID-5 Scale	Correlation
Excitable	Hostility	0.64**
Skeptical	Suspiciousness	0.62**
Cautious	Anxiousness	0.54**
Reserved	Withdrawal	0.61**
Bold	Grandiosity	0.53**
Colorful	Attention-Seeking	0.54**
Imaginative	Unusual Beliefs and Experiences	0.31**
Diligent	Rigid Perfectionism	0.30**
Dutiful	Submissiveness	0.38**

*p < .05; **p < .01

Table 4: *t*-test Comparisons of Item Locations

HDS	Mean Item Location	PID	Mean Item Location	t-test
Excitable (n=14)	.32	Hostility (n=10)	0.99	-2.68**
Skeptical (n=14)	-0.37	Suspiciousness (n=7)	0.80	-2.70**
Cautious (n=14)	-0.37	Anxiousness (n=8)	0.49	-1.29
Reserved (n=14)	-0.37	Withdrawal (n=10)	0.31	-1.90*
Bold (n=14)	0.04	Grandiosity (n=6)	1.07	-2.89**
Colorful (n=14)	0.41	Attention-Seeking (n=8)	1.05	-2.12*
Imaginative (n=14)	-0.32	Unusual Beliefs & Experiences (n=8)	1.17	-5.07**
Diligent (n=14)	-1.20	Rigid Perfectionism (n=10)	0.46	-5.38**
Dutiful (n=14)	-0.6	Submissiveness (n=4)	0.37	-2.64**

