



# The Narcissistic Personality Inventory: Test–retest stability and internal consistency

Peter M. del Rosario \*, Royce M. White

*Marist College Psychology Department, 3399 North Road, Poughkeepsie, NY 12601-1387, USA*

Received 31 March 2004; received in revised form 11 July 2005; accepted 16 August 2005

## Abstract

The Narcissistic Personality Inventory (NPI) is a self-report measure that assesses subclinical levels of narcissism. It evolved through a 54-item, four-factor version to its current 40-item, seven-factor version. This study focuses on the 40-item NPI that produces a full-scale narcissism score and seven-factor-based subscale scores: *authority*, *exhibitionism*, *superiority*, *entitlement*, *exploitativeness*, *self-sufficiency*, and *vanity*. This study is the first to examine its test–retest reliability. The stability coefficients are expected to be high as personality characteristics are presumed to represent enduring and stable traits. The internal consistency is also investigated, particularly as this version has fewer items, yet a greater number of factor-based subscales than its predecessor. The alpha coefficients produced by data from test, retest, and previously published research are compared. The NPI was administered to 175 American college students, and re-administered 13 weeks later. The NPI demonstrated significant test–retest correlations for all scales. Across all results, only the *full-scale* and *authority* subscale possessed adequate ( $\alpha > .70$ ) internal consistency, thus raising questions regarding the utility of the remaining scales.

© 2005 Elsevier Ltd. All rights reserved.

**Keywords:** Narcissistic Personality Inventory; Test–retest reliability; Stability; Reliability; Internal consistency; Components; Factor structure; American college students

\* Corresponding author. Tel.: +1 845 575 3000; fax: +1 845 575 3965.  
E-mail address: [peter.delrosario@marist.edu](mailto:peter.delrosario@marist.edu) (P.M. del Rosario).

## 1. Introduction

The Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979; Raskin & Terry, 1988) is a self-report inventory designed to measure narcissism in non-clinical populations. Though the scale is not a measure of narcissistic personality disorder, the conceptualization of narcissism was based on the behavioral criteria for narcissistic personality from the third version of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III; American Psychiatric Association, 1980). Narcissism is a style of personality that is characterized by grandiosity; fantasies of high levels of power, beauty, and success; high sensitivity to criticism; strong sense of entitlement; and exploitativeness in relationships. These basic criteria have remained essentially unchanged in the current DSM-IV (American Psychiatric Association, 1994).

With the development of theories on narcissism, and measures such as the NPI, there has been a significant increase in psychological research on narcissism over the past few decades. According to Soyer, Rovenpor, Kopelman, Mullins, and Watson (2001), “PsychInfo database references to narcissism and to the narcissistic personality...increased from 405 in the 10-year period from 1969–1978, to 1322 during 1979–1988, and to 1791 during 1989–1998” (p. 245). Our own Psych-Info search revealed that there are currently at least 146 references for research involving the NPI. There are at least 25 references during 1979–1988, 82 during 1989–1998, and at least 39 during 1999–2003. Thus, the NPI has been widely used and appears to have achieved acceptance as a measure of subclinical narcissism.

In the test development process, the NPI originated from a pool of 223 items. As each item presents a pair of self-attitude statements, the respondent chooses the one statement they agree with most. The NPI evolved through a number of analyses, resulting in a 54-item version of the NPI (Raskin & Hall, 1979, 1981). Emmons (1984, 1987) conducted a factor analysis of the 54-item version revealing four components that he termed *Leadership/Authority*, *Superiority/Arrogance*, *Self-Absorption/Self-Admiration*, and *Exploitativeness/Entitlement*. The NPI could then be interpreted as producing a *full-scale* narcissism score and four-factor-based subscale scores. Each subscale was composed of 9–12 non-overlapping items. Using a sample of 451 undergraduates, Emmons (1984) examined the internal consistencies of the *full-scale* and each subscale and found that the Cronbach alpha coefficients (Cronbach, 1951) were 0.86, 0.79, 0.69, 0.69, and 0.74, respectively. Emmons (1987) found very similar alpha coefficients in a later study utilizing a sample of 362 undergraduates. The respective alpha coefficients were 0.87, 0.69, 0.70, 0.81, and 0.68. Emmons concluded that the 54-item NPI internal consistencies were satisfactory. In addition, a number of published studies provided evidence for the 54-item version’s reliability and validity (Emmons, 1984, 1987; Prifitera & Ryan, 1984; Raskin & Hall, 1979, 1981; Watson, Grisham, Trotter, & Biderman, 1984; Watson, Hood, & Morris, 1984).

