

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/221836960>

An Item Response Theory Analysis of the Narcissistic Personality Inventory

Article in *Journal of Personality Assessment* · March 2012

DOI: 10.1080/00223891.2011.645934 · Source: PubMed

CITATIONS

49

READS

1,642

3 authors, including:



Robert A Ackerman

The University of Texas at Dallas

90 PUBLICATIONS 2,953 CITATIONS

[SEE PROFILE](#)



Richard Robins

University of California, Davis

326 PUBLICATIONS 45,283 CITATIONS

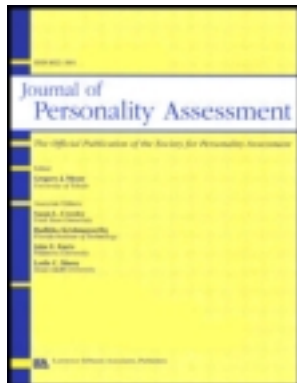
[SEE PROFILE](#)

This article was downloaded by: [The University of Texas at Dallas], [Robert Ackerman]

On: 16 February 2012, At: 08:58

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of Personality Assessment

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/hjpa20>

An Item Response Theory Analysis of the Narcissistic Personality Inventory

Robert A. Ackerman^a, M. Brent Donnellan^a & Richard W. Robins^b

^a Department of Psychology, Michigan State University

^b Department of Psychology, University of California, Davis

Available online: 16 Feb 2012

To cite this article: Robert A. Ackerman, M. Brent Donnellan & Richard W. Robins (2012): An Item Response Theory Analysis of the Narcissistic Personality Inventory, *Journal of Personality Assessment*, 94:2, 141-155

To link to this article: <http://dx.doi.org/10.1080/00223891.2011.645934>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

An Item Response Theory Analysis of the Narcissistic Personality Inventory

ROBERT A. ACKERMAN,¹ M. BRENT DONNELLAN,¹ AND RICHARD W. ROBINS²

¹Department of Psychology, Michigan State University

²Department of Psychology, University of California, Davis

This research uses item response theory methods to evaluate the Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988). Analyses using the 2-parameter logistic model were conducted on the total score and the Corry, Merritt, Mrug, and Pamp (2008) and Ackerman et al. (2011) subscales for the NPI. In addition to offering precise information about the psychometric properties of the NPI item pool, these analyses generated insights that can be used to develop new measures of the personality constructs embedded within this frequently used inventory.

The conceptualization and measurement of narcissistic personality attributes is the subject of ongoing discussion and disagreement (e.g., Ackerman et al., 2011; Brown, Budzek, & Tamborski, 2009; Cain, Pincus, & Ansell, 2008; Miller & Campbell, 2008; Pincus & Lukowitsky, 2010; Rosenthal & Hooley, 2010). One major issue concerns the role that the Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988) should play in future research on narcissism (e.g., Brown et al., 2009; Pincus & Lukowitsky, 2010; Rosenthal & Hooley, 2010). The NPI is the most widely used measure of narcissism in social and personality psychology and it is also increasingly used by some clinical psychologists (e.g., Miller & Campbell, 2008). Nonetheless, concerns have arisen as to whether it: (a) operationalizes narcissism in a way that is consistent with how the disorder is understood in clinical psychology (e.g., Pincus et al., 2009; Rosenthal & Hooley, 2010); (b) successfully captures the core attributes of personality associated with narcissism (Brown et al., 2009); and (c) has an interpretable factor structure and reliable subscales (e.g., Ackerman et al., 2011; Corry, Merritt, Mrug, & Pamp, 2008; del Rosario & White, 2005). As it stands, there is controversy surrounding each of these issues.

These unresolved debates have led some researchers to propose that the field abandon the NPI in favor of alternative measures of narcissistic attributes (e.g., Brown et al., 2009). Although we sympathize with this view, we believe that the field will be reluctant to discard the NPI given its apparent popularity. Recent estimates indicate that over 75% of researchers within social and personality psychology use the instrument to assess narcissism (see Cain et al., 2008). Such estimates suggest that the NPI will continue to play a role in narcissism research for the foreseeable future. It is therefore important that researchers develop a better understanding of the NPI to be able to draw appropriate conclusions from the existing literature and new studies in the future (Ackerman et al., 2011). Moreover, although the NPI might include content that some argue is pe-

ripheral to narcissism (e.g., Rosenthal & Hooley, 2010), we believe that it nevertheless contains some content that reflects relatively uncontroversial manifestations of narcissism (e.g., a sense of entitlement and a willingness to exploit others; see Ackerman et al., 2011). Thus, increased attention toward the psychometric properties of the NPI subscales can be used to inform future scale development efforts. Accordingly, the goal of this research is to use item response theory (IRT) analyses to better understand the properties of this well-used inventory.

BACKGROUND ON THE NARCISSISTIC PERSONALITY INVENTORY

Raskin and Hall (1979) originally developed the NPI to assess the anticipated diagnostic criteria for narcissistic personality disorder (NPD) in the upcoming *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed. [DSM-III]; American Psychiatric Association, 1980). Raskin and his colleagues (e.g., Raskin & Hall, 1981) later winnowed their original item pool to a 40-item version that is in widespread use today (Raskin & Terry, 1988). The response format for the NPI is typically forced-choice, such that participants select one of two options—a narcissistic response or a nonnarcissistic response (e.g., “I like to look at myself in the mirror” vs. “I am not particularly interested in looking at myself in the mirror”). Therefore, the NPI items can be viewed (albeit quite loosely) as a test of narcissism in which items are either answered correctly (i.e., narcissistically) or incorrectly (i.e., not narcissistically).

Several researchers have attempted to identify the dimensional structure of the NPI using exploratory techniques such as principal components analysis (PCA) and exploratory factor analysis (EFA). This body of work has yielded somewhat different structural solutions for the NPI. Emmons (1984, 1987), for example, identified a four-dimension solution: Leadership/Authority, Superiority/Arrogance, Self-Absorption/Self-Admiration, and Exploiteness/Entitlement. Raskin and Terry (1988), in contrast, identified a well-known solution for the NPI consisting of seven scales: Authority, Self-Sufficiency, Superiority, Vanity, Exhibitionism, Entitlement, and Exploiteness.

Psychometric investigations of the NPI in recent years (i.e., Ackerman et al., 2011; Corry et al., 2008) have produced new insights regarding the structure of the NPI by following

Received October 28, 2010; Revised June 2, 2011.

Robert A. Ackerman is now at the University of Texas at Dallas.

Address correspondence to Robert A. Ackerman, Department of Psychological Sciences, The University of Texas at Dallas, 4.126 Green Hall, Richardson, TX 75080-3021; Email: raa110030@utdallas.edu

contemporary recommendations regarding the best practices for conducting factor analyses (e.g., performing analyses on the tetrachoric correlation matrix when the items are dichotomous, using the scree plot and interpretability of the factors as guides in the selection of a factor solution). Corry et al. (2008) uncovered two dimensions behind the NPI item pool: Leadership/Authority and Exhibitionism/Entitlement. These dimensions are useful as they parse the seemingly adaptive content on the NPI from the more explicitly maladaptive content (Ackerman et al., 2011). In generating this solution, Corry et al. placed a premium on identifying dimensions that would yield scales with relatively high levels of internal consistency.

Although internal consistency is important, Ackerman et al. (2011) believed that this emphasis might have prevented Corry et al. (2008) from further isolating important dimensions of personality embedded within the NPI. Ackerman et al. subsequently identified and replicated a three-factor solution consisting of Leadership/Authority, Grandiose Exhibitionism, and Entitlement/Exploitativeness. Whereas their Leadership/Authority factor was very similar to the factor of the same name by Corry et al., the Grandiose Exhibitionism and Entitlement/Exploitativeness factors appeared to split Corry et al.'s factor of Exhibitionism/Entitlement into two parts—one part capturing vanity and exhibitionism and another part capturing entitled beliefs and exploitative behaviors. Ackerman et al. argued that such a distinction was consistent with the theoretical claims made by Brown et al. (2009) regarding the importance of separating grandiosity and entitlement.

One of the major concerns identified in the Ackerman et al. (2011) analyses was that the NPI item pool was somewhat sparse regarding the assessment of many core features of narcissism, such as entitlement and exploitativeness. This relatively small item pool often leads to concerns about the internal consistency of many of the proposed NPI subscales (Corry et al., 2008; del Rosario & White, 2005). However, Ackerman et al. found that their NPI-based Entitlement/Exploitativeness scale repeatedly predicted theoretically relevant constructs despite its relatively low alpha coefficient. This finding is consistent with the emerging recognition that relatively low internal consistency as assessed by alpha is often unrelated to criterion-related validity (McCrae, Kurtz, Yamagata, & Terracciano, 2011). In light of such arguments about the potentially limited utility of this common index of psychometric quality, there is value in evaluating the NPI from the perspective of IRT.

ITEM RESPONSE THEORY METHODS

Existing psychometric work on the NPI has almost exclusively adopted a classical test theory (CTT) approach (e.g., Ackerman et al., 2011; Corry et al., 2008; Emmons, 1984, 1987; Raskin & Terry, 1988). A key assumption of CTT is that measurement error is distributed evenly across all levels of the underlying latent trait. Put differently, this approach assumes that psychological instruments are equally precise at all levels of the latent attribute and therefore equally discriminating at all levels of the underlying trait. This assumption might not hold for many personality measures (see, e.g., Fraley, Waller, & Brennan, 2000; Gray-Little, Williams, & Hancock, 1997). For example, Gray-Little et al. (1997) showed that the Rosenberg (1965) Self-Esteem scale was not particularly effective in discriminating individuals with average levels of self-esteem

from individuals with high levels of self-esteem. Rather, this instrument was much better at distinguishing those with low self-esteem from those with average and above average levels of self-esteem. Consequently, researchers using the scale should place greater confidence in the conclusion that a low score is different from an average score but place less confidence in the conclusion that an average score is much different from a high score. These kinds of findings raise questions regarding the precision of summary scores and their corresponding interpretation in applied research (Morizot, Ainsworth, & Reise, 2007).

IRT methods can be used to provide a more sophisticated perspective on psychological assessments (Embretson & Reise, 2000). In most CTT approaches, "true scores" depend exclusively on the items used to assess the construct. In contrast, IRT methods make a distinction between properties of the items and levels of the latent construct. These techniques represent a family of measurement models in which individuals' responses to a particular item are posited to be a function of the properties of the item as well as individuals' latent standing on a trait. This article makes use of unidimensional IRT models (i.e., models that stipulate that only one common trait underlies responses to a set of items) to investigate the properties of the NPI. In particular, we apply the 2-parameter logistic model (2PL) to the NPI total score and the Corry et al. (2008) and Ackerman et al. (2011) NPI subscales. The units of measurement for the items and latent trait continuum will be logits (i.e., the natural logarithm of the odds of selecting the narcissistic response as opposed to the nonnarcissistic response; Embretson & Reise, 2000). For heuristic purposes, the distribution of logits across the latent trait continuum can be interpreted in a similar way to z scores. We decided to use the 2PL model given its frequent application to other binary response personality inventories like the Multidimensional Personality Questionnaire (MPQ; Johnson, Spinath, Krueger, Angleitner, & Riemann, 2008; Reise & Waller, 1990).