*p < .05; **p < .01

Note: Number of items for each scale are in parentheses

Figure 1: Excitable versus Hostility

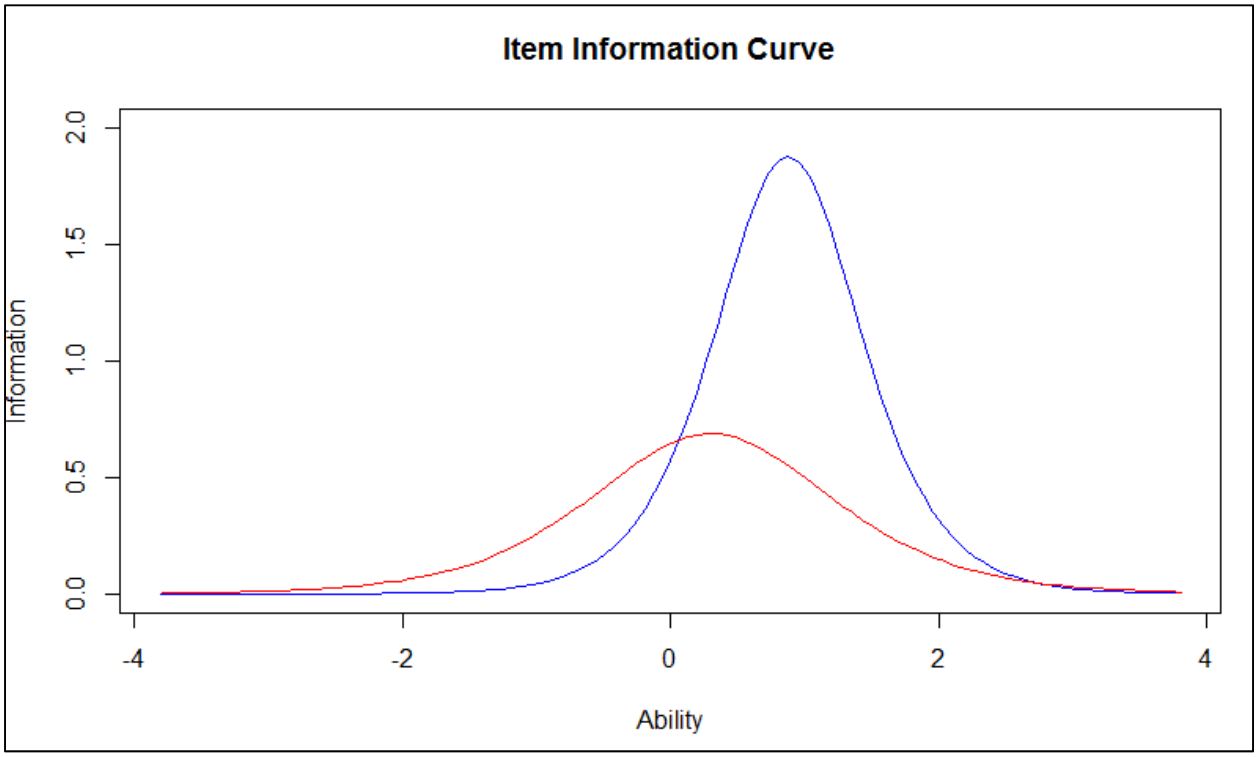

