**Overview**

**Responses to Editor and Review Comments**

Thank you for your feedback on **Manuscript PSPB-14-437**: “Narcissism and Self-Enhancement: A Review and Meta-Analysis.” We were delighted to be invited for a revision. Three main themes in this revision are summarized below:

1. More exhaustive literature search. Reviewer 2, Comment 1 advised us to expand our search methods, in order to potentially include more primary studies. We have responded to this comment by directing a great deal of effort toward more exhaustively searching for studies on the narcissism-self-enhancement relationship (i.e., performing a forward search, sending additional emails to scholars in the field, and using narcissism inventory names as keywords), and this effort paid off with an additional 10 usable primary study samples, for a total of 36 samples. The process is explained in our response to Editor Comment 3.
2. Additional/different analyses. Reviewer 2 provided many helpful recommendations for how we could improve our analyses. In response to these suggestions, we made many changes to our manuscript including the addition of: (a) publication bias analyses (Reviewer 2, Comment 2), (b) exact *p* values and confidence intervals for all of our results (Reviewer 2, Comment 7), and (c) an estimate of proportion of variance accounted for by each moderator (i.e., Pseudo-*R2*; Reviewer 2, Comment 3). Further, we have now performed our moderator analyses separately for agency and communion (Reviewer 2, Comment 5), and added another category to our acquaintanceship moderator such that we now have categories for zero acquaintance, short acquaintance, and long acquaintance (Reviewer 2, Comment 6).
3. Reframing/Rewriting. The Editor, Comment 2 advised that we needed to improve the initial framing of our paper to communicate earlier and more clearly to readers that we are focusing on trait narcissism (see also Reviewer 1, Comment 1) and on ‘self-insight’ self-enhancement. To do this, we have now removed references to pathological and clinical narcissism, as well as changing our title to, “Narcissism and Self-Insight: A Review and Meta-Analysis of Narcissists’ Self-Enhancement Tendencies.” These changes are explained more fully in Editor Comment 2. In addition, throughout the manuscript, we have updated and included more references on certain topics.

**Editor General Comments.**

**08-Jan-2015  
  
Re: "Narcissism and Self-Enhancement: A Review and Meta-Analysis" (MS # PSPB-14-437)  
  
Dear Dr. Grijalva:  
  
I have now received two reviews of your manuscript “Narcissism and Self-Enhancement: A Review and Meta-Analysis” (MS# PSPB-14-437).  Both reviewers are experts in topics relevant to your work, and I am very grateful to have received their feedback on your manuscript.  I also read your manuscript before I read the reviews so that my reading of it would remain unbiased by the reviewers’ comments.  
  
As you will see, both of the reviewers had very favorable reactions to your work.  They both agreed that your topic is important, and that your findings have the potential to make a strong contribution to the literature on narcissism and self-enhancement.  Both reviewers, however, also noted some issues and concerns that led them to conclude that publication would be premature at this point.  Therefore I am offering you a decision of Revise and Resubmit.  
  
The reviewers’ comments are very thorough and I will not reiterate all of them here.  My reading of the reviews suggests that a more favorable decision will require not only some rewriting and reframing, but also some additional analyses and (possibly) inclusion of additional studies in your meta-analysis.  In addition to the issues raised by the two reviewers, I had several thoughts when reading your manuscript and I list these below:**

Thank you for this summary, and for the helpful feedback on the manuscript. In the sections below, we respond to each reviewer comment.

**Editor Comment 1.**

**Since you exclude clinical samples from your search, you might frame the introduction to more clearly communicate your interest in non-clinical narcissism as a personality tendency.  For example, highlighting the clinical definition of narcissism from the DSM in the first sentence of your manuscript might not be ideal, if your focus is not on clinical definitions of narcissism.  In line with this, Reviewer 1 suggests getting rid of all references to clinical narcissism, because most readers of PSPB will expect you to focus on non-clinical narcissism.  This seems reasonable to me – if you open the manuscript with the definition of non-clinical narcissism that you used in your literature search, it would not be necessary to bring up clinical narcissism.**

This is a good point. We have now removed all references to clinical narcissism and the DSM. In addition, we reframed the sections of the paper that had previously relied on the DSM definition of narcissism.

We rewrote the first sentence of the paper to focus on trait narcissism, so that it now reads:

“Self-enhancement is a fundamental characteristic of narcissism. In fact, narcissism has even been called the “self-enhancer personality” (Morf, Horvath, & Torchetti, 2011, p. 399).” (p. 3).

Second, when introducing narcissism on page 4, we removed the definition from the DSM and instead wrote the following:

“Grandiosity is “characterized by affectation of grandeur or splendor or by absurd exaggeration” (Merriam-Webster’s online dictionary) and is the defining feature of the personality trait of narcissism. Narcissists like to be the center of attention, tend to show off, believe that they are special people, and prefer to be in leadership roles and roles imbued with power (Emmons, 1987; Raskin & Terry, 1988; Rhodewalt, 2011).”

**Editor Comment 2.**

**I wonder if the title and initial framing of your study are as precise as possible?  You aren’t interested in “self-enhancement” broadly defined, but instead, in a very specific type of self-enhancement, i.e., self-other discrepancy (what you call “self-insight” self-enhancement).  I think a title that more clearly conveys the type of self-enhancement you’re interested in here would be ideal, because it will give readers a clearer sense of the contents of your manuscript.**

To give readers a clearer sense of the contents of our manuscript we changed the title from “Narcissism and self-enhancement: A meta-analysis and review”, so that the new title is now “Narcissism and Self-Insight: A Review and Meta-Analysis of Narcissists’ Self-Enhancement Tendencies.”

Further, in our abstract we added the following sentence:

“Specifically, we focus on self-insight self-enhancement (i.e., whether narcissists perceive themselves more positively than they are perceived by others), thus we only include studies that compare narcissists’ self-reports to observer-reports or objective measures.”

Also, early in the paper’s introduction (i.e., the first page of our paper) we added the following sentence:

“The current work therefore comprehensively reviews and meta-analyzes the narcissism-self-enhancement bias literature. Specifically, we will focus on self-insight self-enhancement, which is measured by comparing self-reports to external criteria (i.e., observer-reports and objective measures).” (p. 3).