The NPI continued to evolve to its current form as Raskin and Terry (1988) performed a principal-components analysis, leading them to further reduce the NPI to a parsimonious 40-item, seven-factor version. In addition to the NPI *full-scale* score, these factors constitute the seven component subscales of the NPI: *authority*, *exhibitionism*, *superiority*, *entitlement*, *exploitativeness*, *self-sufficiency*, and *vanity*. Each component scale is comprised of 3–8 non-overlapping items. Raskin and Terry (1988) reported the Guttman lambda-3 (alpha) coefficients (Guttman, 1945) for the *full-scale* and component subscales to be 0.83, 0.73, 0.63, 0.54, 0.50, 0.52, 0.50, and 0.64, respectively ( $n = 1018$ ). All scales, except for the *full-scale* and *authority* scale, appear to be low enough

to raise questions regarding the utility of those subscales. It should be noted that the Guttman lambda-3 is the equivalent of the Cronbach alpha coefficient (SPSS, 1999; Traub, 1994, p. 87).

Surprisingly, little published information has been available regarding the reliability of the 40-item NPI. To date, no investigation of the NPI's test–retest stability has been published. Therefore, the first purpose of our study was to examine the test–retest stability of the 40-item NPI and its component scales. The reliability coefficients were expected to be relatively high as narcissism is presumed to be a stable and enduring personality characteristic. Given the relatively low Guttman lambda-3 (alpha) coefficients of six of the eight scales (Raskin & Terry, 1988), the second purpose of our study was to examine the NPI's internal consistency. While the standard practice in psychometric research is to report the internal consistency of scales following one administration, our study allows the inspection of the internal consistencies by comparing the alpha coefficients produced from our test and retest data, and those from the research of Raskin and Terry (1988).

## 2. Method

### 2.1. Participants

A total of 175 American undergraduate students (144 females, 30 males, and one missing the gender datum) at a small, private, liberal arts college in the northeastern United States participated in the study. Age of participants ranged between 18 and 53 years ( $M = 21.11$ ,  $SD = 5.81$ ). Students were permitted to self-identify their race in more than one category. The racial/ethnic composition was 92.6% White ( $n = 162$ ), 3.4% Latino ( $n = 6$ ), 1.1% African-American ( $n = 2$ ), 1.1% White and Latino ( $n = 2$ ), 0.6% African-American and Latino ( $n = 1$ ), 0.6% African-American and Native American Indian ( $n = 1$ ), and 0.6% Asian American ( $n = 1$ ). For comparison purposes, it should be noted that the participants in Raskin and Terry (1988) research were 1018 American undergraduate students (529 females, 479 males) with an age range between 17 and 49 years ( $M = 20$ ,  $SD = 6.7$ ).

### 2.2. Procedure

The participants volunteered to participate in a study that required them to complete a demographic questionnaire, the NPI, and two other measures at two points in time. Participants responded anonymously and all scales were administered in a classroom setting. They took the NPI during the first week of a college semester and again near the end of that semester, approximately 13 weeks later, for a mean test–retest interval of 89.9 days ( $SD = 9.8$ ), a median interval of 92.5 days, with a minimum of 63 and maximum of 99 days.

## 3. Results and discussion

Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS), version 10.1. Table 1 presents the test and retest means, standard deviations, and test–retest

Table 1

Test–retest means, standard deviations, and correlations for NPI scales ( $N = 175$ )

Scale	Test		Retest		Test/retest $r$
	$M$	SD	$M$	SD	
Full-scale	14.66	6.11	14.23	6.36	0.81*
Authority	4.17	2.13	4.07	2.17	0.80*
Exhibitionism	1.85	1.67	1.82	1.77	0.77*
Superiority	2.04	1.37	2.14	1.41	0.70*
Entitlement	1.49	1.34	1.36	1.28	0.57*
Exploitativeness	1.49	1.23	1.38	1.10	0.60*
Self-sufficiency	2.45	1.41	2.38	1.44	0.60*
Vanity	1.18	1.05	1.09	1.12	0.73*

\*  $p < 0.01$ , two-tailed.