Figure 2: Skeptical versus Suspiciousness

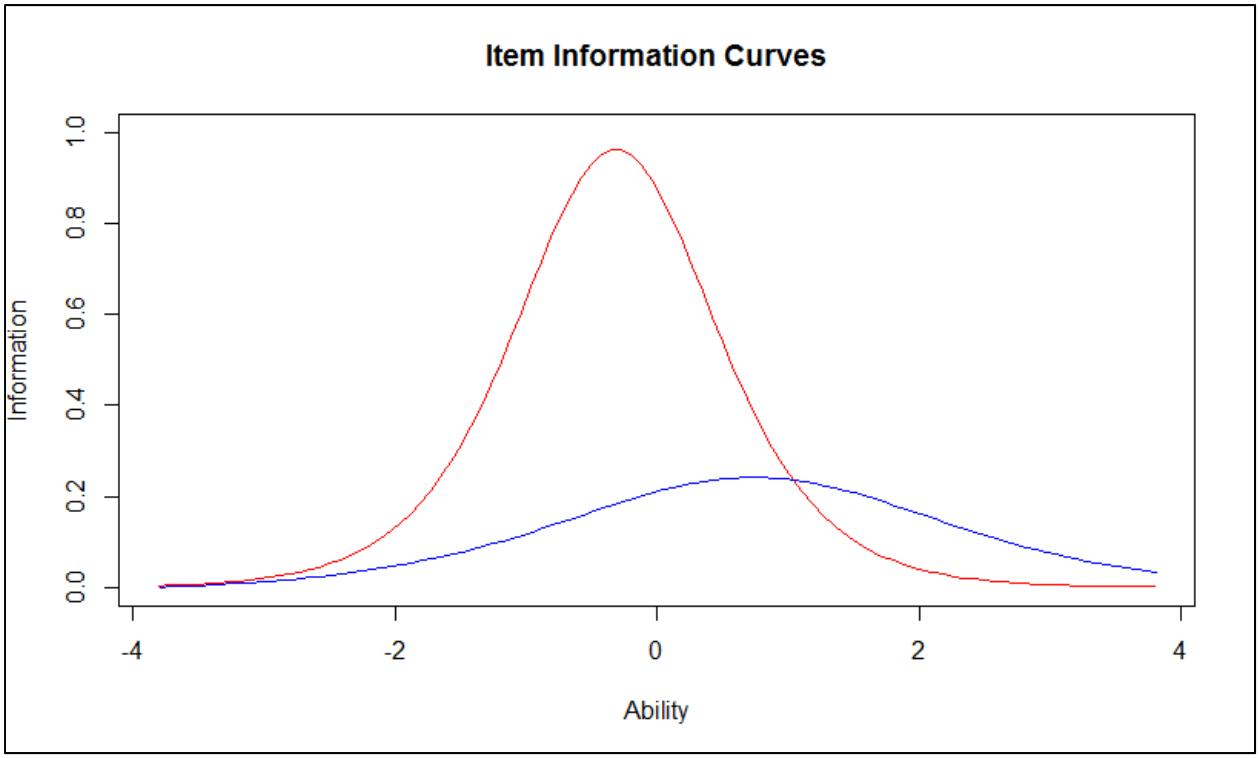

Figure 3: Cautious versus Anxiousness

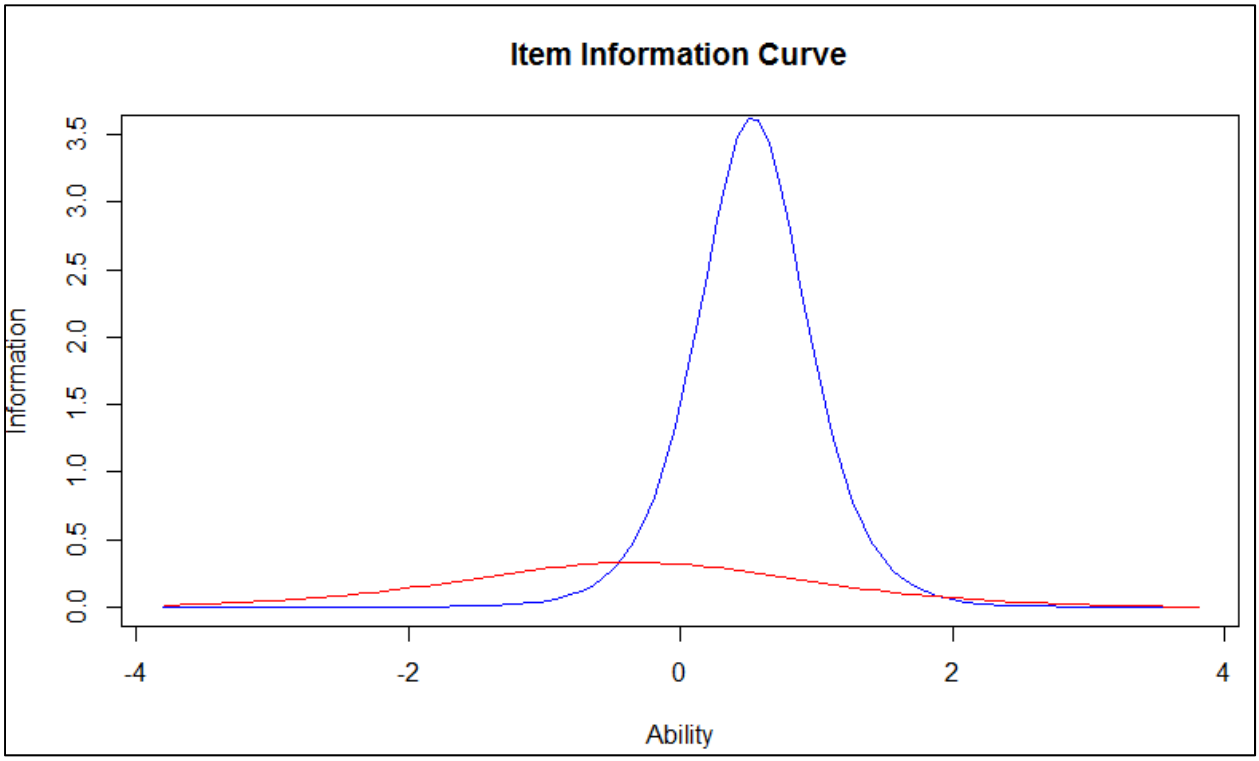