**Editor Comment 3.**

**I was a bit surprised that your literature search only yielded 26 usable data sets.  This concern is echoed by Reviewer 2, who asks whether you did a forward search of your articles, and suggests that you expand your search to include the clinical literature.  Even though you are not interested in clinical narcissism, some clinical scholars operationalize narcissism as a personality trait rather than a clinical diagnosis, and their works could be relevant here.  It also occurred to me that you might broaden your database searches beyond keyword searches to searches of “all text,” and look for any articles that include the NPI (or other measures of narcissism that you included).  I understand that meta-analysis can be frustrating because the continual publication of new articles makes your “final” sample somewhat of a moving target.  Nonetheless, it would strengthen your case if you make another attempt to locate all usable data sets.**

We were also surprised that our first meta-analytic literature search only yielded 26 samples. Although many papers measure narcissism and collect self- and observer-reports, it is less common for the focus of the paper to be on self-enhancement, thus, quite frequently, necessary effect size information is not included in published papers.

We worked at length to address concerns about whether the literature review was exhaustive enough in terms of its inclusion of studies.

This entailed extending our search to encompass (a) additional measures of narcissism including some clinical measures (see list below)—in effect this meant that we also searched through clinical psychology papers, but (consistent with Reviewer 2’s recommendation) we continued to exclude clinical samples and samples that used measures of ‘vulnerable narcissism’, (b) we carried out a forward search of all articles included in the meta-analysis by looking for more recent papers that cited those papers, (c) to further make sure we did not miss any recent studies, we searched every single paper published in the last five years that contained the words ‘narcissism’ or ‘narcissist’ or ‘narcissistic’, and (d) directly contacted 19 authors to request effect sizes that were not included in published papers.

**Narcissism inventories included in the keyword list:**

“OMNI”

“Personality Diagnostic Questionnaire-4” (PDQ-4)

"Structured Clinical Interview for DSM Disorders" (SCID)

"Diagnostic Interview for DSM- IV Personality Disorders" (DIPD-IV)

"International Personality Disorders Examination" (IPDE)

"Personality Disorder Interview- IV" (PDI-IV)

"Structured Interview for DSM- IV Personality Disorders" (SIDP-IV)

"Diagnostic Interview for Narcissism" (DIN)

“Schedule for Nonadaptive and Adaptive Personality” (SNAP)

“Millon Clinical Multiaxial Inventory” (MCMI-III)

“Minnesota Multiphasic Personality Inventory” (MMPI-2)

“The Dirty Dozen”

“Narcissistic Grandiosity Scale” (NGS)

“Hogan Developmental Survey” (HDS)

“California Personality Inventory” (CPI)

**General search for narcissism inventory keywords:**

PsycINFO = 3,548 papers searched

ProQuest Dissertation & Theses Full Text = 2,848 papers searched

**Forward search**

PsycINFO = 2,650 papers searched

**All papers published in the last five years (2010-2015) that included the words narcissism or narcissist or narcissistic:**

PsycINFO = 2,318 papers searched

This expanded literature search of 11,364 primary sources resulted in an additional 10 usable samples for a total of 36 samples.

**Editor Comment 4.**

**On page 8 you write “Regardless of whether or not narcissists are better than average, the potential threat is eliminated…”  What is the “potential threat” you refer to here?**

Thank you for bringing this confusing sentence to our attention. In the previous version of the paper, we were attempting to point out that if a narcissist actually is better than average on a particular characteristic then it would not be considered self-enhancement for a narcissist to perceive himself or herself more positively than he or she perceives the average person on that characteristic. This is one of the drawbacks associated with social comparison self-enhancement indices. We rewrote the paragraph to hopefully make this point more clearly, and in doing, so we removed the confusing reference to an unspecified “potential threat”.

The revised paper now reads as follows:

“The current meta-analytic review will exclusively focus on self-insight indices of self-enhancement because of the advantages outlined above, as well as practical issues related to the availability of primary studies that measure narcissism. In addition, self-enhancement as measured by social comparison has the undesirable property of lacking “a reality criterion against which the validity of the self-description can be evaluated” (Kurt & Paulhus, 2008, p. 840). In other words, social comparison measures could be particularly problematic when studying narcissists’ tendency to self-enhance, because it is possible that narcissists really are better than average (if a narcissist actually is better than average, then it would not be considered self-enhancement for that individual to perceive himself or herself more positively than he or she perceives the average person). For example, a recent meta-analysis established that narcissism is linked to physical attractiveness, (*r* = .15, *k* = 18, *N* = 1,039; Holtzman & Strube, 2010), so, assuming a normal distribution for physical attractiveness scores, it is potentially more likely that a narcissist is of slightly above average attractiveness. At the same time, the relationship between narcissism and attractiveness is relatively small, and the bulk of existing evidence suggests that narcissism is not strongly correlated with other positive criteria such as course grades (*r* = -.04, Farwell & Wohlwend-Lloyd, 1998), intelligence (*r* = .11, Gabriel et al., 1994), task performance in a group discussion exercise [*r* = 12 (staff ratings), *r* = .04 (peer ratings), John & Robins, 1994], leadership effectiveness [*r* = .02, *k* = 32, *N* = 5,593; Grijalva et al., 2015a), or creativity [*r* = .08 (fluency) and *r* = .06 (flexibility); Goncalo et al., 2010]. To avoid the pitfalls associated with not having a measure of reality against which to compare self-reports, in the current paper we consider external criteria (i.e., observer-reports and objective measures) to be essential—particularly when examining narcissism—because these external criteria constitute an “explicit standard” against which to establish the magnitude and direction of self-enhancement (Robins & Beer, 2001, p. 340).” (p. 8-9).

**Editor Comment 5.**

**On page 14, you write that you were primarily interested in social comparison measures, but I think you mean self-insight measures?**

You are correct, and we apologize for this mistake. In the specified sentence, we have now changed “social comparison measures” to “self-insight measures”.

**Editor Comment 6.**

**On page 23, you write that narcissists self-enhanced on each of the agentic constructs, without exception, after having just reported that they did not seem to self-enhance on task performance on the RAT.  This could use a bit of rephrasing.**

We have now rewritten this section. When we went back to update our results, we realized that for the study using the Remote Associate Task (RAT; i.e., Nunez, 2007) we had included effect sizes based on two different measures of narcissism: the Narcissistic Personality Inventory (NPI) and the NPDS. However, the NPDS is considered a measure of vulnerable narcissism (see Grijalva et al., 2015, Table 5 in *Psychological Bulletin*), and vulnerable narcissism measures were excluded from the current meta-analysis. Wink (1991) states that the NPDS indicates diminished self-esteem and depression, inadequacy, unhappiness, and worry. Further, the NPDS was developed empirically by contrasting clinical patients’ scores with those of individuals not in treatment.