The 2PL models an individual's response to a binary item as a function of the person's latent trait level (θ_p), the item's difficulty level (β_i), and the item's discrimination level (α_i). This item response function is written mathematically as follows (Embretson & Reise, 2000, p. 70):

$$P(X_{ip} = 1 | \theta_p, \beta_i, \alpha_i) = \frac{\exp[\alpha_i(\theta_p - \beta_i)]}{1 + \exp[\alpha_i(\theta_p - \beta_i)]}$$

Two characteristics of the items are being modeled: difficulty and discrimination. Item difficulty is defined as the standing on the latent trait necessary for a person to have a .50 probability of selecting the narcissistic option. Item discrimination can be interpreted as the degree to which an item is more closely related to the latent trait and is therefore analogous to factor loadings in exploratory and confirmatory applications of the common factor model. These discrimination parameters capture the familiar observation that some items are more informative about the latent trait in question than others (i.e., some items have higher factor loadings than other items).

Once item difficulty and discrimination parameters are obtained, item information curves (IICs) can be generated for each item. The IIC formula is (Embretson & Reise, 2000, p. 184)

$$I(\theta) = \alpha_i^2 P_i(\theta)(1 - P_i(\theta))$$

where $I(\theta)$ is the amount of information for a particular level of theta, α_i is the item discrimination parameter, and $P_i(\theta)$ is the item response function for a particular item. These curves provide a visual illustration of how much information each item provides at different levels of the latent trait. In a typical IIC, the x -axis represents the continuum of the latent trait (this normally ranges from -3 to 3 logits, where 0 is the average standing on the latent variable) and the y -axis represents the amount of information that each item provides. Thus, if an item provided high levels of information at the average level of the trait, this would be reflected in a peaked curve around the 0 point on the x -axis. This would in turn suggest that the item performs best at differentiating those individuals who possess average levels of the trait from those who possess below average and above average levels. In this article, the metric of the latent trait is identified via person-anchoring, which means that 0 is defined as the average level of theta in the population under investigation. Likewise, 1 is defined as the standard deviation of theta. As discussed later, this centering has implications for issues of generalization.

Given the additive nature of IICs, they can be summed up to create a test information function (TIF) that conveys the total amount of information that the scale provides across different levels of the latent trait. Such information is useful because it helps to pinpoint areas where the scale is most precise. The TIF is directly related to the standard error of measurement (i.e., $SE(\theta) = 1/\sqrt{TIF(\theta)}$; see Embretson & Reise, 2000, p. 185), such that higher levels of information at different parts of the latent trait continuum correspond to lower standard errors of measurement. In other words, higher levels of test information indicate higher levels of measurement precision, whereas lower levels indicate lower levels of measurement precision.

One important assumption underlying 2PL models is that the scales being analyzed are unidimensional, and this turns out to be a critical issue for evaluating the NPI. Pure and unqualified unidimensionality is difficult to achieve in practice for many personality measures (see Slocum-Gori, Zumbo, Michalos, & Diener, 2009). Fortunately, measures can be considered essentially unidimensional if there is a strong dominant first factor and relatively minor additional factors (Slocum-Gori et al., 2009). The relative dominance of the first factor is often evaluated by conducting an EFA of the correlation matrix and then comparing the size of the first eigenvalue to that of the second eigenvalue. When measures have subscales, other methods for assessing essential unidimensionality can be especially useful. Reise, Moore, and Haviland (2010), for instance, have advocated the use of bifactor models to explore the issue of unidimensionality for multifaceted measures that are nevertheless believed to assess a general construct.

In a restricted bifactor model, each item is specified to load on one general factor and one subscale factor (what Reise et al. [2010] referred to as a "group factor"). Importantly, the general factor and the subscale factors are specified to be orthogonal. This model allows researchers to evaluate the extent to which a common factor is responsible for the covariation among the items as well as how much covariation among particular groups of items is explained by a specific independent subscale factor. Reise et al. (2010) argued that this information is helpful for researchers who are interested in evaluating the suitability of scales for unidimensional IRT analyses. To be sure, there is no reason to believe that a composite of the 40 NPI items assesses

just one trait (e.g., Ackerman et al., 2011; Corry et al., 2008; Bradlee & Emmons, 1992; see especially Raskin & Terry, 1988, p. 892). However, it is possible that a bifactor model might show that the NPI total score is suitable for unidimensional IRT analyses if subscales capture little meaningful information over and above a general dimension of narcissism. In the analyses reported here, we sought to evaluate such a possibility using a bifactor model. We then proceeded to perform IRT analyses on the NPI total score and the Corry et al. (2008) and Ackerman et al. (2011) subscales, as they satisfied the practical requirement of conducting analyses on more or less unidimensional scales.

As an extension of the unidimensionality assumption, IRT methods also assume that the scales being analyzed satisfy the condition of local independence. This means that item responses should be independent of one another conditioning on the common factor (i.e., item responses should be uncorrelated once the common latent variable is taken into account). Researchers using confirmatory factor analytic (CFA) approaches account for a lack of local independence by including correlations between the residuals for two indicators that assess the same latent variable (e.g., correlated residuals for two items measuring the same construct). When this assumption of local independence is violated, as could be the case when items share similar wording, item discrimination parameters might become biased along with other IRT parameters (Embretson & Reise, 2000).¹ The analyses reported here show results for the complete scales. When applicable, however, we note instances in which we believe the assumption of local independence was seriously violated.

PRESENT STUDY

This study uses the 2PL IRT model to better understand the psychometric properties of the NPI in college students, the primary population studied with the NPI. Preliminary analyses are performed to evaluate the unidimensionality of the NPI total score and the Corry et al. (2008) and Ackerman et al. (2011) subscales. Afterward, item difficulty and discrimination parameters are estimated for the NPI total score and the Corry et al. and Ackerman et al. subscales. Item fit will then be evaluated with Yen's (1981) Q statistic and inspection of the item response functions superimposed on the nonparametric response functions for each item. If item fit is deemed satisfactory, the IICs and a TIF are then generated for each scale.²

¹Violating the assumption of local independence can have several consequences. First, local nonindependence might change the meaning of the latent trait to reflect that of the locally dependent items (e.g., instead of measuring the broader trait of Extraversion, a scale might come to be defined by the narrower concept of talkativeness; see Chen & Thissen, 1997; Morizot et al., 2007). Second, local nonindependence might artificially inflate the magnitude of the item discrimination statistics associated with the locally dependent items and attenuate the discrimination statistics of the other items, thereby producing a misleading picture of the items most closely related to the purported latent trait (Chen & Thissen, 1997). Last, the TIFs connected with the subscales including locally dependent items might also be misleading given that the IICs being summed no longer reflect independent information about the latent trait and therefore do not truly reflect increased precision of measurement (Morizot et al., 2007).

²In an additional set of analyses (available on request), we evaluated the extent to which each of the NPI scales satisfied the condition of local independence. This was evaluated by fitting a one-factor CFA-based model to the scale and examining the expected parameter change values associated with the modification indexes as computed in *Mplus*, version 6.0. Close attention was

TABLE 1.—Descriptive statistics for the Narcissistic Personality Inventory (NPI) total score, the Corry et al. (2008) subscales, and the Ackerman et al. (2011) subscales.

	<i>M</i>	<i>SD</i>	Percentile			Skewness	Kurtosis	α	\bar{r}_{ij}	<i>r</i> with NPI Total Score
			25th	50th	75th					
NPI total score	.38	.17	.25	.38	.50	0.30	−0.34	.84	.12	
Corry et al. subscales										
Leadership/Authority	.49	.28	.22	.44	.67	0.10	1.00	.77	.26	.79
Exhibitionism/Entitlement	.30	.21	.14	.29	.43	0.53	−0.38	.71	.15	.79
Ackerman et al. subscales										
Leadership/Authority	.45	.27	.27	.45	.64	0.16	−0.91	.77	.24	.84
Grandiose Exhibitionism	.35	.25	.10	.30	.50	0.46	−0.58	.74	.22	.73
Entitlement/Exploitativeness	.27	.28	.00	.25	.50	0.84	−0.12	.49	.20	.45

Note. α = Cronbach's alpha coefficient; \bar{r}_{ij} = average interitem correlation. The correlations of each scale with each of the others in sequentially listed order from the NPI total score to the Ackerman Entitlement/Exploitativeness scale are as follows: .79, .79, .84, .73, .45; .43, .94, .39, .26; .45, .89, .55; .42, .24; and .23.

To streamline our discussion of the results, we use the following five categories to characterize the range of item difficulties within a scale: very easy (i.e., $\beta_i < -2.0$ logits), easy (i.e., -2.0 logits $< \beta_i < -1.0$ logits), average difficulty (i.e., -1.0 logits $< \beta_i < 1.0$ logits), hard (i.e., 1.0 logits $< \beta_i < 2.0$ logits), and very hard (i.e., $\beta_i > 2.0$ logits). Likewise, we use the following categories to describe the latent trait continuum: well below average (i.e., $\theta_p < -2.0$ logits), below average (i.e., -2.0 logits $< \theta_p < -1.0$ logits), average level (i.e., -1.0 logits $< \theta_p < 1.0$ logits), above average (i.e., 1.0 logits $< \theta_p < 2.0$ logits), and well above average (i.e., $\theta_p > 2.0$ logits). Moreover, we consider items with discrimination values between 0.8 and 2.5 to be good discriminators of the level of the latent trait in question (see de Ayala, 2009). As a point of comparison, Johnson et al. (2008) recently performed IRT analyses on self-reports of the MPQ using a sample from Minnesota. These analyses showed that some MPQ scales displayed relatively wider measurement bandwidth (e.g., Social Closeness, Control, Social Potency, and Absorption) than others (e.g., Achievement and Aggression). For example, the item difficulties for the Control scale ranged from -1.963 to $.296$, whereas item difficulties for the Aggression scale ranged from $.535$ to $.775$. In addition, item discrimination statistics for the MPQ scales ranged from .270 for an item on the Control scale to 1.912 for an item on the Well-Being scale.