Figure 4: Reserved versus Withdrawal

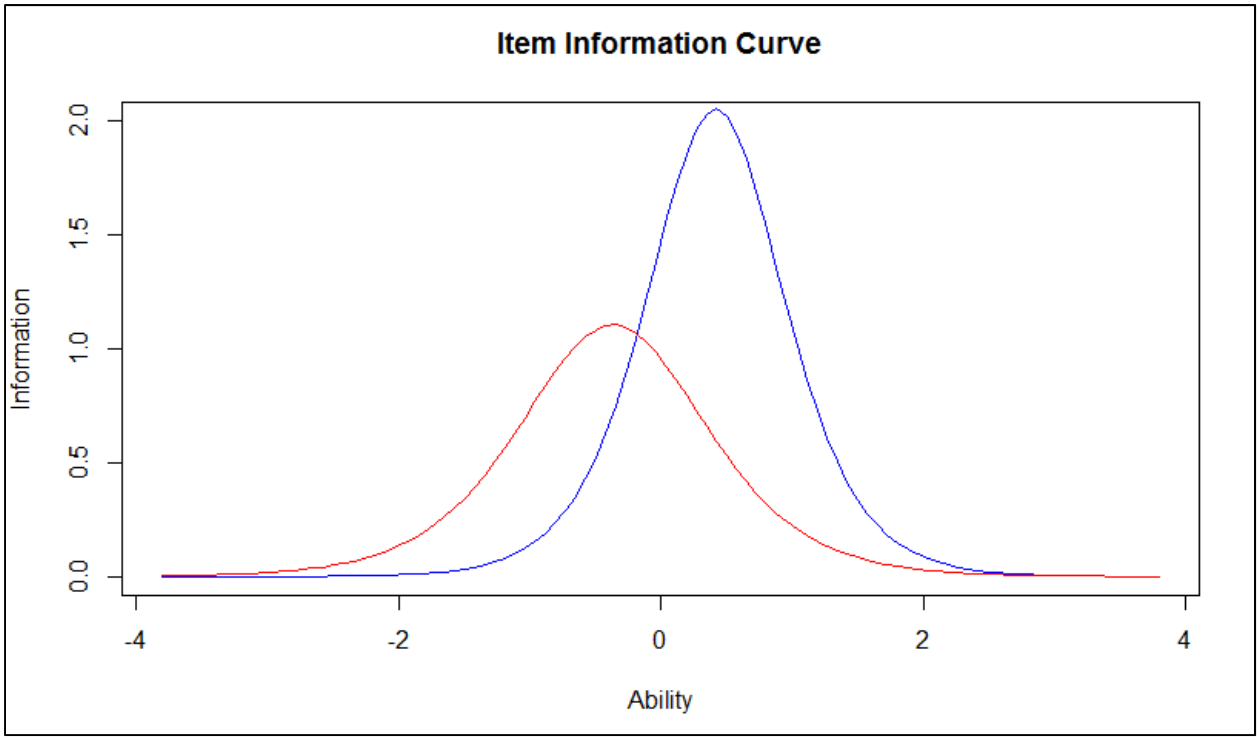

Figure 5: Bold versus Grandiosity

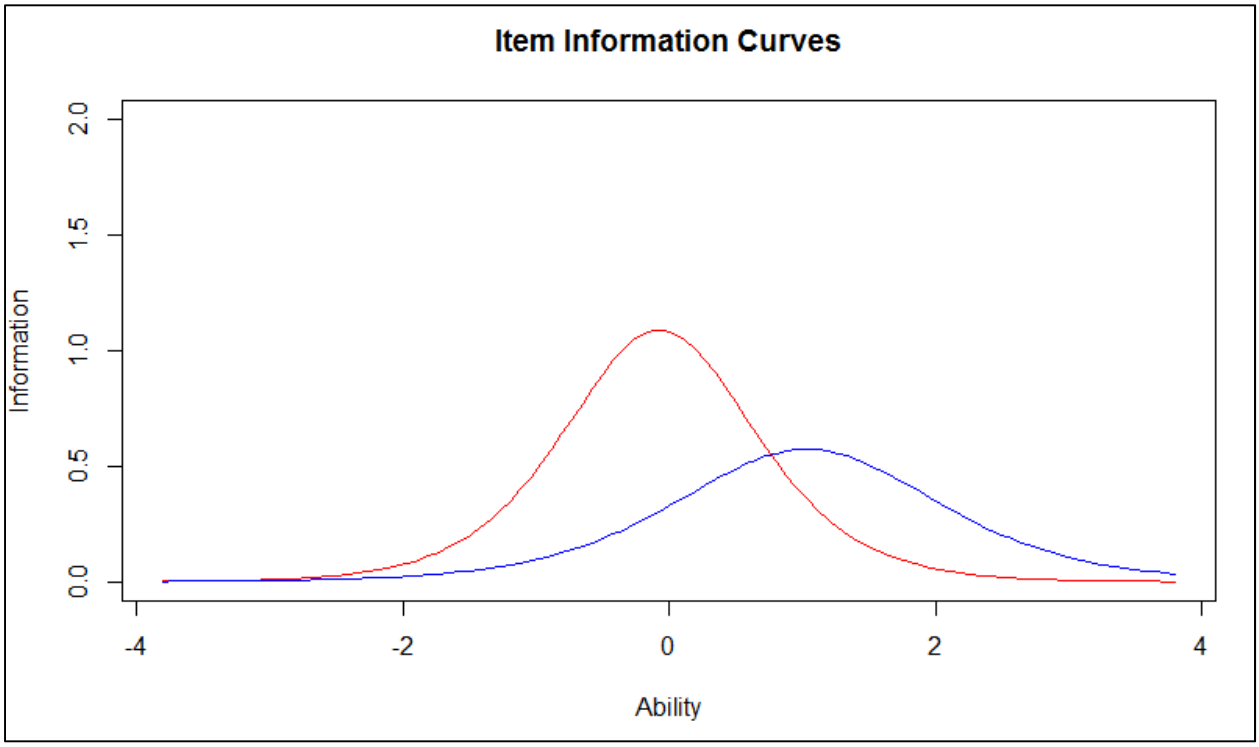