When we excluded the narcissism-self-enhancement correlations based on the NPDS (and added an additional study from a different author), our results for task performance increased slightly from (*ρ* = .14, SE = .07) to our current result of (*ρ* = .17, SE = .05, 95% CI = [.03, .30]) and the standard error decreased. After making this correction, there was a statistically significant relationship between narcissism and self-enhancement in task performance.

Our choice to report the RAT separately in the previous version of the paper was primarily based on an inclusion error, and as pointed out by reviewer 2, was post hoc, so we have now removed this analysis and discussion from the current paper. [Note: this response is the same for Reviewer 2, Comment 10].

**As I mentioned above, both of the reviewers also listed several issues that merit your consideration.  Reviewer 1, for example, suggests a slight theoretical reframing, and recommends several alternative citations that you might include.  (Indeed, some of these suggested citations might contain relevant data sets for inclusion in your meta-analysis.)  Reviewer 2 makes several suggestions regarding additional analyses you might do, and raises some issues concerning the rationale for some of your analyses as well as your interpretations of some of your findings.  I read over all of the reviews carefully, and I believe that every one of them deserves your full consideration in a revision.  Please pay careful attention to all of these comments as you revise your manuscript.  If you do revise and resubmit your manuscript, please include a detailed cover letter that explains exactly how you responded to every one of the reviewers’ comments (or clearly explains why you chose not to address a particular point).**

Thank you for this summary. We attempted to addresses each of the reviewers concerns below. **I would like to receive the revision within 90 days from receipt of this letter (or 180 days with data collection).  If it is not possible for you to submit your revision by the due date, you will need to request an extension.  The revision should be submitted via SageTrack, once the revised manuscript is prepared.  Log into** [**http://mc.manuscriptcentral.com/pspb**](http://mc.manuscriptcentral.com/pspb) **and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions."  Under "Actions," you can click on "Create a Revision."  Your manuscript number has been appended to denote a revision.  Your original files are available to you when you upload your final manuscript.  Please delete any redundant files before completing the submission.  Please include in your cover letter a description of all of the revisions you have made.  If you have any questions during this process, please contact the editorial office at** [**Kristen.Marchetti@sagepub.com**](mailto:Kristen.Marchetti@sagepub.com)**.  
  
Of course, I cannot guarantee a favorable decision on a revised manuscript.  Given the overall enthusiasm for this topic, however, I am hopeful that a revision will be met with a more favorable response.  In the meantime, please let me know if you have any questions about the contents of this letter, my decision, or the reviews themselves.  I look forward to seeing a revision of your manuscript.  Thank you for considering PSPB as an outlet for your work.  
  
Sincerely,  
  
Dr. Jennifer Bosson  
Associate Editor  
Personality and Social Psychology Bulletin**

**REVIEWER(S)’ COMMENTS TO THE AUTHOR:**

**Reviewer 1**

**Reviewer 1, General Comments.**

**Comments to the Author  
This submission has a lot of merit. It addresses an important and timely issue, offers some interesting theorizing, and reports a competently conducted meta-analysis. I had comments that will hopefully be of some relevance to the authors (and the action editor).**

**GENERAL ISSUES  
  
Reviewer 1, Comment 1.**

**In the opening statement and in several spots all over the manuscript (and a couple of footnotes), the authors refer to pathological narcissism. I suggest removing all references to it. The readership of PSPB is interested in normal or everyday narcissism, not in pathological one. By now (well, since the 90s), the distinction between narcissism varying on a continuum (i.e., narcissism as a personality trait) and pathological narcissism has been well-established, and it is acceptable to just mention narcissism as a trait and taking form there. Also, removal of pathological narcissism will streamline the manuscript. (By the way, it should be “clinical”, not pathological, narcissism.)**

We thank you for this suggestion, which we believe helped us to more clearly frame our paper. We have now removed all references to clinical narcissism, pathological narcissism, and the DSM definition of narcissism. In addition, we rewrote several sections of the paper to highlight that we are focusing on narcissism as a personality trait. For a description of these changes, see our response to Editor Comment 1.

**Reviewer 1, Comment 2.**

**On p. 13, top, I wondered: If the self-criterion residual method is superior (to difference scores), why wouldn’t the author focus on it and delegate reporting of difference scores to a footnote? Then, reading on and reaching pp. 21-22, I realize that the authors have a theoretical reason for comparing effect sizes for the two methods. This theoretical rationale ought to be carried forward.**

We have now rewritten the section of the manuscript entitled “Difference Scores versus the Self-Criterion Residual Method” in a way that we hope clarifies the theoretical rational for this analysis. Although difference scores are oft criticized, we were interested in investigating how similar the effect sizes produced using the self-criterion residual method were to those produced using difference scores.

“Difference Scores versus the Self-Criterion Residual Method

We will also investigate a methodological moderator that will allow us to compare and contrast different methods of measuring self-enhancement. To calculate self-enhancement, researchers tend to use one of two methods: (1) difference scores that are calculated by subtracting external-ratings from self-ratings or (2) the self-criterion residual method (John & Robins, 1994; Paulhus & John, 1998). The self-criterion residual is calculated by regressing self-reports onto an external criterion (e.g., others’ perceptions). The resulting residual reflects the degree of self-other bias present because all of the shared variance has been removed—effectively making the residuals an estimate of self-enhancement (or in some cases self-effacement, if the residual is negative, meaning that an individual’s self-report was lower than the observer-report). These residuals (also known as bias scores) can be correlated with narcissism scores, or any other variable, to calculate its relationship with self-enhancement.

Of these two approaches, difference scores have been criticized more frequently for their methodological weaknesses (Cronbach, 1958, 1992; Cronbach & Furby, 1970; Edwards, 1994; Edwards, 1995; Furr & Bacharach, 2013 Johns, 1981; Zuckerman & Knee, 1996). These criticisms have been discussed at length elsewhere, but a key concern (among others) is that the difference score is less reliable than either of its two components, when the individual components are correlated, as will most likely be the case when comparing self-reports and observer-reports. Because of methodological problems, difference scores have been accused of producing “ambiguous and potentially misleading” results (Edwards, 1995, p. 307). Given the criticisms of difference scores, the self-criterion residual method (John & Robins, 1994; Paulhus & John, 1998) has become the preferred approach to calculating self-enhancement bias. At the same time, it remains unclear how much, on average, the results produced using the self-criterion residual method will differ from those produced using difference scores—scholars tend to report their results using only one of the two methods. Therefore, the current meta-analysis will examine the effect of the two different methods of calculating self-enhancement bias on the magnitude and direction of the narcissism-self-enhancement relationship. We will be attempting to address the question, how similar (or different) are effect sizes produced using the self-criterion residual method, as compared to those produced using difference scores? Thus,

*Research Question 1*: Do meta-analytic effect sizes based on the self-criterion residual method differ from those using the difference score method?”