Examination of the IICs is used to identify those items that provide the most information about the respective latent traits in question. These analyses therefore identify item content that is particularly relevant for assessing narcissistic tendencies. Moreover, the TIF for each scale is evaluated to determine which levels of the latent trait are being measured with the most and least precision. We consider a value of 1.0 to indicate a reasonable level of information and a value of 2.0 or more to indicate a good level of information. As Reise and Haviland (2005) remark, it might not always be possible to develop items for the entire range of a personality trait—especially those considered pathological. In such cases, researchers might be satisfied with an instrument that assesses those regions of the latent trait defined

as average and above average. Because Leadership/Authority has previously been found to be a reasonably adaptive dimension of personality (e.g., Ackerman et al., 2011), we expect the scale to assess those regions of the latent trait defined as below average, average, and above average more precisely. In contrast, we expect those subscales assessing more maladaptive traits (e.g., exhibitionism, entitlement, and exploitativeness) to provide more precision for above average levels of the trait.

METHOD

Participants, Measure, and Procedure

Participants included 19,001 college students (36.0% men) from a large university in California who were administered the NPI-40 (Raskin & Terry, 1988) online as part of a prescreening procedure that was mandatory for participation in the subject pool.³ The sample size is an important strength of this study given that IRT methods often require large samples. Based on participants' responses to a question about their race, we found that the sample was primarily European American (38.0%) and Asian American (39.3%). Table 1 provides descriptive statistics for the NPI total score and the Corry et al. (2008) and Ackerman et al. (2011) NPI subscales. Item text and response frequencies for this sample are reported in the Appendix. The Appendix also provides a key to the various NPI subscales derived from previous factor analytic work.

RESULTS

The analyses proceeded in a series of steps. We first evaluated the degree to which the NPI total score could be considered unidimensional using a bifactor model (details of the EFA results using the overall NPI score can be found in Ackerman et al., 2011). We subsequently evaluated the unidimensionality of the Corry et al. (2008) and Ackerman et al. (2011) subscales by examining the ratio of the first to the second eigenvalues derived from the respective within-scale matrix of tetrachoric correlations. After determining whether the scales were suitable for the analyses, the 2PL model was fit to each of the scales using marginal maximum likelihood estimation with the

given to those correlations between item residuals that were equal to or above |.20| (see Morizot et al., 2007, p. 413). If a pair of items was found to possess a residual correlation equal to or above this value, the item with the stronger factor loading was kept in a trimmed scale (items that were dropped are noted in the Appendix). We compared all results on the trimmed scales to the results based on the complete scales that included content violating the assumption of local independence. We note discrepancies where applicable.

³The primary data analyzed in this article originally appeared in other publications (e.g., Ackerman et al., 2011; Trzesniewski, Donnellan, & Robins, 2008); however, this is the first time that IRT-based methods have been used on this data set.

latent variable modeling package within R (see Rizopoulos, 2006).

Item fit was evaluated using Yen's (1981) Q statistic, which rank orders participants according to their estimated level of theta, divides them into 10 groups, and then compares the average proportion of participants endorsing the narcissistic option in each of the 10 groups to that proportion that is predicted by the item response function (see Reise, 1990). Yen's Q statistic is distributed as a χ^2 with the degrees of freedom being equal to the number of groups made minus the number of parameters being estimated (Reise, 1990). Because χ^2 is sensitive to sample size (see Reise, 1990) and we had thousands of participants in our sample, we chose to limit our attention to those values that were greater than or equal to 100. In such cases, we plotted the predicted item response function alongside a parallel nonparametric item response function that was derived using kernel binomial regression (cf. Ramsay, 2000). We paid particular attention to the degree of discrepancy between the curves for the area of the latent trait continuum between -2 and 2 logits. Items were judged to be sufficiently fitting if the two curves were overlapping or close to one another for most of this region.

Is the NPI Total Score Unidimensional Enough for IRT Analyses?

Reise et al. (2010) provided a tutorial for the evaluation of essential unidimensionality using bifactor models. A unidimensional model of the scale is first estimated with CFA to obtain estimates of the item loadings when subscale traits are not controlled for in the analysis. A restricted bifactor model (i.e., a solution in which each item loads on a general factor and one and only one subscale factor) is then specified to the items. The logic here is that if the loadings on the general factor in the bifactor solution deviate markedly from the loadings for the unidimensional solution, multidimensionality is clearly distorting the parameter estimates.

Because we were interested in the total score for the NPI with 40 items, we used the Raskin and Terry (1988) solution to specify the subscale factors for the bifactor model analysis.⁴ The bifactor model was specified using CFA within *Mplus* 6.0; WLSMV (weighted least squares with mean and variance adjustment) estimation was used to account for the dichotomous nature of the items. In this model, each of the 40 NPI items was specified to load on a general factor and one additional orthogonal subscale factor following the Raskin and Terry (1988) solution. The variance on each of the latent factors was fixed to 1.

Table 2 presents results from these analyses. Following Reise et al. (2010), loadings on the general factor that are greater than or equal to .50 are considered strong and thus highlighted in bold; likewise, loadings on any of the subscale factors that exceeded the respective loading for the general factor in strength are highlighted in bold. As can be seen, 14 loadings were observed to load strongly on the general factor. If the criteria were relaxed such that loadings $\geq .40$ were considered strong, this

TABLE 2.—Standardized factor loadings from a unidimensional model solution and a restricted bifactor model solution using the Raskin and Terry (1988) solution.

Item	Unidimensional Model	Bifactor Model							
		Gen	AU	EP	ET	EX	SS	SU	VA
1	.592	.550	.320						
2	.399	.403				.158			
3	.407	.405				.178			
4	.491	.523						.077	
5	.418	.447			.030				
6	.487	.486		.437					
7	.751	.602				.702			
8	.377	.350	.211						
9	.536	.506						.705	
10	.657	.535	.619						
11	.560	.493	.404						
12	.620	.562	.381						
13	.525	.529		.486					
14	.198	.185			.461				
15	.483	.450							.496
16	.315	.316		.256					
17	.190	.188					.184		
18	.342	.363			.045				
19	.508	.458							.887
20	.508	.504				.236			
21	.306	.305					.495		
22	.085	.067					.344		
23	.433	.449		.160					
24	.191	.171			.543				
25	.326	.313			.772				
26	.372	.381						.147	
27	.557	.585			.248				
28	.351	.340				.202			
29	.423	.388							.503
30	.781	.653				.636			
31	.349	.357					.297		
32	.643	.583	.400						
33	.680	.571	.570						
34	.414	.426					.324		
35	.404	.400		.418					
36	.661	.609	.379						
37	.356	.374						.074	
38	.448	.440				.245			
39	.434	.460					.192		
40	.552	.527						.565	

Note. Loadings on the general factor that are greater than or equal to .50 are in bold; loadings on any of the subscale factors that exceeded the respective loading for the general factor are in bold. Fit of the unidimensional model: $\chi^2(740, N = 19,001) = 54,126.692, p < .001$; comparative fit index (CFI) = .754; Tucker–Lewis Index (TLI) = .741; root mean square error of approximation (RMSEA) = .062. Fit of the restricted bifactor model: $\chi^2(700, N = 19,001) = 29,366.853, p < .001$; CFI = .868; TLI = .853; RMSEA = .046. Gen = General factor; AU = Authority; EP = Exploitativeness; ET = Entitlement; EX = Exhibitionism; SS = Self-Sufficiency; SU = Superiority; VA = Vanity.

number would be increased to 26. Examination of the subscale factor loadings revealed that few items on the subscales possessed loadings that exceeded the strength of the general factor (notable exceptions included items on the Entitlement and Vanity subscales). There also appeared to be minimal differences between loadings for the general factor and loadings for the unidimensional model. This suggests that the NPI total score could be considered unidimensional enough for IRT analyses. However, an inspection of the loadings for the general factor in the bifactor model suggests that this latent trait is primarily defined by content related to leadership and attention-seeking behavior.

⁴All other solutions for the NPI do not make use of the full 40 items. Because a restricted bifactor model requires that each item loads on a general factor and one and only one subscale factor (see Reise et al., 2010), only the Raskin and Terry (1988) solution seemed appropriate for this analysis using all 40 items.

Are the Corry et al. and Ackerman et al. Subscales Meaningful?

The previous analyses based on the Reise et al. (2010) bifactor approach suggested that the 40-item NPI total score is unidimensional enough for IRT analyses. They also showed that the majority of the Raskin and Terry (1988) subscales do not account for substantial covariation among the items once the general factor of narcissism is controlled for in the analysis. Reise et al. made the argument that subscales are only meaningful to the extent that they explain covariation in the items over and above some general factor. We therefore elected to perform similar restricted bifactor analyses for the subscales that make up the Corry et al. (2008) and Ackerman et al. (2011) solutions. To be able to justify the creation of subscales, we reasoned that these solutions must contain subscales in which the majority of the items load more strongly on the subtrait factor than the general factor.

When we applied the restricted bifactor model to the 23 items that make up the Corry et al. (2008) solution, we found that 4 items possessed strong loadings (i.e., $\geq .50$) on the general factor (results available on request). In addition, eight out of nine items on the Leadership/Authority subtrait factor possessed loadings that exceeded the strengths of those found for the general factor. The Exhibitionism/Entitlement subtrait factor also had three loadings (out of 14) that exceeded the strength of those found on the general factor. These findings suggest that the Leadership/Authority items are picking up on a particular aspect of the NPI better than the general trait. Although we did not find strong evidence to suggest that this was the case for Exhibitionism/Entitlement, we believe that it is important to isolate the subtrait of Leadership/Authority from the other elements of narcissism given current debates regarding whether Leadership/Authority should be considered a narcissistic trait.

Applying the restricted bifactor model to the 25 items that make up the Ackerman et al. (2011) solution (results available on request), we found that 5 items possessed strong loadings (i.e., $\geq .50$) on the general factor. In addition, 8 out of 11 items on the Leadership/Authority subtrait factor had loadings that exceeded the strength of those found on the general factor. Moreover, 3 out of 10 items on the Grandiose Exhibitionism subtrait factor and three out of four items on the Entitlement/Exploitativeness subtrait factor had loadings that exceeded the strength of those found on the general factor. Given that the majority of the items for the Leadership/Authority and Entitlement/Exploitativeness subscales loaded more strongly on the specific group factors than the general trait, we believe it is empirically justifiable to create all three subscales from the Ackerman et al. solution for the NPI.

Taken together, the results of these bifactor analyses suggest that the NPI subscales from the Corry et al. (2008) and Ackerman et al. (2011) solutions are potentially meaningful. Most of the subscales (three out of the five) contained a majority of items on which the loadings for the subtrait factor exceeded those for the general factor in strength. These findings lend support to the utility of conducting IRT analyses on these particular subscales.

Application of the 2-Parameter Logistic Model to the NPI Total Score

After running the 2PL model on the NPI total score, we found that 11 items possessed χ^2 values for item fit that were greater than or equal to 100 (Items 1, 7, 10, 11, 12, 26, 30, 32, 33, 36,

TABLE 3.—2-parameter logistic model (2PL) item parameter estimates for the Narcissistic Personality Inventory total score.