Figure 6: Colorful versus Attention-Seeking

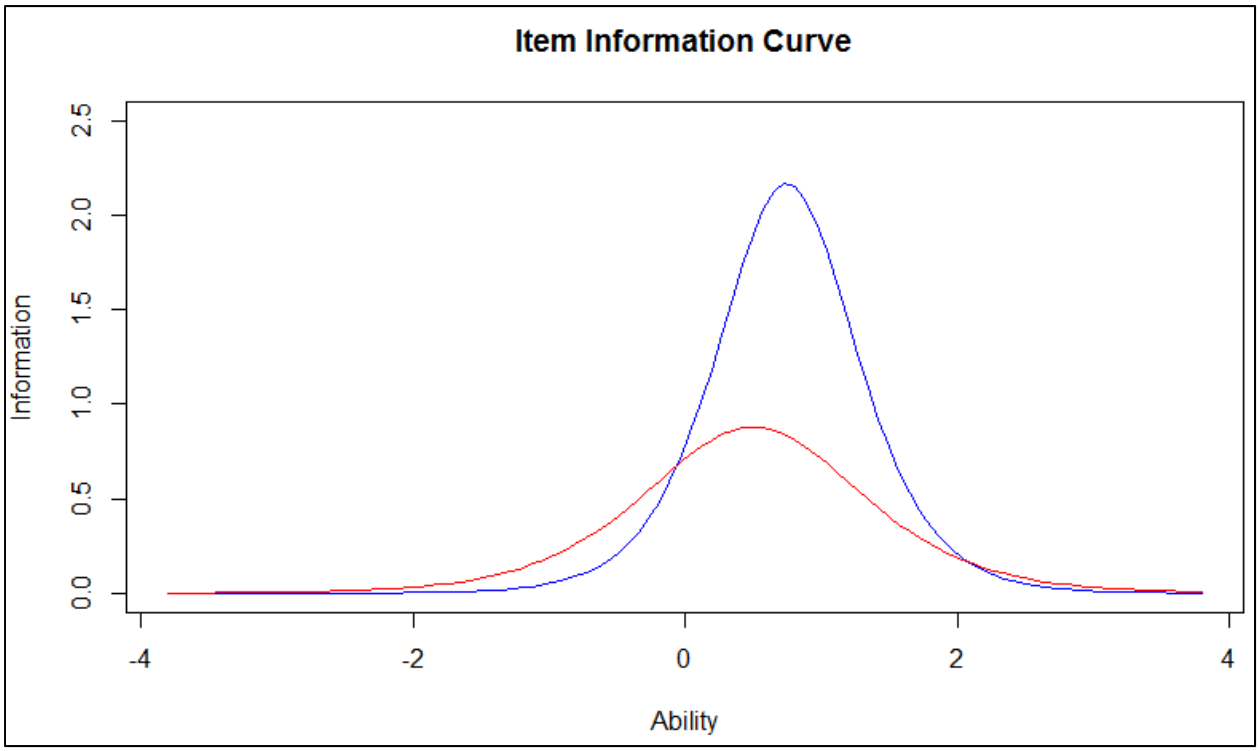

Figure 7: Imaginative versus Unusual Beliefs and Experiences