(p. 15-16).

**Reviewer 1, Comment 3.**

**p. 23, middle. The authors report “the narcissism-self-enhancement relationship for other agentic constructs …” Should p < (rather than p >) for all traits (intelligence, leadership, extraversion, attractiveness, and openness)?**

Thank you for catching this error. We have now fixed the mistake, and carefully double checked our Results section to make sure that all of our analyses were reported correctly.

**MINOR ISSUES  
  
Reviewer 1, Comment 4.**

**p. 3. In the second paragraph, the authors present the term ego-involving contexts without defining it. In the next sentence, they make a definitional attempt, but it’s roundabout. I suggest defining this term in a more direct manner.**

We removed the aforementioned sentence because the term “ego-involving contexts” was from a quote and was not well-defined.

We have now replaced the aforementioned sentence with the following sentence on page 3:

“In general, a person predominantly self-enhances characteristics that are most central to his or her self-concept (Gaertner, Sedikides, & Chang, 2008; Gebauer, Sedikides, Verplanken, & Maio, 2012; Sedikides, Gaertner, & Toguchi, 2003).”

**Reviewer 1, Comment 5.**

**p. 4, first half: “b” is unclear. It only becomes clear with the last sentence of p. 11.**

Previously, the sentence read, “(b) examining whether the length of acquaintanceship affects observer-ratings and how this may magnify or diminish self-enhancement.”

We have now rewritten this sentence to more clearly explicate the role that we think length of acquaintanceship will play in observer-ratings of narcissists.

“(b) examining the role played by length of acquaintanceship, whether observers who have known a narcissist longer perceive them more negatively; a tendency that could result in a greater discrepancy between narcissists’ self-reports and observer-reports (i.e., greater observed self-enhancement) for well-acquainted as opposed to minimally acquainted participants” (p. 4). **Reviewer 1, Comment 6.**

**p. 5, first few lines. The choice of citations is somewhat unusual.  
\*I am not sure the three cites (Campbell, Reeder, Sedikides, & Elliot, 2000; Gramzow & Willard, 2006; John & Robins, 1994) are the most relevant to the point of a “general tendency for people to have inflated vies of themselves.” Shouldn’t the authors refer to reviews/books rather than primary sources (which are a bit old)? Alternative citations are: Alicke & Sedikides, 2009 (European Review of Social Psychology), Alicke & Sedikides (2011; Handbook of Self-Enhancement and Self-Protection), Sedikides & Gregg, 2008 (Perspectives on Psychological Science).**

Thank you for pointing us towards this relevant literature. We agree that the recommended citations are a much better choice for a general tendency to have inflated self-views. We have now replaced the previous citations with those recommended above, (i.e., “Alicke & Sedikides, 2009; Alicke & Sedikides, 2011; Sedikides & Gregg, 2008”) (p. 5).

**Reviewer 1, Comment 7.**

**\*For the point that people “endorse self-serving attributions”, I’d recommend replacing the two cites (of which the Blaine & Crocker is particularly peculiar) with: Mezulis, Abramson, Hyde, & Hankin, 2004 (Psychological Bulletin) and Campbell & Sedikides (2009, Review of General Psychology).**

Again, thank you for recommending more appropriate citations. We have now replaced our previous citations with Campbell and Sedikides (1999) and Mezulius, Abramson, Hyde, and Hankin (2004) (p. 6). Please note that the date for the Campbell and Sedikides paper we cite is 1999 (not 2009), but that we believe this is the paper referenced by Reviewer 1 because it is the only Review of General Psychology paper related to the topic of self-enhancement by Campbell and Sedikides.

**Reviewer 1, Comment 8.**

**\*for the BTAE, I suggest replacing the Krueger & Mueller cite with: Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), The self in social judgment (pp. 85–106). Philadelphia, PA: Psychology Press.**

We have now updated the citation for the better-than-average effect to include Alicke and Govorun (2005), and removed Krueger and Mueller (2002) (p. 6).  
 **Reviewer 1, Comment 9.**

**p. 5, last par. The author will do well to replace most, if no all, of the primary sources a couple of recent reviews. See the Rhodewalt chapter in Handbook of Self and Identity (2013) and the Morf et al. chapter in Handbook of Self-Enhancement and Self-Protection (2011).**

Thank you for recommending both the Rhodewalt (2011) and the Morf, Horvath, & Torchetti (2011) chapters.

In the first paragraph of page 4, we have now replaced most of the citations for narcissism being a defensive, vulnerable type of self-esteem with citations for the two aforementioned chapters: Rhodewalt (2011) and Morf et al. (2011).

In addition, we found Morf et al.’s (2011) review of the narcissism and self-enhancement literature to be particularly helpful. We also now cite this chapter in several other places throughout our manuscript.

**Reviewer 1, Comment 10.**

**p. 11, middle. When the authors claim that “narcissists make positive first impressions that deteriorate as people get to know them better”, they could also consider citing Back, Schmukle, and Egloff (2010, JPSP).**

When discussing how narcissists make positive first impressions that deteriorate over time, we now also cite Back, Schmukle, and Egloff (2010) (p. 14).

**Reviewer 1, Comment 11.**

**p. 24, first part. Interesting that it is important for narcissists to be likable. Reminds one of the quintessential narcissism, Bill Clinton, and his drive (for lack of better word) to be likable.**

Thank you, we found this interesting too!

**Reviewer 2**

**Reviewer 2, General Comment.**

**Comments to the Author  
This is an excellent paper meta-analyzing the association between narcissism and self-enhancement, and testing several moderators.  I am not an expert on meta-analysis so cannot evaluate the statistics, but I have several suggestions for the authors to consider.  In my view, a slightly improved version of this paper should be published in PSPB.**

Thank you for these comments.