Item	Item Fit $\chi^2(8)$	2PL Model			
		β	σ_β	α	σ_α
1	124.71	−0.84	0.02	1.26	0.03
2	59.73	2.26	0.06	0.78	0.03
3	52.43	1.69	0.05	0.79	0.02
4	65.66	1.84	0.04	1.02	0.03
5	48.98	0.71	0.03	0.79	0.02
6	52.01	0.97	0.03	0.95	0.02
7	135.42	0.63	0.02	1.60	0.03
8	98.02	−2.06	0.06	0.73	0.02
9	97.25	0.25	0.02	1.00	0.02
10	184.82	−0.16	0.01	1.47	0.03
11	99.65	0.06	0.02	1.17	0.02
12	102.79	0.06	0.01	1.35	0.03
13	84.22	0.90	0.02	1.05	0.02
14	12.65	2.64	0.15	0.32	0.02
15	62.46	1.52	0.04	0.89	0.02
16	35.48	0.15	0.03	0.55	0.02
17	27.49	0.92	0.06	0.33	0.02
18	35.95	1.27	0.04	0.61	0.02
19	89.80	0.75	0.02	0.89	0.02
20	55.40	1.42	0.03	1.04	0.03
21	33.93	1.88	0.07	0.55	0.02
22	12.50	1.25	0.17	0.15	0.02
23	58.53	1.89	0.05	0.89	0.03
24	17.65	3.38	0.20	0.31	0.02
25	26.30	2.40	0.09	0.58	0.02
26	117.52	−0.68	0.03	0.65	0.02
27	97.86	0.43	0.02	1.15	0.02
28	44.00	0.97	0.04	0.62	0.02
29	69.24	−0.35	0.02	0.70	0.02
30	150.73	0.34	0.01	1.73	0.03
31	50.84	−0.18	0.03	0.61	0.02
32	111.31	0.82	0.02	1.53	0.03
33	144.94	0.11	0.01	1.57	0.03
34	83.72	−0.22	0.02	0.75	0.02
35	57.38	1.35	0.04	0.76	0.02
36	161.91	1.01	0.02	1.64	0.03
37	54.14	0.56	0.03	0.63	0.02
38	59.94	1.69	0.04	0.87	0.02
39	54.31	1.46	0.04	0.84	0.02
40	109.10	0.36	0.02	1.06	0.02

Note. β = difficulty parameter; σ_β = standard error for difficulty parameter; α = discrimination parameter; σ_α = standard error for discrimination parameter.

and 40). However, visual inspection of the nonparametric item response functions against the 2PL item response functions for these items revealed that they generally fit the model (these and all other figures assessing item fit are available on request from the first author). Discrimination and difficulty estimates along with item-fit statistics are reported in Table 3. The majority of the items were of average (57.5%) or hard (30.0%) difficulty; only one item was very easy (Item 8) and four items were very hard (Items 2, 14, 24, and 25).

Most of the items did a good job of discriminating individuals at average and above average levels of the trait. As Table 3 shows, the most discriminating items were those that reflected aspects of leadership ability and attention-seeking behavior. It is notable that a few items connected to the Raskin and Terry (1988) Entitlement and Self-Sufficiency subscales were very poor discriminators of the underlying trait (i.e., Items 14, 17, 22, and 24). Figure 1 shows the TIF for the NPI total score. The scale provided good precision for much of the latent trait

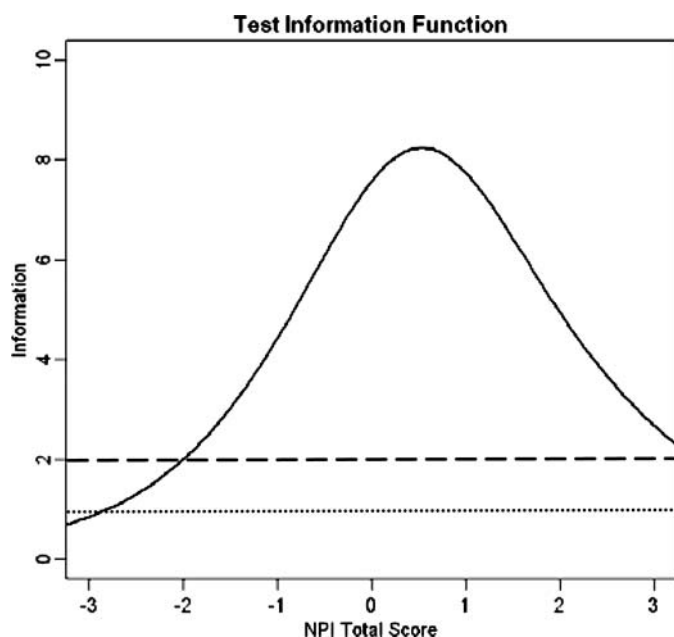


FIGURE 1.—Test information function for the Narcissistic Personality Inventory total scale. *Note.* The dashed line in the figure provides a reference point for which the information is equal to 2.0. The line with square dots beneath the dashed line provides a reference point for which the information is equal to 1.0.

continuum (i.e., between well below average, or $\theta_p \approx -2.0$, and well above average, or $\theta_p \approx 3.0$). Moreover, the greatest precision appeared to occur for those individuals possessing average and somewhat above average levels of the trait.

Additional analyses revealed that there were several items pairs that produced violations of local independence. IRT analyses performed on the trimmed NPI scale revealed very similar estimates for the difficulty and discrimination parameters. The TIF for the trimmed scale was appreciably smaller (as would be expected given the removal of redundant item content) and showed that the trimmed NPI scale provided good precision for people possessing between somewhat below average (i.e., $\theta_p \approx -1.3$) and well above average (i.e., $\theta_p \approx 2.8$) levels of the trait. Violating the assumption of local independence was therefore considered to have a minimal impact on the overall pattern of results.

Examination of the item content shows that many of the easier items (i.e., $.10 > \beta_i$) belong to the Raskin and Terry (1988) subscales of Authority and Self-Sufficiency. These items reflect normatively positive attitudes toward the self. The items closer to the middle of the continuum (i.e., $.10 < \beta_i < 1.0$) primarily belong to the Exploitativeness, Exhibitionism, and Superiority subscales. These items reflect somewhat grandiose self-statements that entail a sense of specialness and facile control over other people. The hardest items (i.e., $\beta_i > 1.0$) belong to the Exhibitionism and Entitlement scales. These items represent even more immodest and grandiose self-statements.

Application of the 2-Parameter Logistic Model to the Corry et al. NPI Subscales

Before fitting the models, we used *Mplus* to conduct EFAs on the tetrachoric correlation matrices for the item sets making up the Corry et al. (2008) Leadership/Authority and Exhi-

TABLE 4.—2-parameter logistic model (2PL) item parameter estimates for the Leadership/Authority and Exhibitionism/Entitlement subscales from the Corry et al. solution for the Narcissistic Personality Inventory.

	Item	Item Fit $\chi^2(8)$	2PL Model			
			β	σ_β	α	σ_α
Leadership/Authority	1	252.93	-0.82	0.02	1.31	0.03
	8	143.93	-1.89	0.05	0.81	0.03
	10	684.73	-0.14	0.01	2.17	0.05
	11	285.72	0.05	0.01	1.36	0.03
	12	595.13	0.05	0.01	1.57	0.03
	27	328.92	0.43	0.02	1.15	0.03
	32	367.19	0.76	0.02	1.80	0.04
	33	552.71	0.09	0.01	2.36	0.05
Exhibitionism/Entitlement	36	335.03	0.98	0.02	1.75	0.04
	3	126.14	1.77	0.05	0.75	0.02
	4	146.64	1.81	0.04	1.05	0.03
	7	664.48	0.55	0.01	2.30	0.06
	14	44.59	2.37	0.13	0.36	0.02
	15	466.94	1.13	0.02	1.39	0.04
	19	355.54	0.64	0.02	1.15	0.03
	20	266.99	1.18	0.02	1.39	0.03
	24	20.87	3.00	0.17	0.35	0.02
	25	59.76	2.39	0.09	0.58	0.02
	28	107.90	0.76	0.03	0.85	0.02
	29	326.61	-0.25	0.02	1.08	0.03
	30	654.52	0.31	0.01	2.44	0.07
	38	228.77	1.32	0.03	1.25	0.03
	39	70.36	2.16	0.09	0.53	0.02

Note. β = difficulty parameter; σ_β = standard error for difficulty parameter; α = discrimination parameter; σ_α = standard error for discrimination parameter.

bitionism/Entitlement subscales. We then examined the ratios of the eigenvalues for the first to second factors to evaluate their essential unidimensionality. For Leadership/Authority, the first two eigenvalues were 4.51 and 0.96, producing a ratio of 4.69 and a first eigenvalue that explained 50.11% of the variance. Moreover, the first two eigenvalues for Exhibitionism/Entitlement were 4.47 and 1.89, thus producing a ratio of 2.36 and a first eigenvalue that explained 31.93% of the variance. These findings suggested to us relatively dominant first factors, so we considered each of the scales to be essentially unidimensional.

After running the 2PL model on the Leadership/Authority subscale, we found that all the items possessed χ^2 values for item fit above 100. Moreover, when we ran the 2PL model on the Exhibitionism/Entitlement subscale, we found that 10 items possessed χ^2 values for item fit above 100 (Items 3, 4, 7, 15, 19, 20, 28, 29, 30, and 38). Nevertheless, visual inspection of the nonparametric item response functions against the 2PL item response functions revealed that all of the items generally fit their corresponding models. Difficulty and discrimination statistics, along with item fit statistics, are reported in Table 4.

Leadership/Authority. The left-hand side of Panel A in Figure 2 shows the IICs for the items on the Leadership/Authority subscale. Items 10 and 33 provided reasonable levels of precision at the average level of the trait. Panel A in Figure 2 further shows that Items 32 and 36 were better at differentiating individuals with slightly higher levels of Leadership/Authority. Item 1 was shown to differentiate individuals with lower levels of Leadership/Authority. The TIF for the Leadership/Authority scale is presented on the left-hand side of Panel B in Figure 2.

