Short Report

Do Narcissists Dislike Themselves ‘‘Deep Down Inside’’?

W. Keith Campbell,1 Jennifer K. Bosson,2 Thomas W. Goheen,3 Chad E. Lakey,1 and Michael H. Kernis1

Narcissism is a personality trait associated with an inflated, grandiose self-concept and a lack of intimacy in interpersonal relationships. A popular assumption is that narcissists’ positive explicit (conscious) self-views mask implicit (nonconscious) self-loathing. This belief is typically traced to psychodynamic theory, especially that of Kohut (1966; Morrison, 1983). Empirically, this view predicts that narcissists will reveal negative self-views when these are measured with unobtrusive instruments—such as the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998)—that record people’s automatic, uncontrolled responses. Using the IAT, however, researchers found no simple relation between narcissism and implicit self-esteem (r = −.13 and .03; Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003; Zeigler-Hill, 2006).1

According to another line of thought, narcissists’ explicit self-views are not uniformly positive; rather, narcissism is associated with positive self-views in agentic domains (e.g., status, intelligence), but not in communal domains (e.g., kindness, morality). Evidence for this idea comes from both explicit trait ratings, which show an association between narcissism and positive self-views only on agentic traits (Campbell, Rudich, & Sedikides, 2002), and from analyses showing that narcissism is particularly strongly associated with self-esteem measures that capture dominance (Brown & Zeigler-Hill, 2004). Bradlee and Emmons (1992) and Paulhus and Williams (2002) have also reported personality data supporting this distinction.

As narcissists do not evaluate themselves uniformly positively across all dimensions—and the self-esteem IAT measures the strength of cognitive associations between the self and an evaluative dimension—the lack of correlation between narcissism and implicit self-esteem might reflect the words used in the IAT. Specifically, IATs employing more agentic words may correlate positively with narcissism, whereas those using more communal words may correlate negatively or not at all with narcissism. Indeed, this pattern is seen in narcissists’ implicit responses on the Thematic Apperception Test. On this test, narcissism correlates positively with nPower and negatively with nIntimacy and nAffiliation (Carroll, 1987).

Researchers often use IAT words that activate respondents’ communal self-views more than their agentic ones. For example, the IAT words used by Jordan et al. (2003; friend, gift, happy, holiday, joy, love, party, smile, sunshine, warmth, agony, cockroach, death, disaster, disease, evil, garbage, pain, sink, and vomit) and by Zeigler-Hill (2000; happy, joy, paradise, pleasure, smile, sunshine, agony, death, grief, pain, sickness, and tragedy) include several communal terms and few agentic terms. Twenty-four pilot respondents rated both word lists along agentic and communal dimensions. Both lists conveyed significantly more communion than agency, t(23)s > 4.50, p_{rep}s > .997, which might explain the weak associations these researchers observed between narcissism and IAT scores.

We tested our logic that narcissism correlates positively with implicit agency and negatively or not at all with implicit communion in two studies. Using the IAT words from Jordan et al. (2003), we tested the link between narcissism and implicit self-esteem in a sample of undergraduates. Next, we created separate IATs to measure agentic and communal implicit self-views and tested their associations with narcissism.

STUDY 1

Subjects (N = 154) completed the Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988), the Rosenberg (1965) self-esteem scale (RSES), and the IAT (Jordan et al., 2003). Results replicated past findings: Narcissism was uncorrelated with implicit self-esteem, r = .13, p_{rep} = .872, and positively correlated

1Although our main interest here is in the simple correlation between narcissism and implicit self-esteem, these researchers focused on the interaction of implicit and explicit self-esteem in predicting narcissism. We tested this interaction in the two data sets reported here, but did not find it. We did, however, find a similar interaction pattern using explicit self-esteem and the communion IAT in Study 2.

Address correspondence to W. Keith Campbell, Department of Psychology, University of Georgia, Athens, GA 30602, e-mail: wkc@uga.edu.
with explicit self-esteem, $r = .30, p_{rep} = .987$. A meta-analysis of the link between the NPI and IAT, combining our current data with those of Jordan et al. (2003) and Zeigler-Hill (2006), yielded an overall association not different from zero ($r = .04, p_{rep} = .636$, combined $N = 331$). Thus, IATs that assess primarily communal self-views do not correlate with narcissism.

**STUDY 2**

We tailored three new IATs to measure implicit agency, implicit communion, and implicit self-esteem. Our agency IAT used words that reflected high versus low agency (assertive, active, energetic, outspoken, dominant, enthusiastic, quiet, reserved, silent, withdrawn, submissive, and inhibited); pilot ratings confirmed that these words were more agentic than communal, $t(23) = 9.20, p_{rep} > .999, d = 3.84$. The communion IAT used words that reflected high versus low communion (kind, friendly, generous, cooperative, pleasant, affectionate, mean, rude, stingy, quarrelsome, grouchy, and cruel); pilot ratings confirmed that these words were more communal than agentic, $t(23) = 4.95, p_{rep} = .998, d = 2.06$. The self-esteem IAT used positive and negative words (good, wonderful, great, right, bad, awful, terrible, and wrong) that reflected agency and communion equally, $t(23) = 1.20, p_{rep} = .796, d = 0.50$.

Subjects ($N = 114$) completed the three IATs (in varied order, which did not affect the findings), then rated themselves explicitly on the agentic and communal IAT words and completed the NPI and RSES. As Figure 1 shows, the NPI correlated positively with explicit and implicit agency ($r = .52$ and $r = .29, p_{rep} = .999$ and $.967$), but not with explicit or implicit communion ($r = -.12$ and $r = .01, p_{rep} = .816$ and $.530$). The association between narcissism and the agency IAT was reliably stronger than that between narcissism and the communion IAT, $t(112) = 2.21, p_{rep} = .939, d = 0.42$. Moreover, narcissism correlated with both the RSES and the self-esteem IAT ($r = .25$ and $r = .21, p_{rep} = .971$ and $.944$).

**DISCUSSION**

These results suggest that narcissists do not uniformly dislike themselves “deep down inside.” Rather, narcissists report high explicit and implicit self-views on measures of agency, and neutral self-views on measures of communion. Thus, the association of narcissism with implicit self-esteem depends, in part, on the proportions of agency and communion captured by the implicit instrument. As Figure 1 shows, the existing self-esteem IAT, which activates primarily communal self-views, produced a nonsignificant correlation similar to that found with our communion IAT. However, our self-esteem IAT—which captures agency and communion equally—correlated positively with narcissism.

In sum, it may be imprecise to conceptualize narcissism as a positive explicit self-concept blanketing a negative implicit self-concept. Rather, narcissists exhibit a somewhat imbalanced self at both explicit and implicit levels, with favorable agentic self-views that are not necessarily matched by favorable communal self-views.

**Acknowledgments**—The first two authors contributed equally to this manuscript.

**REFERENCES**


(Received 5/12/06; Revision accepted 7/20/06; Final materials received 7/25/06)