**Reviewer 2, Comment 1.**

**I wonder if the authors could find some more samples to include.  Did the authors do a forward search of all articles they found (i.e., looked for more recent papers that cited those papers)?  There are some references that the authors are missing that make me wonder if their literature search may have missed some results.  In addition, I think it would be worth looking through the clinical psychology literature (excluding clinical samples, but some of that literature is based on non-clinical samples) and including other measures of narcissism that are more often used by clinicians (though be careful to avoid measures that also tap into 'vulnerable narcissism' which, in my view, is not narcissism and will muddy the picture).  I can't think of the names of the scales that clinical psychologists use off the top of my head, but a quick search should turn them up.**

These were good suggestions. We worked at length to perform additional searches, and have now added 10 samples.

For a thorough description of our updated search, please see Editor Comment 3.

**Reviewer 2, Comment 2.**

**I think the authors need to address the question of publication bias.  If possible, I would like to see a p-curve analysis, and some estimate of the effect size corrected for publication bias.  My guess is there will not be much publication bias for the communal constructs, but there could be some for the agentic constructs, because those are likely the focal constructs in most published studies.**

We have now added three different publication bias analyses: funnel plots, Duval and Tweedie’s (2000) trim and fill analysis (which provides estimates of effect size corrected for publication bias), and *p-curve* analyses.

**“Publication Bias**

Publication bias was investigated for the narcissism-overall self-enhancement relationship, as well as separately for self-enhancement in agentic and communal criteria. First, as reported above, we compared published to unpublished studies—the concern being that studies with larger effect sizes might be more likely to get published. This was not the case in the current paper because we found that the effect sizes did not differ between published and unpublished studies (see results for ‘Publication Type’ in Tables 3 through 6). Second, we examined funnel plots with standard error on the vertical axis and the effect size on the horizontal axis. In general, studies with a larger number of participants are expected to cluster toward the top of the graph and near the mean, whereas studies with a smaller number of participants are expected to be located at the bottom of the graph and exhibit greater variability. Publication bias is indicated by a lack of symmetry about the mean with smaller studies tending to have larger than average effect sizes (i.e., to be located on the right side of the graph). This would occur because statistically significant studies are more likely to be published, and the concern is that some studies with smaller effect sizes (i.e., studies on the left side of the graph) are missing from the meta-analysis. However, based on the funnel plots in the current study, there does not appear to be a large amount of publication bias for overall self-enhancement (Figure 1), self-enhancement in agentic criteria (Figure 2), or for self-enhancement in communal criteria (Figure 3). Third, to quantify the magnitude of any potential publication bias we also performed Duval and Tweedie’s (2000) trim and fill analysis that is designed to detect where missing studies are likely to be located and then impute these missing studies in order to compute a less biased, “adjusted” meta-analytic effect. The observed point estimate for overall self-enhancement was .21 (95% CI = [.16, .26]), whereas the adjusted value was .17 (95% CI = [.12, .23]); for agentic effect sizes the observed point estimate was .29 (95% CI = [.24, .34]), whereas the adjusted value was .25 (95% CI = [.19, .30]); for communal effect sizes the observed point estimate was .004 (95% CI = [-.08, .08]), whereas the adjusted value was .004 (95% CI = [-.08, .08]). Although for overall self-enhancement and self-enhancement in agentic criteria the adjusted effect sizes decreased slightly, none of the differences between observed and adjusted effect sizes reached statistical significance. Finally, we conducted *p-curve* analyses. A “*p-curve* is the distribution of statistically significant *p* values for a set of independent findings” (Simonsohn et al., 2014, p.535). If the *p-curve* distribution is right-skewed, this is an indication that publication bias is less likely because a larger number of the *p* values are near .01 rather than the high .04s. For overall self-enhancement, 13 of 36 effect sizes were excluded from the *p-curve* analysis because they were not statistically significant, leaving 23 effect sizes to include in the analysis. The results are summarized in Figure 4 where one can see that the distribution is right-skewed; 74% of effect sizes had *p* values smaller or equal to 0.01, and 91% no larger than 0.02. The curve is significantly right-skewed based on both the binomial test (which tests the share of significant results for *p* values less than .025; *p* < 0.0001) and the continuous test (Z = -12.78, *p* < .0001). Next, for self-enhancement in agentic criteria, 26 effect sizes were included in the analysis. Table 5 shows that this distribution is also right-skewed with 85% of the effect sizes having *p* values smaller or equal to 0.01, and 97% no larger than 0.02. The curve is significantly right-skewed based on both the binomial test (p< 0.0001) and the continuous test (Z = -14.73, *p* < .0001). We could not perform this analysis for communal self-enhancement because only 3 effect sizes were significant, but based on our previous publication bias analyses, it seems as though the threat is very minimal for communal criteria. In sum, we can conclude that publication bias is likely not a great threat to the validity of the current study.” (p. 26-28).

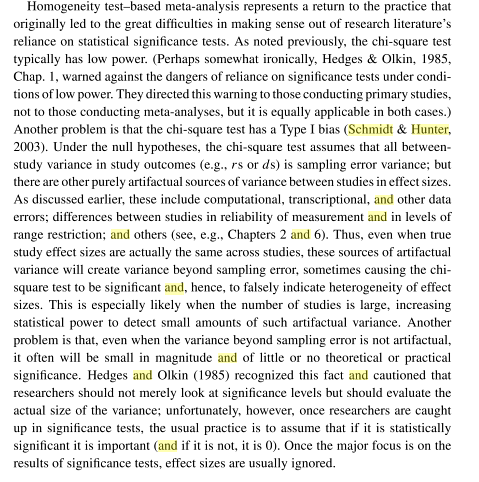
**Reviewer 2, Comment 3.**

**I am not an expert on meta-analysis, but I thought that, before examining potential moderators, it is typical to estimate the amount of heterogeneity in effect sizes.  I think it would be useful to present such an analysis.  Then, it would also be useful to know how much of that variance is accounted for by each moderator, and how much heterogeneity there is within the subgroups for each moderator (e.g., within the agentic effect sizes and the communal effect sizes, etc.) for the categorical moderators.**

This was a great suggestion. Because we conducted our meta-analysis using relatively new multilevel meta-analytic techniques, we were unaware of an accepted method of reporting the amount of heterogeneity in effect sizes in a similar way to those of more traditional meta-analytic methods. Thus, we also calculated our most important findings using the Hunter and Schmidt (2004) approach to meta-analysis, which allowed us to report credibility intervals in Footnote 4. Notably, whereas confidence intervals are dependent on sample size and are a reflection of sampling error, credibility intervals are an indication of true dispersion across studies due to moderators. The credibility interval for each of these analyses were relatively wide suggesting that moderator variables were present.