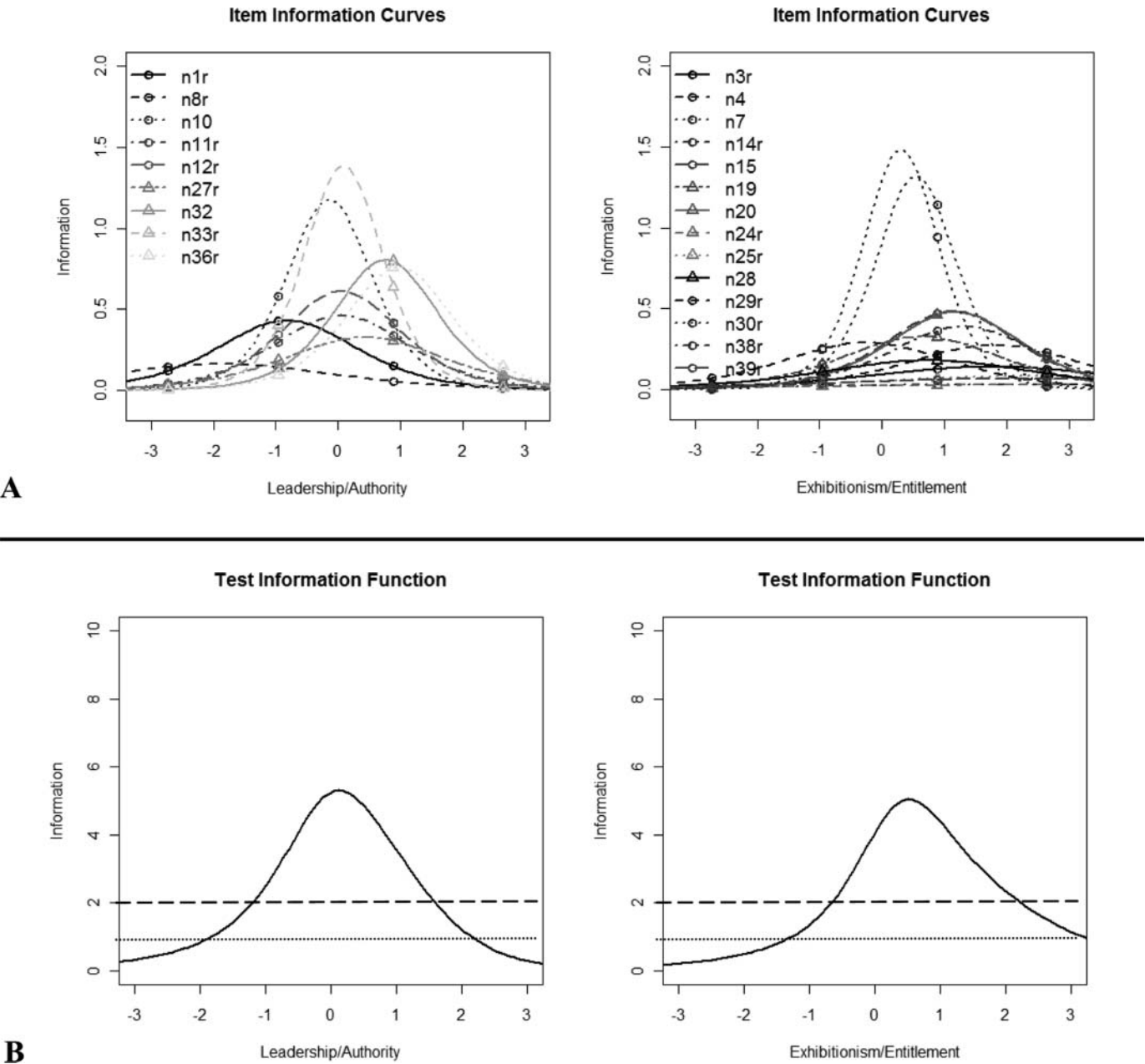


FIGURE 2.—Item information curves and test information functions for Corry et al. (2008) Narcissistic Personality Inventory subscales. *Note.* The dashed line in the figures in Panel B provides a reference point for which the information is equal to 2.0. The line with square dots beneath the dashed line in Panel B provides a reference point for which the information is equal to 1.0.

The scale was seen to provide good levels of measurement precision (i.e., information approximately ≥ 2.0) for individuals possessing between average (i.e., $\theta_p \approx -1$) and above average (i.e., $\theta_p \approx 1.5$) levels of the trait. Analyses did not reveal any violations of local independence.

Exhibitionism/Entitlement. The right-hand side of Panel A in Figure 2 shows the IICs for the Exhibitionism/Entitlement items. Items 7 and 30 provided a great deal of information at the average level of the trait. Moreover, Items 15, 20, and 38 provided some information at the above average level of the trait, whereas Item 4 was seen to provide some information near the

well above average end. Of note, Items 14, 24, and 25 provided minimal levels of information. The right-hand side of Panel B in Figure 2 shows the TIF for the Exhibitionism/Entitlement scale. The scale provided good levels of measurement precision (i.e., information approximately ≥ 2.0) for individuals possessing between average (i.e., $\theta_p \approx -0.5$) and above average (i.e., $\theta_p \approx 2.0$) levels of the trait.

Additional analyses revealed several item pairs that violated the assumption of local independence. IRT analyses on the trimmed scale showed very similar results for the item parameters. The only noteworthy exception was the attenuated discrimination parameter estimate for Item 30 in the reduced

model ($\alpha = 1.94$). The IICs also showed a similar pattern of results. The TIF, on the other hand, showed that the trimmed scale did not provide good precision for any region of the latent trait after removal of these items. Instead, it only provided reasonable levels of precision (i.e., information approximately ≥ 1.0) for individuals possessing between average (i.e., $\theta_p \approx -0.6$) and well above average (i.e., $\theta_p \approx 2.5$) levels of the trait. Therefore, although violating the condition of local independence did not adversely affect the item parameter estimates, it did result in a TIF that was artificially inflated by redundant item content.

Application of the 2-Parameter Logistic Model to the Ackerman et al. NPI Subscales

The unidimensionality of each of the subscales was once again evaluated by examining the ratio of the first to second eigenvalues. The first two eigenvalues for Leadership/Authority were 4.91 and 1.14, producing a ratio of 4.32 and a first eigenvalue that explained 44.64% of the variance. The first two eigenvalues for Grandiose Exhibitionism were 3.84 and 1.11, producing a ratio of 3.45 and a first eigenvalue that explained 38.4% of the variance. Last, the first two eigenvalues for Entitlement/Exploitativeness were 2.01 and 0.79, resulting in a ratio of 2.54 and a first eigenvalue that explained 50.25% of the variance. We believe that these findings provide evidence of dominant first factors, so we considered the scales suitable for the 2PL analyses.

After running the 2PL model on the Leadership/Authority subscale, we found that eight items possessed χ^2 values for item fit above 100 (Items 1, 10, 11, 12, 27, 32, 33, and 36). Moreover, after applying the 2PL model to the Grandiose Exhibitionism and Entitlement/Exploitativeness subscales, we found that all of the items on these scales demonstrated poor fit. Nevertheless, visual inspection of the nonparametric item response functions against the 2PL item response functions for each of these items revealed that all of the items generally fit their corresponding models. Difficulty and discrimination statistics and item fit statistics for each subscale are reported in Table 5.

Leadership/Authority. Panel A in Figure 3 shows the IICs for the Leadership/Authority items. These results parallel those found for the Corry et al. (2008) Leadership/Authority subscale. Panel B in Figure 3 shows the TIF for the scale. The scale was seen to provide good levels of measurement precision (i.e., information approximately ≥ 2.0) for individuals possessing between average (i.e., $\theta_p \approx -1.3$) and above average (i.e., $\theta_p \approx 1.8$) levels of the trait. Additional analyses revealed one item pair that violated the local independence assumption. IRT analyses on the trimmed scale revealed virtually identical item parameter estimates. Moreover, the IICs and TIF for the trimmed scale were essentially unchanged.

Grandiose Exhibitionism. Panel A in Figure 3 shows the IICs for the Grandiose Exhibitionism items. These results parallel those found for the Corry et al. (2008) Exhibitionism/Entitlement subscale. The TIF is presented in Panel B of Figure 3. Overall, the scale provided good precision (i.e., information approximately ≥ 2.0) for individuals possessing between average (i.e., $\theta_p \approx -0.80$) and above average (i.e., $\theta_p \approx 2.0$) levels of the trait.

TABLE 5.—2-parameter logistic model (2PL) item parameter estimates for the Leadership/Authority, Grandiose Exhibitionism, and Entitlement/Exploitativeness subscales from the Ackerman et al. (2011) solution for the Narcissistic Personality Inventory.

	Item	Item Fit $\chi^2(8)$	2PL Model			
			β	σ_β	α	σ_α
Leadership/Authority	1	188.78	-0.81	0.02	1.33	0.03
	5	97.57	0.71	0.03	0.79	0.02
	10	486.89	-0.13	0.01	2.30	0.05
	11	186.77	0.05	0.01	1.36	0.03
	12	252.96	0.05	0.01	1.48	0.03
	27	195.29	0.43	0.02	1.12	0.03
	32	354.90	0.76	0.02	1.80	0.04
	33	381.05	0.09	0.01	2.22	0.05
	34	88.90	-0.23	0.02	0.76	0.02
	36	275.55	0.96	0.02	1.82	0.04
Grandiose Exhibitionism	40	88.42	0.43	0.02	0.83	0.02
	4	172.15	1.74	0.04	1.11	0.03
	7	749.05	0.54	0.01	2.34	0.07
	15	551.82	1.12	0.02	1.42	0.04
	19	433.76	0.60	0.02	1.26	0.03
	20	254.63	1.21	0.02	1.33	0.03
	26	146.30	-0.50	0.02	0.94	0.02
	28	1039.64	0.75	0.03	0.86	0.02
	29	429.62	-0.24	0.02	1.18	0.03
	30	698.46	0.31	0.01	2.38	0.07
Entitlement/Exploitativeness	38	316.77	1.44	0.03	1.09	0.03
	13	4249.88	1.13	0.04	0.78	0.03
	14	7313.86	1.01	0.03	0.99	0.03
	24	4204.04	1.04	0.03	1.29	0.05
	25	3298.97	1.01	0.02	2.20	0.11

Note. β = difficulty parameter; σ_β = standard error for difficulty parameter; α = discrimination parameter; σ_α = standard error for discrimination parameter.

Additional analyses revealed numerous item pairs that violated the condition of local independence. IRT analyses on the trimmed scale revealed item parameter estimates that were very similar to those for the complete scale. The only notable exception was the attenuated item discrimination estimate found for Item 30 on the trimmed scale ($\alpha = 1.70$). Likewise, the IICs demonstrated a similar pattern of results, except that Item 30 provided reduced information. The overall level of information for the TIF of the trimmed scale was substantially reduced. Indeed, the scale no longer provided good precision (i.e., information approximately ≥ 2.0) for any region of the latent trait. Instead, the scale only provided reasonable precision (i.e., information approximately ≥ 1.0) for individuals possessing between average (i.e., $\theta_p \approx -0.40$) and well above average (i.e., $\theta_p \approx 2.3$) levels of the trait.

Entitlement/Exploitativeness. Panel A in Figure 3 displays the IICs for the Entitlement/Exploitativeness items. As can be seen, all items were providing information at the above average level of the trait. Nevertheless, Item 25 provided the most information and was therefore the best item at differentiating individuals with above average levels of Entitlement/Exploitativeness from those with average levels or well above average levels of the trait. This pattern of information provided by the items is further reflected in the TIF in Panel B of Figure 3. All in all, the scale provided reasonable precision (i.e., information ≥ 1.0) for individuals possessing between average (i.e., $\theta_p \approx 0.0$) and above average (i.e., $\theta_p \approx 2.0$) levels of the trait. Analyses did not reveal any violations of local independence.