stability coefficients for the NPI Scales. The stability correlations for the seven components scales and the *full-scale* ranged from 0.57 to 0.81 and all correlations were significant,  $p < 0.01$ , two-tailed. The *full-scale* ( $r = 0.81$ ) scores produced a high ( $r > 0.80$ ) test–retest correlation. Four of the NPI component scales, *authority* ( $r = 0.80$ ), *exhibitionism* ( $r = 0.77$ ), *superiority* ( $r = 0.70$ ), and *vanity* ( $r = 0.73$ ), yielded test–retest correlations in the high moderate range (between  $r = 0.70$  and  $r = 0.80$ ). Two component scales, *exploitativeness* ( $r = 0.60$ ), and *self-sufficiency* ( $r = 0.60$ ), produced test–retest correlations in the low moderate range (between  $r = 0.60$  and  $r = 0.70$ ). The *entitlement* scale ( $r = 0.57$ ) correlation fell just below the low moderate range. The significant test–retest correlations in Table 1 provide evidence that the 40-item NPI scale and its seven components possessed satisfactory stability. Furthermore, considering that these results were obtained over a test–retest interval of about 14 weeks, they support the notion that narcissism, like other personality characteristics, is an enduring set of personality characteristics that remain relatively stable over time.

Table 2 presents a comparison of the test and retest Cronbach coefficient alphas (Cronbach, 1951) for the NPI scales. Raskin and Terry (1988) reported the internal consistency Guttman lambda-3s (alphas) for the NPI scales that were based on a sample of 1018 college students. These are also presented in Table 2. As mentioned previously, Cronbach alpha and Guttman lambda-3 are equivalent coefficients of internal consistency (SPSS, 1999; Traub, 1994, p. 87). The alpha coefficients from our initial test ranged from 0.39 to 0.80, with an average of 0.57. The alpha coefficients from our retest ranged from 0.30 to 0.82, with an average of 0.59. Raskin and Terry reported Guttman lambda-3 (alphas) that ranged from 0.50 to 0.83, with an average of 0.61. As can be seen in Table 2, the internal consistencies from the test, retest, and previous research were quite comparable across each of the following scales: *full-scale*, *authority*, *exhibitionism*, *superiority*, *entitlement*, and *self-sufficiency*. The coefficients for the *exploitativeness* ( $\alpha = 0.45$ , 0.30, and 0.52, respectively) and *vanity* ( $\alpha = 0.58$ , 0.74, and 0.64, respectively) scales were less comparable.

As shown in Table 2, across the test, retest, and previously reported (Raskin & Terry, 1988) results, the internal consistency of the *full-scale* narcissism scores were satisfactory with alpha coefficients of 0.80, 0.82, and 0.83, respectively. Of the seven component scales, only the *authority* scale possessed adequate ( $\alpha > 0.70$ ) internal consistency across the results, with alpha coefficients of 0.72, 0.73, and 0.73, respectively. Finally, the remaining six scales, *exhibitionism*, *vanity*

Table 2

Test, retest, and previously reported coefficient  $\alpha$ s, and number of items for NPI scales

Scale	Test $\alpha^a$	Retest $\alpha^a$	Reported Guttman lambda-3 ( $\alpha$ ) <sup>b</sup>	No. of items
Full-scale	0.80	0.82	0.83	40
Authority	0.72	0.73	0.73	8
Exhibitionism	0.62	0.68	0.63	7
Superiority	0.53	0.55	0.54	5
Entitlement	0.48	0.46	0.50	6
Exploitativeness	0.45	0.30	0.52	5
Self-sufficiency	0.39	0.46	0.50	6
Vanity	0.58	0.74	0.64	3

*Note.* Previously reported Guttman lambda-3 (alpha) coefficients are from “A Principal-Components Analysis of the Narcissistic Personality Inventory and Further Evidence of its Construct Validity,” by Raskin and Terry (1988).

<sup>a</sup>  $n = 175$ .

<sup>b</sup>  $n = 1018$ .

*superiority, entitlement, exploitativeness, and self-sufficiency*, produced inadequate alpha coefficients across the test, retest, and previous results.

A concern that arose in Raskin and Terry (1988) research continued to appear in the present study, that is, the production of low internal consistency coefficients for several of the component scales. They theorized that their alphas were lower than desired due to the fact that they were based on too few items per scale. Indeed, Nunnally and Bernstein (1994) point out that scales that produce low alphas may be too short or the items may have little in common. Furthermore, all things being equal, a major way to make tests more reliable is to make them longer. Examination of the number of items comprising the component scales showed that each of the seven scales consists of only 3–8 items. The finding of inadequate internal consistency in six of the seven component scales across these studies raises questions regarding the utility of those six scales. Interestingly, the *authority* scale produced a set of satisfactory alpha coefficients, yet was comprised of only 8 items. Perhaps future work on the NPI might utilize the Spearman–Brown formula to estimate how many more items each scale would need to bring the reliability to a satisfactory level (Nunnally & Bernstein, 1994, p. 263). As is, the utility of the 40-item NPI may be more limited to the *full-scale* narcissism and *authority* scale scores. Finally, these results raise the question of whether or not the earlier 54-item, four-factor NPI version (with satisfactory internal consistency alpha coefficients) might be preferable.