“Although the multilevel approach to meta-analysis has several advantages, it is relatively new, thus we also report the results for our main findings using Hunter and Schmidt’s (2004) procedure. The overall relationship between narcissism and self-enhancement was .21 and corrected for unreliability in the predictor was .24 (*k* = 36, *N* = 6,844, 95% CI = [.19, .29], 80% credibility interval = [.04, .44]). For effect sizes with agentic criteria, the relationship between narcissism and self-enhancement was .27 and corrected for unreliability in the predictor was .30 (*k* = 28, *N* = 5,754, 95% CI = [.25, .35], 80% credibility interval = [.14, .46]), whereas for effect sizes with communal criteria relationship between narcissism and self-enhancement was .01 and corrected for unreliability in the predictor was .01 (*k* = 11, *N* = 1,645, 95% CI = [-.09, .11], 80% credibility interval = [-.13, .16]). Thus, the Hunter and Schmidt (2004) approach to calculating meta-analytic effect sizes produced results that did not significantly differ from the multilevel approach. Notably, whereas confidence intervals are dependent on sample size and are a reflection of sampling error, credibility intervals are an indication of true dispersion across studies due to moderators. The credibility interval for each of these analyses was relatively wide suggesting that moderator variables were present.

In addition, we want to clarify why we chose to report credibility intervals rather than the homogeneity statistics preferred by proponents of the Hedges and Olkin’s (1985) approach to meta-analysis. First, these homogeneity tests are a method of significance testing for moderator analyses, and thus are subject to all of the same criticisms as significance testing. Further, they rely on chi-square statistics, which are sometimes subject to a Type I bias. Specifically, Schmidt and Hunter (2004) state that,



(p. 460).

To address the second part of Reviewer 2, Comment 3 (i.e., a request that we estimate how much variance is accounted for by each moderator), we added pseudo-*R2* values to the relevant Tables. Pseudo-*R2* values are an estimate of variance explained in the context of multilevel modeling [i.e., pseudo-*R2* = 1 – (residual variance with predictors)/(residual variance without predictors)].

Further, as recommended, we reported our moderator analyses and pseudo-*R2* values for overall self-enhancement (Table 3), self-enhancement in agentic criteria (Table 5), and self-enhancement in communal criteria (Table 6).

References:

Hunter, J.E. & Schmidt, F.L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.

**Reviewer 2, Comment 4.**

**For research question 1, you need a better criterion to determine whether the effect sizes for studies using residuals are more valid than effect sizes for studies using difference scores.  Just because the residual studies have larger effects doesn't mean they are more valid.  A better test might be whether there is less heterogeneity in the studies using residuals than difference scores.  If you could show that there was less heterogeneity in results based on residuals (controlling for other moderators that might be confounds), that would be more convincing. Another approach would be to just describe the difference between the two kinds of studies without arguing that the larger effects for residuals is evidence that they are more valid. (Also, difference scores have some of the same undesirable properties as difference scores, but fewer of them, it seems. Still, they are not perfect and I think it would be important to point this out when discussing them on pp. 27-28).**

In our discussion section, we now just describe our results for difference scores compared to regression residuals without arguing that the larger effects for residuals are evidence that they are more valid.

Further, we now state that regression residuals have also faced methodological criticism, and spend more time introducing alternative approaches to calculating self-enhancement.

On pages 34-35 we now say,

“A third contribution of the current meta-analysis was to investigate the impact of using difference scores, as opposed to regression residuals. We found that there was not a statistically significant difference in the magnitude of the effect sizes produced using the two different approaches to calculating self-enhancement (see Table 3). Although difference scores are criticized more frequently, the self-criterion residual method has also faced methodological criticism [see Krueger and Wright (2011)]. It is notable that the two approaches did not diverge more markedly. That being said, there are now more advanced methods than difference scores and the self-criterion residual method for indexing self-enhancement. First, drawing on Kenny’s (1994) Social Relations Model (SRM), a new method was proposed by Kwan and colleagues (2004) that requires round-robin data (i.e., data collected in a small group in which all group members provide self-reports, as well as reports for every other group member). If round-robin data are available, then using Kwan and colleagues (2004) method allows one to get a more precise estimate of self-enhancement by taking into consideration both perceiver effects (how one tends to perceive others) and target effects (how one tends to be perceived; for a more detailed description see Kwan et al., 2004 and Borkenau, Zaltauskas, & Leising, 2009). A second option is to use polynomial regression and response surface methodology (Edwards & Parry, 1992), which have become popular in the person-environment fit literature in industrial and organizational psychology (and do not require round-robin data). Using polynomial regression, researchers can analyze their results from a three-dimensional perspective, relating self-ratings and observer-ratings to an outcome of interest, such as narcissism. Polynomial regression and response surface methodology allow researchers to directly test the relationships that difference scores are supposed to evaluate without the same restrictive (often untested) assumptions inherent to the use of difference scores (see Edwards 2002 for a thorough description of polynomial regression).”

**Reviewer 2, Comment 5.**

**For the last analysis reported on page 19, it seems to me that it would be better to leave the constructs in the 'neither' category out of the analysis altogether.  With them in there (coded as '0'), it is possible that they are partly responsible for the effect.  i also think the analysis just before that one is unnecessary (showing that there is no difference between the 'neither' constructs and all other constructs) - if it is important, the rationale for this analysis needs to be made more explicit.**

Thank you for this helpful feedback. In retrospect, it doesn’t make sense to include the neither category in our analyses because criteria that are neither agentic nor communal (or both) are not of theoretical interest in the current paper. We have now run our analyses with the agentic effect sizes (see Table 5) and communal effect sizes (see Table 6) in different files, thus excluding the ‘neither’ category. In addition, this simplification meant that we no longer needed to perform the contrast coding analysis referred to in the second part of your comment. Thus, the result showing that there is no difference between the ‘neither’ constructs and all other constructs has now been removed from the paper.