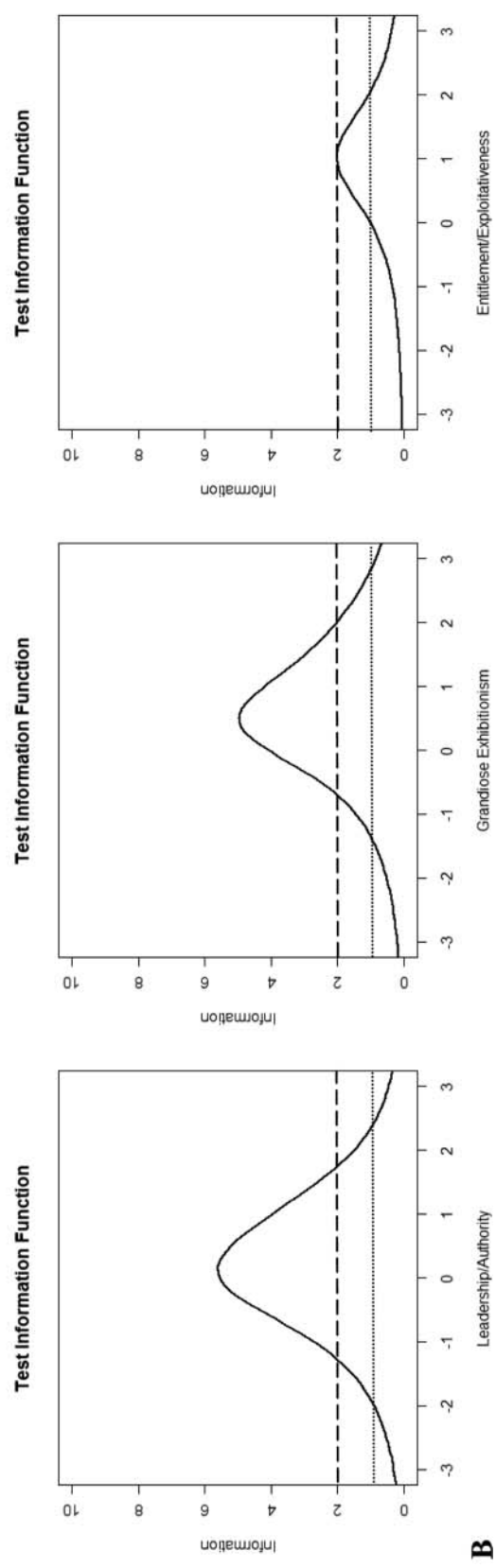
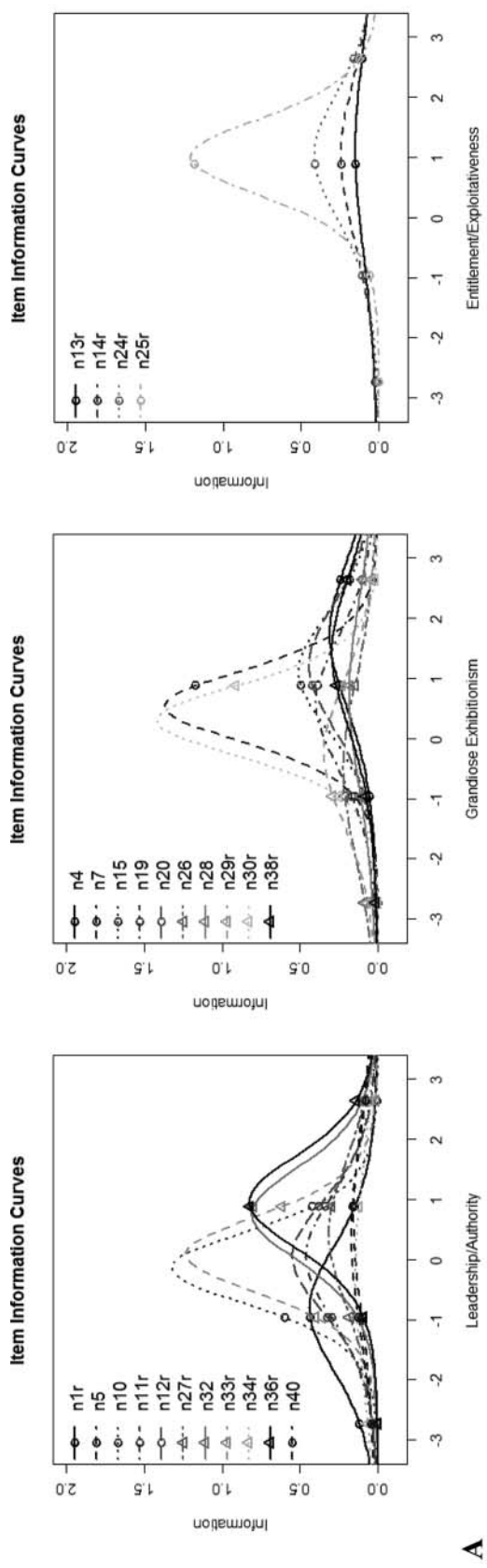


FIGURE 3.—Item information curves and test information functions for Ackerman et al. (2011) scales *Note.* The dashed line in the figures in Panel B provides a reference point for which the information is equal to 2.0. The line with square dots beneath the dashed line in Panel B provides a reference point for which the information is equal to 1.0.

DISCUSSION

The purpose of this study was to use IRT methods to gain insight into the properties of the NPI. We evaluated the NPI total scale as well as the Corry et al. (2008) and Ackerman et al. (2011) subscales. Initial results based on the outcome of a bifactor model-based approach for evaluating unidimensionality suggested that a common dimension underlies responses to the 40 NPI items. However, that latent trait is primarily defined by content related to leadership ability and attention seeking. The common dimension does not capture content encompassed in the Raskin and Terry (1988) subscales of Self-Sufficiency and Entitlement. We have previously argued that any use of the NPI total score is potentially problematic as it conflates adaptive and maladaptive aspects of personality (Ackerman et al., 2011). The findings of this research further this argument by showing that much of what is considered to be interpersonally maladaptive about narcissism (namely entitlement; Campbell, Bonacci, Shelton, Exline, & Bushman, 2004; Konrath, Bushman, & Campbell, 2006) is not strongly related to the underlying dimension captured by the NPI total score. This finding is especially interesting given that the entitlement content of the NPI has the strongest associations with socially toxic outcomes (Ackerman et al., 2011; Konrath et al., 2006) and seems to share the most overlap with pathological measures of narcissism (Ackerman et al., 2011; Maxwell, Donnellan, Hopwood, & Ackerman, 2011).

We suggest that there is much to be gained from using subscales based on some of the existing structural solutions for the NPI rather than using the NPI total score (Ackerman et al., 2011; Barry, Frick, & Killian, 2003). In particular, we think that the Corry et al. (2008) and Ackerman et al. (2011) solutions to the NPI offer a useful partitioning of the item content that is justified on both empirical and theoretical grounds. Results showed that the NPI subscales differed in the area of the latent trait they were assessing most precisely and the amount of precision that they offered overall. Because the results revealed a high degree of overlap between the Corry et al. and Ackerman et al. Leadership/Authority subscales, as well as between the Corry et al. Exhibitionism/Entitlement subscale and the Ackerman et al. Grandiose Exhibitionism subscale, those scales are discussed together. Next we discuss these issues in more detail and describe the implications of our results for use of the NPI and future scale development efforts.

Use of the NPI Total Score and Subscales in Narcissism Research

A considerable amount of research in social and personality psychology makes use of the NPI total score to study narcissism. An important question is whether use of the NPI total score is justified given the potentially multidimensional nature of the scale. The bifactor analysis that we conducted on the NPI-40 suggests that the NPI total score can be considered more or less unidimensional. That is, despite the fact that the instrument is multifaceted, there are indications of a common dimension underlying responses to the NPI items. The natural question then becomes what this common dimension is assessing.

The analyses reported here provided some insight into this question. Examination of how the item content differed depending on its difficulty suggested a continuum of normative to grandiose self-regard. People with lower levels of “narcis-

sism” were more likely to endorse NPI items that reflected self-perceived potential for leadership (e.g., “I have a natural talent for influencing people”) or normative positive feelings toward the self (e.g., “I will be a success”). In contrast, only people with higher levels of “narcissism” were likely to endorse NPI items that reflected strongly held feelings of leadership potential (e.g., “I am a born leader”) or a sentiment of specialness and entitlement (e.g., “I will never be satisfied until I get all that I deserve”).

Consistent with previous researchers’ criticisms (e.g., Rosenthal & Hooley, 2010), we are skeptical that the underlying latent trait that is captured by the NPI total score reflects uncontroversial manifestations of maladaptive narcissism. When evaluating the factor loadings in the bifactor model, items that reflect being a leader or authority figure and being the center of attention had the strongest associations with the general factor. In contrast, many of the items capturing entitlement that have been found to be good discriminators of the latent trait underlying the Ackerman et al. (2011) Entitlement/Exploitativeness subscale possessed the weakest loadings on this general factor.⁵ Consequently, we suspect that this general factor is saturated with content linked to elements of Extraversion from the Five-Factor Model. The issue, as Trull and McCrae (2002) argued, is that “nothing in the *DSM-III-R* definition [of narcissistic personality disorder] suggests high [Extraversion]: Sociability, leadership, and excitement seeking are not mentioned at all” (p. 53). Likewise, Corbitt (2002) claimed that “the Extraversion domain is not theoretically central to narcissism” (p. 297). Thus, there are ongoing disagreements about whether many of the elements of personality assessed by the NPI are manifestations of narcissism as it is commonly understood in clinical psychology.

Given these concerns, we believe it is best to parse the NPI total score into subscales to gain a more refined understanding of narcissism. We suggest that if researchers are primarily interested in assessing narcissistic entitlement and its relations with criterion variables, then the use of the NPI total score is probably counterproductive because items assessing those attributes will be substantially outnumbered by items assessing more adaptive elements of personality. Likewise, if researchers are interested in gaining a more refined understanding of the features of the NPI unrelated to leadership and authority, it makes sense to break up the instrument into its subscales in research. The advantage of this approach is that it affords greater clarity over which attributes assessed by the NPI are contributing to its overall associations with criterion variables.

Psychometric Characteristics of the NPI Subscales

Given our reservations about the NPI total score, an important goal of this research was to evaluate the characteristics of the NPI subscales using IRT methods. We evaluated the overall amount of precision that they were providing, the range of the trait that was being afforded the most precision, and the items

⁵Further evidence for this assertion comes from examination of the zero-order correlations between the theta estimates for the NPI total score and the NPI subscales. The NPI total score trait was most strongly related to the Corry et al. (2008) Leadership/Authority ($r = .85$) and Ackerman et al. (2011) Leadership/Authority ($r = .87$) traits. It also demonstrated strong relations with the Corry et al. Exhibitionism/Entitlement ($r = .77$) and Ackerman et al. Grandiose Exhibitionism ($r = .75$) traits. In contrast, it was much more modestly related to the Ackerman et al. Entitlement/Exploitativeness trait ($r = .34$).

that were most informative regarding the underlying traits. We expected that the subscales assessing more adaptive traits like Leadership/Authority would provide good precision for the below average, average, and above average levels of the trait. In contrast, we expected that the subscales assessing more maladaptive traits would provide good precision for the average and above average levels.