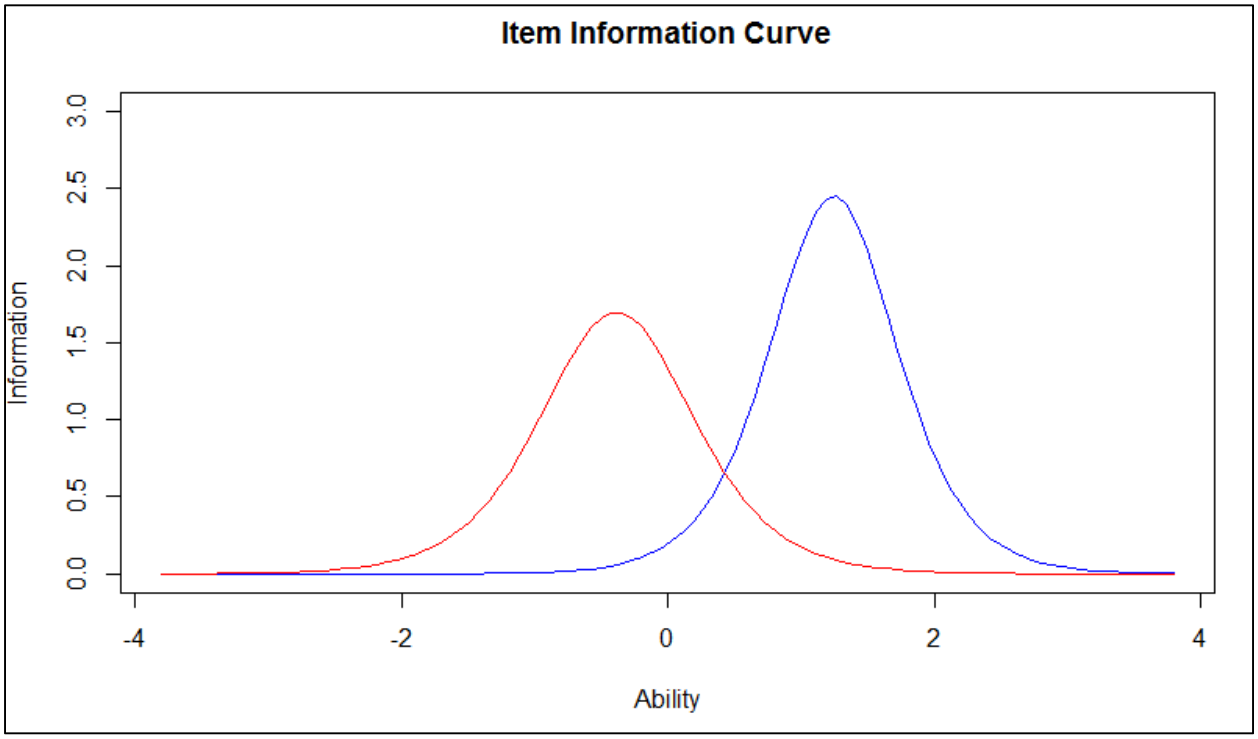

Figure 8: Diligent versus Rigid Perfectionism

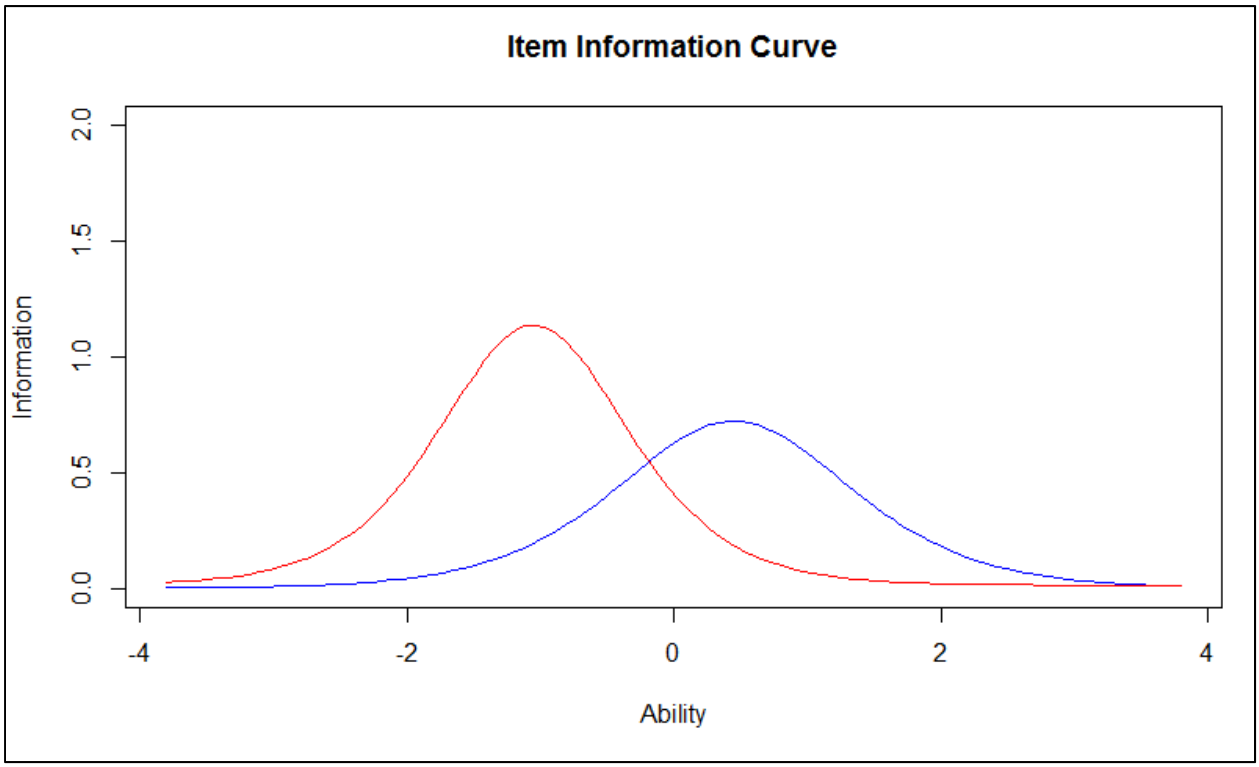

Figure 9: Dutiful versus Submissivness