**Reviewer 2, Comment 6.**

**Was there enough detail provided in the studies to use a more fine-grained measure of acquaintance in the moderator analyses?  If a continuous variable is not possible, perhaps the studies/effect sizes can be grouped into more than two categories?  I do think the current analysis (very low acquaintance vs. everyone else) is very interesting and if it's the best information that's available, then the authors should still retain this analysis, but if it would be possible to also compare medium vs. long acquaintance, for example, that would be useful too.**

We have now used a more fine-grained measure of acquaintance in our moderator analyses. We first tried to use a continuous measure of mean time acquainted, but there was too much missing data to use this approach. Instead, we created three categories: zero acquaintance, short acquaintance, and long acquaintance.

“Further, we coded the length of raters’ relationships with the focal participant using three categories: (1) *zero acquaintanceship*—the rater and target were strangers who had not interacted [e.g., participants’ personalities were rated by strangers based on their Facebook profiles (Carlson et al., 2011a); participants’ physical attractiveness was rated by strangers based on photographs (Gabriel et al., 1994)], (2) *short acquaintanceship*—when the rater and participant had interacted, but known each other for less than one week [e.g., without prior interaction, individuals participated in a leaderless group discussion exercise and then rated each group member’s task performance (Robins & Beer, 2001); without prior interaction, pairs of participants talked for five minutes and then rated their partner’s personality (Carlson et al., 2011b)], and (3) *long acquaintanceship*—when the rater and participant had known each other longer than one week (e.g., friend, family member, and coworker ratings).” (p. 20)

However, despite using a more fine-grained measure of acquaintance, our Hypothesis 3 was still not supported (see results below).

“Next, we assessed the impact of length of acquaintance on the narcissism-self-enhancement relationship. It was predicted that individuals who had known a narcissist for a shorter period of time would see the narcissist more positively, resulting in a smaller discrepancy between their observer-reports and narcissists’ self-reports, as compared to more knowledgeable observers. For this analysis, we identified three broad categories of acquaintance for which we ran separate regression analyses: zero acquaintance (corrected *B* = .24, *SE* = .03, *k* = 39, number of samples = 6, 95% CI = [.18, .30], short acquaintance (corrected *B* = .15, *SE* = .05, *k* = 30, number of samples = 7, 95% CI = [-.001, .30]), and long acquaintance (corrected *B* = .23, *SE* = .03, *k* = 53, number of samples = 17, 95% CI = [.16, .31]). Hypothesis 3 was not supported, as evidenced by the overlapping confidence intervals for the three categories of acquaintanceship. The self-enhancement bias was not larger in magnitude when based on ratings from family/friends/coworkers who were in the long acquaintance category, as compared to newer acquaintances or strangers (i.e., short acquaintance or zero acquaintance). Interestingly, short acquaintance was the only category that was not significantly different from zero, whereas both long acquaintance and zero acquaintance exhibited positive relationships. ” (p. 28)

In our limitations section on page 36, we now say,

“In addition, although it is logically intuitive that peoples’ perceptions of narcissists would change as they get to know a narcissist better, because it takes time for narcissists more negative qualities to become apparent, our acquaintance hypothesis (Hypothesis 3) was not supported. We would have liked to use a continuous measure of length of acquaintance rather than three categories (zero, short, and long acquaintance), but this was not possible because of missing information for many studies and the fact that many studies combined information from dyads with different lengths of acquaintance. Due to the somewhat crude nature of the acquaintance categories in the current work, we hope that researchers will continue to address the role played by length of acquaintance.”

**Reviewer 2, Comment 7.**

**It would be useful if the authors would give confidence intervals and exact p-values for all results, including non-significant results.**

We have now included confidence intervals and exact p-values in our relevant Tables. In our results section, we only report the confidence intervals (and not the exact *p* values) because confidence intervals simultaneously summarize information about significance (whether or not the intervals include zero) and the sampling error associated with a point estimate. In addition, we recognize that the field is trying to move away from Null Hypothesis Significance Testing (NHST) logic, which is why we chose to focus on confidence intervals. Of course, interested readers can find the exact *p* values in our Tables, and if you would prefer, we can also add the exact p-values to our Results section.

**Reviewer 2, Comment 8.**

**It is very interesting that the type of criterion (observer-report vs. objective measure) did not moderate the magnitude of the effect.  I think the authors should discuss this more.**

We have now added a discussion of the type of criterion (observer-reports vs. objective measures) to our discussion section.

“Although not related to a specific a priori hypothesis, the current study also indicated that narcissists self-enhanced to similar degrees when self-enhancement was measured using different sources of comparison—observer-reports as opposed to objective measures (e.g., intelligence tests). Dufner et al. (2013) found that observer-reports (which they referred to as perceived self-enhancement) and objective measures (which they referred to as actual self-enhancement) had overlapping nomological networks, but were only weakly related (*r* = .11). Thus, just because someone self-enhances on objective measures does not mean they will necessarily self-enhance according to observer-reports, and vice versa. Our meta-analytic results however showed that narcissists self-enhanced regardless of whether researchers used objective measures or observer-reports. This was particularly interesting because Dufner and colleagues (2013) further found that these two ways of measuring self-enhancement were associated with different social impressions. There was a positive relationship between objective self-enhancement and perceptions of emotional stability, whereas there was an inverted u-shaped relationship between observer-reported self-enhancement and emotional stability such that people with moderate levels of self-enhancement were perceived to be more emotionally stable than people with very high or low levels of observer-reported self-enhancement.” (p. 32-33)

**Reviewer 2, Comment 9.**

**In the competing moderators model, why include moderators that were not significant on their own?**

This is a good point. In retrospect, the competing moderators model appears to have been unnecessary, so we have now removed this analysis from our paper.

**Reviewer 2, Comment 10.**

**Excluding the remote associates task from the task performance analyses seems very post-hoc.  Please explain the rationale for considering it a different type of task, and please also explicitly say that there is greater chance of false positive since the decision to split the task performance results was based on examining the raw data (effect sizes) after you found a non-significant effect.  (For example, the overall effect size for conscientiousness was very similar to the overall effect size for task performance, but you didn't go and look at the individual studies to try to figure out if there are a subset of studies that did find large effects for conscientiousness, because this result was consistent with your hypothesis.  Related to this, I would encourage you not to rely too heavily on statistical significance, and consider than an effect size of .13 or .14 might actually be meaningful.)**

When we went back to update our results, we realized that for the study using the Remote Associate Task (RAT; i.e., Nunez, 2007) we had included effect sizes based on two different measures of narcissism: the Narcissistic Personality Inventory (NPI) and the NPDS. However, the NPDS is considered a measure of vulnerable narcissism (see Grijalva et al., 2015, Table 5 in *Psychological Bulletin*). Wink (1991) states that the NPDS indicates diminished self-esteem and depression, inadequacy, unhappiness, and worry. Further, the NPDS was developed empirically by contrasting clinical patients’ scores with those of individuals not in treatment. When we excluded the narcissism-self-enhancement correlations based on the NPDS (and added an additional study from a different author), our results for task performance increased slightly from (*ρ* = .14, SE = .07) to our current result of (*ρ* = .17, SE = .05, 95% CI = [.03, .30]) and the standard error decreased. After making this correction, there was a statistically significant relationship between narcissism and self-enhancement in task performance.