The Leadership/Authority subscales provided the most precision among all the subscales. These subscales provided good precision for the average and above average regions of the latent trait, as well as reasonable precision for the below average level of the trait. When items were sorted according to their difficulty, the item content suggested a latent continuum ranging from self-perceived leadership potential to arrogant certainty about one's leadership ability. People with relatively lower levels of the trait were likely to endorse items that possessed content reflecting traits conducive to leadership (e.g., being assertive and influential) and statements that involved self-perceptions of leadership ability (e.g., seeing oneself as a good leader). People had to possess higher levels of the trait to be more likely to endorse items that reflected more deeply held convictions of leadership ability (e.g., "I am a born leader"). Given the range of precision already provided by the scales, we do not believe that extra items are necessary. However, we recommend dropping Items 5 and 40 from the Ackerman et al. (2011) Leadership/Authority scale because these items provided a small amount of information and their content seemed to better capture sentiments of grandiosity on closer inspection.

The Corry et al. (2008) Exhibitionism/Entitlement and Ackerman et al. (2011) Grandiose Exhibitionism subscales provided a good deal of precision along much of their latent trait continuums; however, they tended to more precisely measure the slightly higher ends of their traits. Unlike Leadership/Authority, the results showed that much of this precision was an artifact due to redundant item content on the scales. The underlying continuums for these traits appear to reflect grandiosity. To be sure, people with average levels of these traits were likely to endorse items that reflected normative levels of vanity and being the center of attention. People had to possess higher levels of these traits to be likely to endorse items that were characterized by boastful showing off, demanding attention, and grandiose feelings toward the self.

The IICs for the Corry et al. (2008) Exhibitionism/Entitlement and Ackerman et al. (2011) Grandiose Exhibitionism subscales showed that none of the items were providing information at the lower ends of the traits. Given the nature of these traits, we believe that this is to be expected. It would be ideal to develop more items that provide information for the average levels of these trait continuums. These items could focus on a normative preference for attention and habits or behaviors that indicate self-absorption but do not necessarily involve the exploitation or denigration of others. As Lessard, Greenberger, Chen, and Farruggia (2011) have argued, it is possible that thinking positively of oneself can be healthy and adaptive (or at the very least unproblematic), whereas it is the psychological phenomenon of holding positive feelings toward the self that occur at the expense of others that is socially toxic and destructive (e.g., tending to denigrate others to feel good about the self). Thus, greater conceptual clarity over this issue could help further guide scale refinement efforts.

Although the above average area of these traits was assessed with a reasonable level of precision, it would be useful to develop even more items that provide information for this area. We believe that the item content should reflect the use of intrapersonal (e.g., beliefs denoting superiority) and interpersonal strategies (e.g., bragging about accomplishments) aimed at propping up a potentially inflated self-image. We recommend dropping Items 14, 24, and 25 from the Corry et al. (2008) Exhibitionism/Entitlement subscale as they provided negligible levels of information. Such a finding is consistent with the idea that these items reflecting entitlement do not belong on a scale that appears to better reflect grandiosity (Brown et al., 2009).

The Ackerman et al. (2011) Entitlement/Exploitativeness scale provided reasonably effective measurement for individuals who possessed somewhat above average levels of the trait. This subscale provided the narrowest bandwidth of precision among the subscales. Individuals endorsing these items are likely to be very high on this dimension. This might explain why the scale seems to "work" from the perspective of criterion-related validity—higher scorers will likely have very high levels of this seemingly maladaptive aspect of personality linked with low self-esteem and higher rates of counterproductive school behaviors (Ackerman et al., 2011; Maxwell et al., 2011). We recommend that future researchers concentrate their efforts on developing items that provide more precise information about this area of the trait as well as the average level of this trait.

Summary and Conclusions

In the end, the only NPI subscales that provided a good range of precision across the largest spectrum of the underlying trait were the Leadership/Authority scales from the Corry et al. (2008) and Ackerman et al. (2011) factor solutions. Collectively, these results therefore provide additional evidence for concerns over the psychometric properties of the NPI given that the scales shown to have the best precision and coverage are also believed by some to actually reflect constructs that should be separated from narcissism per se (see Brown et al., 2009; Rosenthal & Hooley, 2010). In short, the attributes that are well assessed by the NPI might not be considered narcissistic by all researchers. In contrast, the elements of personality that are considered uncontroversial manifestations of narcissism, such as entitlement and exploitativeness, are probably not as well measured by the NPI item pool. Thus, there is a need to create better measures of those constructs.

One of the limitations of this work is the sample used to study the psychometric properties of the NPI scales. Namely, it is based on a large convenience sample of college students. In our defense, the NPI was developed using college student samples and the majority of research in social and personality psychology uses such samples (see Gosling, Vazire, Srivastava, & John, 2004; Sears, 1986). Likewise, the vast majority of research with the NPI uses college student samples. Even so, future work should use more diverse samples and strive to obtain criterion variables using a multimethod strategy.

A related issue that needs to be addressed is the consequence of using a college sample to derive the item parameter estimates for the NPI. With person-anchoring, the 0 point on the latent scales for the NPI total score and its subscales were defined to be the mean score on the respective traits for the sample that we used. It is important to recognize that our inferences regarding

the relative difficulty and discrimination of the items on the NPI might be different if the scale was given to a clinical sample.

Unfortunately, few studies have administered the NPI to a clinical population. Of those few that have, the average scores for the NPI total score were actually higher in the nonclinical population than the clinical population (this is noteworthy considering that in the same study, structured clinical interviews found the clinical population to be more narcissistic than the nonclinical population; see Miller, Gaughan, Pryor, Kamen, & Campbell, 2009). This has again led some to question the utility of the overall NPI score as a measure of narcissism (Rosenthal & Hooley, 2010). Indeed, elevated scores for college students as opposed to a clinical sample is consistent with the idea that the overall NPI total score might include content that is controversial from the perspective of clinical accounts of narcissism.

With this caveat in mind, we expect that many of the items found to be of average difficulty in the current research might be less frequently endorsed by a clinical population and therefore appear “harder” as a result. Specifically, those items that are reflective of dominance and leadership ability (the majority of which are found in the NPI total scale and the Leadership/Authority subscales) will likely be found to have higher item difficulties in a clinical sample given that the trait underlying responses to these items will likely be lower in such a population. On the other hand, because clinical populations might possess higher levels of distress and interpersonal problems, they might endorse a higher proportion of items on the Entitlement/Exploitativeness subscale than a college student population. Consequently, the Entitlement/Exploitativeness items might appear “easier” in that population. In the end, if researchers would like to continue administering the NPI to clinical populations, it is critical that future research empirically evaluates whether the NPI and its subscales function differently across these two populations.

Another issue is that we developed the Ackerman et al. (2011) subscales using this data set, which might mean that the psychometric properties of the scales will be less optimal in additional samples. However, Ackerman et al. showed that their factor solution generally replicated across multiple data sets. Moreover, we repeated the IRT analyses reported here using the data set in Maxwell et al. (2011) and obtained similar results. With minor deviations in regard to measurement bandwidth and overall precision, the TIF for the NPI total score, and the IICs and TIFs for the NPI subscales, were essentially identical. The virtue of this sample is the very large sample size (and therefore enhanced precision of the IRT parameters) compared to the 586 participants in Maxwell et al. (50% women). The general convergence in results (available on request) should minimize concerns that these results will not generalize to a sample taken at a different college campus (viz., a Midwestern university with a predominantly White subject pool).

In closing, the application of modern psychometric tools to study the properties of the NPI yielded several insights. First, not all of the NPI scales were equally informative about their corresponding traits across a reasonable range of the latent trait continuum. A second insight is that the NPI scales differed in what location of the latent trait continuum they were measuring most precisely. Finally, particularly good items were identified, which offers researchers clues as to the content that can be included in further scale development work. In sum, the results reported here underscore recent arguments regarding the need to

revise the NPI to make further progress on studying the attributes of personality linked to narcissism (e.g., Brown et al., 2009).

ACKNOWLEDGMENT

Christopher Hopwood provided helpful comments on a previous draft.

REFERENCES

- Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., & Kashy, D. A. (2011). What does the Narcissistic Personality Inventory really measure? *Assessment, 18*, 67–87.
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- Barry, C. T., Frick, P. J., & Killian, A. L. (2003). The relation of narcissism and self-esteem to conduct problems in children: A preliminary investigation. *Journal of Clinical Child and Adolescent Psychology, 32*, 139–152.
- Bradlee, P. M., & Emmons, R. A. (1992). Locating narcissism within the interpersonal circumplex and the five-factor model. *Personality and Individual Differences, 13*, 821–830.
- Brown, R. P., Budzek, K., & Tamborski, M. (2009). On the meaning and measure of narcissism. *Personality and Social Psychology Bulletin, 35*, 951–964.
- Cain, N. M., Pincus, A. L., & Ansell, E. B. (2008). Narcissism at the crossroads: Phenotypic description of pathological narcissism across clinical theory, social/personality psychology, and psychiatric diagnosis. *Clinical Psychology Review, 28*, 638–656.
- Campbell, W. K., Bonacci, A. M., Shelton, J., Exline, J. J., & Bushman, B. J. (2004). Psychological entitlement: Interpersonal consequences and validation of a self-report measure. *Journal of Personality Assessment, 83*, 29–45.
- Chen, W., & Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics, 22*, 265–289.
- Corbitt, E. M. (2002). Narcissism from the perspective of the Five-factor model. In P. T. Costa, Jr. & T. A. Widiger (Eds.), *Personality disorders and the Five-factor model of personality* (2nd ed., pp. 293–298). Washington, DC: American Psychological Association.
- Corry, N., Merritt, R. D., Mrug, S., & Pamp, B. (2008). The factor structure of the Narcissistic Personality Inventory. *Journal of Personality Assessment, 90*, 593–600.
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: Guilford.
- del Rosario, P. M., & White, R. M. (2005). The Narcissistic Personality Inventory: Test–retest stability and internal consistency. *Personality and Individual Differences, 39*, 1075–1081.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.
- Emmons, R. A. (1984). Factor analysis and construct validity of the Narcissistic Personality Inventory. *Journal of Personality Assessment, 48*, 291–300.
- Emmons, R. A. (1987). Narcissism: Theory and measurement. *Journal of Personality and Social Psychology, 52*(1), 11–17.
- Frable, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*, 350–365.
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust Web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist, 59*, 93–104.
- Gray-Little, B., Williams, V. S. L., & Hancock, T. D. (1997). An item response theory analysis of the Rosenberg Self-Esteem scale. *Personality and Social Psychology Bulletin, 23*, 443–451.
- Johnson, W., Spinath, F., Krueger, R. F., Angleitner, A., & Riemann, R. (2008). Personality in Germany and Minnesota: An IRT-based comparison of MPQ self-reports. *Journal of Personality, 76*, 665–706.
- Konrath, S., Bushman, B. J., & Campbell, W. K. (2006). Attenuating the link between threatened egotism and aggression. *Psychological Science, 17*, 995–1001.
- Lessard, J., Greenberger, E., Chen, C., & Farruggia, S. (2011). Are youths’ feelings of entitlement always “bad”? Evidence for a distinction between