One methodological concern that should be noted has to do with the sample size of the current study. In particular, the comparisons of internal consistency estimates are based on sample sizes that vary greatly: the current sample of 175 and Raskin and Terry (1988) sample of 1018. It should be noted that as the sample size in a reliability study is increased, the precision of the population reliability estimate will increase (Charter, 1999, 2003). Kline (1986) recommends a sample size of at least 200 for the determination of adequately precise reliability coefficients. Nunnally and Bernstein (1994) recommend a sample size of at least 300 for reliability estimation. Charter's (1999, 2003) investigations of minimum sample sizes needed for reliability studies have led him to conclude that a sample size of 400 is required for sufficiently precise reliability estimation. Bearing

this in mind, it is likely that Raskin and Terry's (1988) reliability estimates have reached a high level of precision. In contrast, the current study's alpha coefficients based on a sample of 175 falls short of the recommended minimum sample sizes and are therefore, less precise (i.e. a wider confidence interval for the alpha coefficients).

Another methodological limitation should be noted in relation to these findings. Specifically, the generalizability of the results is limited by the nature of the sample used: mainly American college-aged women who identified their race/ethnicity as White. The sample's gender imbalance in particular, may be problematic as some research using the 54-item NPI found that narcissism was more an issue for men than for women (Carroll, 1987, 1989; Joubert, 1989; Watson et al., 1984; Watson, Taylor, & Morris, 1987). Furthermore, the DSM-IV indicates that of those diagnosed with narcissistic personality disorder, 50–75% are males (American Psychiatric Association, 1994, p. 660). Such methodological shortcomings would have been avoided had the present research planned specifically to investigate the psychometric properties of the NPI from the outset. Additional research is needed to examine the reliability of the NPI, particularly with larger, more representative samples from non-clinical and clinical populations.

## Acknowledgements

The authors wish to thank the graduate assistants, particularly Lindsay Hayes, for their invaluable assistance. We also thank all of the students who volunteered for the study, and our faculty. Finally, our gratitude goes to Robert Raskin for allowing us to conduct research with the NPI.

## References

- American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Carroll, L. (1987). A study of narcissism, affiliation, intimacy, and power motives among students in business administration. *Psychological Reports*, 61, 355–358.
- Carroll, L. (1989). A comparative study of narcissism, gender, and sex-orientation among bodybuilders, athletes, and psychology students. *Psychological Reports*, 64, 999–1006.
- Charter, R. A. (1999). Sample size requirements for precise estimates of reliability, generalizability, and validity coefficients. *Journal of Clinical and Experimental Neuropsychology*, 21, 556–559.
- Charter, R. A. (2003). Study samples are too small to produce sufficiently precise reliability coefficients. *Journal of General Psychology*, 130, 117–129.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- 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, 11–17.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255–282.
- Joubert, C. E. (1989). The Famous Sayings test: sex differences and some correlations with other variables. *Psychological Reports*, 64, 763–766.
- Kline, P. (1986). *A handbook of test construction: Introduction to psychometric design*. NY: Methuen.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). NY: McGraw-Hill.

- Prifitera, A., & Ryan, J. (1984). Validity of the Narcissistic Personality Inventory NPI in a psychiatric sample. *Journal of Clinical Psychology*, 40, 140–142.
- 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., & 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(5), 890–902.
- Soyer, R. B., Rovenpor, J. L., Kopelman, R. E., Mullins, L. S., & Watson, P. J. (2001). Further assessment of the construct validity of four measures of narcissism: replication and extension. *Journal of Psychology*, 135, 245–259.
- SPSS (1999). SPSS base 10.0 applications guide. Chicago: SPSS, Inc.
- Traub, R. E. (1994). *Reliability for the social sciences: theory and applications*. CA, Sage: Thousand Oaks.
- Watson, P. J., Grisham, S. O., Trotter, M. V., & Biderman, M. D. (1984). Narcissism and empathy: validity evidence for the Narcissistic Personality Inventory. *Journal of Personality Assessment*, 48, 301–305.
- Watson, P. J., Hood, R. W., & Morris, R. J. (1984). Religious orientation, humanistic values, and narcissism. *Review of Religious Research*, 25, 257–264.
- Watson, P. J., Taylor, D., & Morris, R. J. (1987). Narcissism, sex roles, and self functioning. *Sex Roles*, 16, 335–350.