- exploitive and non-exploitive dimensions of entitlement. *Journal of Adolescence*, 34, 521–529.
- Maxwell, K., Donnellan, M. B., Hopwood, C. J., & Ackerman, R. A. (2011). The two faces of Narcissus? An empirical comparison of the Narcissistic Personality Inventory and the Pathological Narcissism Inventory. *Personality and Individual Differences*, 50, 577–582.
- McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, 15, 28–50.
- Miller, J. D., & Campbell, W. K. (2008). Comparing clinical and social-personality conceptualizations of narcissism. *Journal of Personality*, 76, 449–476.
- Miller, J. D., Gaughan, E. T., Pryor, L. R., Kamen, C., & Campbell, W. K. (2009). Is research using the Narcissistic Personality Inventory relevant for understanding narcissistic personality disorder? *Journal of Research in Personality*, 43, 482–488.
- Morizot, J., Ainsworth, A. T., & Reise, S. T. (2007). Toward modern psychometrics: Application of item response theory models in personality research. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 407–423). New York, NY: Guilford.
- Pincus, A. L., Ansell, E. B., Pimentel, C. A., Cain, N. M., Wright, A. G. C., & Levy, K. N. (2009). Initial construction and validation of the pathological narcissism inventory. *Psychological Assessment*, 21, 365–379.
- Pincus, A. L., & Lukowitsky, M. R. (2010). Pathological narcissism and narcissistic personality disorder. *Annual Review of Clinical Psychology*, 6, 421–446.
- Ramsay, J. O. (2000). *TestGraf: A program for the graphical analysis of multiple choice test and questionnaire data*. Montreal, Canada: McGill University.
- Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports*, 45, 590.
- Raskin, R. N., & Hall, C. S. (1981). The Narcissistic Personality Inventory: Alternate form reliability and further evidence of construct validity. *Journal of Personality Assessment*, 45, 159–162.
- Raskin, R. N., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology*, 54, 890–902.
- Reise, S. P. (1990). A comparison of item- and person-fit methods of assessing model-data fit in IRT. *Applied Psychological Measurement*, 14, 127–137.
- Reise, S. P., & Haviland, M. G. (2005). Item response theory and the measurement of clinical change. *Journal of Personality Assessment*, 84, 228–238.
- Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, 92, 544–559.
- Reise, S. P., & Waller, N. G. (1990). Fitting the two-parameter model to personality data. *Applied Psychological Measurement*, 14, 45–58.
- Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, 17(5), 1–25.
- Rosenberg, M. (1965). *Society and adolescent self-image*, Princeton, NJ: Princeton University.
- Rosenthal, S. A., & Hooley, J. M. (2010). Narcissism assessment in social-personality research: Does the association between narcissism and psychological health result from a confound with self-esteem? *Journal of Research in Personality*, 44, 453–465.
- Sears, D. O. (1986). College sophomores in the lab: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51, 515–530.
- Slocum-Gori, S. L., Zumbo, B. D., Michalos, A. C., & Diener, E. (2009). A note on the dimensionality of quality of life scales: An illustration with the Satisfaction with Life Scale (SWLS). *Social Indicators Research*, 92, 489–496.
- Trull, T. J., & McCrae, R. R. (2002). A Five-factor perspective on personality disorder research. In P. T. Costa, Jr. & T. A. Widiger (Eds.), *Personality disorders and the Five-factor model of personality* (2nd ed., pp. 45–58). Washington, DC: American Psychological Association.
- Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2008). Is “generation me” really more narcissistic than previous generations? *Journal of Personality*, 76, 903–918.
- Yen, W. M. (1981). Using simulation results to choose a latent trait model. *Applied Psychological Measurement*, 5, 245–262.

APPENDIX A.—Item content for the Corry et al. (2008) and Ackerman et al. (2011) scales.

Item	Response Options	Proportion Endorsed	Corry et al. Scale	Ackerman et al. Scale
1	I have a natural talent for influencing people^a	.69	Leadership/Authority	Leadership/Authority
	I am not good at influencing people.	.31		
2	Modesty doesn't become me^a	.17		
	I am essentially a modest person.	.83		
3	I would do almost anything on a dare^a	.23	Exhibitionism/Entitlement	
	I tend to be a fairly cautious person.	.77		
4	I know that I am good because everybody keeps telling me so^a	.17	Exhibitionism/Entitlement	Grandiose Exhibitionism
	When people compliment me I sometimes get embarrassed	.83		
5	If I ruled the world it would be a much better place^a	.38		Leadership/Authority
	The thought of ruling the world frightens the hell out of me	.62		
6 ^b	I can usually talk my way out of anything^a	.32		
	I try to accept the consequences of my behavior	.68		
7 ^{b,c,e}	I like to be the center of attention^a	.33	Exhibitionism/Entitlement	Grandiose Exhibitionism
	I prefer to blend in with the crowd	.67		
8	I will be a success^a	.80	Leadership/Authority	
	I am not too concerned about success.	.20		
9 ^b	I think I am a special person^a	.45		
	I am no better or no worse than most people.	.55		
10 ^b	I see myself as a good leader^a	.54	Leadership/Authority	Leadership/Authority
	I am not sure if I would make a good leader.	.46		
11 ^b	I am assertive^a	.49	Leadership/Authority	Leadership/Authority
	I wish I were more assertive.	.51		
12 ^b	I like having authority over people^a	.49	Leadership/Authority	Leadership/Authority
	I don't mind following orders.	.51		

APPENDIX A.—Item content for the Corry et al. (2008) and Ackerman et al. (2011) scales. (Continued)

Item	Response Options	Proportion Endorsed	Corry et al. Scale	Ackerman et al. Scale
13	I find it easy to manipulate people^a I don't like it when I find myself manipulating people.	.31 .69		Entitlement/Exploitativeness
14 ^{b,c}	I insist upon getting the respect that is due me^a I usually get the respect that I deserve	.30 .70	Exhibitionism/Entitlement	Entitlement/Exploitativeness
15 ^{b,c,e}	I like to display my body^a I don't particularly like to show off my body.	.24 .76	Exhibitionism/Entitlement	Grandiose Exhibitionism
16	I can read people like a book^a People are sometimes hard to understand.	.48 .52		
17	I like to take responsibility for making decisions^a If I feel competent I am willing to take responsibility for making decisions.	.57 .43		
18	I want to amount to something in the eyes of the world^a I just want to be reasonably happy.	.33 .67		
19 ^{c,e}	I like to look at my body^a My body is nothing special.	.36 .64	Exhibitionism/Entitlement	Grandiose Exhibitionism
20	I am apt to show off if I get the chance^a I try not to be a show off.	.23 .77	Exhibitionism/Entitlement	Grandiose Exhibitionism
21 ^b	I always know what I am doing^a Sometimes I am not sure of what I am doing.	.27 .73		
22 ^b	I rarely depend on anyone else to get things done^a I sometimes depend on people to get things done.	.45 .55		
23	Everybody likes to hear my stories^a Sometimes I tell good stories.	.19 .81		
24 ^{b,c}	I expect a great deal from other people^a I like to do things for other people	.26 .74	Exhibitionism/Entitlement	Entitlement/Exploitativeness
25 ^{b,c}	I will never be satisfied until I get all that I deserve^a I take my satisfactions as they come	.21 .79	Exhibitionism/Entitlement	Entitlement/Exploitativeness
26 ^{b,e}	I like to be complimented^a Compliments embarrass me	.60 .40		Grandiose Exhibitionism
27	I have a strong will to power^a Power for its own sake doesn't interest me	.40 .60	Leadership/Authority	Leadership/Authority
28 ^b	I like to start new fads and fashions^a I don't very much care about new fads and fashions	.37 .63	Exhibitionism/Entitlement	Grandiose Exhibitionism
29 ^{b,c,e}	I like to look at myself in the mirror^a I am not particularly interested in looking at myself in the mirror.	.56 .44	Exhibitionism/Entitlement	Grandiose Exhibitionism
30	I really like to be the center of attention^a It makes me uncomfortable to be the center of attention.	.41 .59	Exhibitionism/Entitlement	Grandiose Exhibitionism
31 ^b	I can live my life in any way I want to^a People can't always live their lives in terms of what they want.	.53 .47		
32	People always seem to recognize my authority^a Being an authority doesn't mean that much to me.	.29 .71	Leadership/Authority	Leadership/Authority
33	I would prefer to be a leader^a It makes little difference to me whether I am a leader or not.	.47 .53	Leadership/Authority	Leadership/Authority
34 ^{b,d}	I am going to be a great person^a I hope I am going to be successful.	.54 .46		Leadership/Authority
35	I can make anybody believe anything I want them to^a People sometimes believe what I tell them.	.29 .72		
36	I am a born leader^a Leadership is a quality that takes a long time to develop.	.24 .76	Leadership/Authority	Leadership/Authority
37	I wish somebody would someday write my biography^a I don't like people to pry into my life for any reason.	.42 .58		
38 ^b	I get upset when people don't notice how I look when I go out in public^a I don't mind blending into the crowd when I go out in public.	.22 .78	Exhibitionism/Entitlement	Grandiose Exhibitionism
39 ^{b,c}	I am more capable than other people^a There is a lot that I can learn from other people.	.26 .74	Exhibitionism/Entitlement	
40	I am an extraordinary person^a I am much like everybody else.	.42 .58		Leadership/Authority

Note. We recommend that Items 5 and 40 from the Ackerman et al. Leadership/Authority subscale and Items 14, 24, and 25 from the Corry et al. Exhibitionism/Entitlement subscale be dropped in future research.

^aItems that indicate the narcissistic response. ^bItems that were discarded for the item response theory (IRT) analysis with the trimmed Narcissistic Personality Inventory total score.

^cItems that were discarded for the IRT analysis with the trimmed Corry et al. Exhibitionism/Entitlement scale. ^dItems that were discarded for the IRT analysis with the trimmed Ackerman et al. Leadership/Authority subscale. ^eItems that were discarded for the IRT analysis with the trimmed Ackerman et al. Grandiose Exhibitionism subscale.