As you pointed out, our choice to report the RAT separately was post hoc (and based on an inclusion error), so we have removed this analysis and discussion from the current paper.

**Reviewer 2, Comment 11.**

**When listing all agentic constructs (p. 23), you report all p-values as greater than .05, but then state in the next sentence that all effects were significant. reporting exact p-values and confidence intervals will help eliminate this kind of error.**

Thank you for catching this. We have now reported exact *p* values and confidence intervals for all of our effect sizes (see Reviewer 2, Comment 7 above for more information).

**Reviewer 2, Comment 12.**

**In the discussion, I don't think you can conclude that narcissists' self-reports of communal constructs are likely to be accurate.  Just because narcissism is not correlated with self-enhancement for agreeableness and conscientiousness doesn't mean that narcissists are accurate.  For example, it could be that all self-reports of agreeableness are inaccurate.  In other words, there are other ways to be inaccurate besides self-enhancing.  A more correct statement would be that narcissists are not more likely to be inaccurate on these constructs than non-narcissists. (See also p. 26 where you again refer to narcissists' self-reports on communal traits as 'accurate'.)**

We have not removed all of our previous references to accuracy, and instead state the following:

In the Introduction:

“Perhaps more surprising than the finding that narcissists endorse agentic characteristics, is the fact that individuals high in narcissism possess a much less discrepant idea of how others perceive their communal traits compared to their agentic traits (Carlson, Naumann, & Vazire, 2011a; Carlson et al., 2011b).” (p. 12).

In the Discussion:

“From a practical perspective, this means that researchers should potentially be more cautious about relying on a narcissists’ self-reports of agentic qualities, as well as likeability and conscientiousness, because these are the characteristics narcissists are more likely to positively distort. In contrast, narcissists are not more likely to provide inaccurate self-ratings on communal constructs than non-narcissists.” (p. 30-31).

**MINOR ISSUES  
  
Reviewer 2, Comment 13.**

**In the introduction, Kurt & Paulhus (2008) is very relevant.  It might also meet the inclusion criteria for your meta-analysis.**

Thank you for pointing us toward this relevant article.

We have now added a description of Kurt and Paulhus (1998) to the introduction section entitled “Two types of self-enhancement bias” (p. 7).

“Further, in one of the few studies that collected both social comparison and self-insight information from the same participants, Kurt and Paulhus (2008) reported that when they controlled for self-reported personality (i.e., the Big Five), their social comparison index was no longer related to self-rated or peer-rated adjustment outcomes. However, self-insight measures did explain incremental variance in both of these indicators of psychological adjustment. Thus, leading Kurt and Paulhus (2008) to conclude that the self-insight index is “a more defensible operationalization of self-enhancement than is the social comparison index” and that it predicts poorer interpersonal adjustment, particularly when these interpersonal adjustment ratings are obtained from peer-reports (p. 848).”

Unfortunately, Kurt and Paulhus (1998) did not meet our inclusion criteria because, although the paper was focused on self-enhancement, it did not measure narcissism. However, we were able to add Kurt’s (2004) dissertation to our updated meta-analysis.

**Reviewer 2, Comment 14.**

**In the introduction (especially pages 9-10) and the discussion, Carlson (2013; Self and Identity) is very relevant and directly addresses the question of whether narcissists value communal traits.**

We agree that Carlson (2013) is very relevant, and we have now incorporated its findings into our introduction section where we discuss whether or not narcissists value communal traits (p. 12-13).

“This idea, however, that narcissists are avoiding vulnerability by eschewing communal traits is difficult to test. Instead, research is accumulating that the tendency to emphasize agentic traits and simultaneously deemphasize communal traits might be a conscious life strategy employed by narcissists that is focused on maximizing personal gain. Interestingly, narcissists appear to know full well that this personal gain often occurs at others’ expense (Jonason, Li, & Teicher, 2010). Carlson (2013) found that narcissists are not only aware that they are narcissistic (i.e., they admit to bragging and acting condescending), but with surprising self-insight, individuals high in narcissism also acknowledged that narcissism produced positive consequences for themselves that were accompanied by a fairly negative impact on others. A negative impact these narcissists appeared to find acceptable, as they also reported a desire to become more narcissistic in the future (Carlson, 2013).

We now also refer to this paper in the discussion section on page 33-34.

“If narcissists really considered communal traits to reflect negative qualities, then we would have expected to see statistically significant negative effect sizes. Based on our results, narcissists may perceive communal qualities as simply being unimportant. This is consistent with Campbell and Foster’s (2007) observation that one of the defining features of narcissism is a “*lack of interest* in warm and caring interpersonal relationships” (p. 118, emphasis added). Recall, that recent research suggests narcissists are aware of the social costs of narcissism, but that they believe the personal rewards associated with narcissism outweigh any social costs (Carlson, 2013).”

**Reviewer 2, Comment 15.**

**I think there is a lot more literature on the concept of narcissism as 'unmitigated agency' than the authors discuss.**

We tried to address this concern in several ways.

First, we have now added a few sentences after our definitions of agency and communion in which we highlight narcissism’s connection with the interpersonal circumplex and unmitigated agency (p. 10-11)

“In a seminal work clarifying the boundaries between these two concepts, Wiggins (1991) referred to agency as “the condition of being a differentiated individual, and it is manifest in strivings for mastery and power which enhance and protect that differentiation”, whereas communion was defined as “the condition of being part of a larger social or spiritual entity, and it is manifested in strivings for intimacy, union, and solidarity with that larger entity” (p. 89; see also Bakan, 1966). Within this framework, narcissism is a vector positioned between the high-agency and low-communion axes (Paulhus, 2001; Paulhus & Williams, 2002; Wiggins & Pincus, 1994). In effect, narcissism has a strong positive correlation with the agency axis of the interpersonal circumplex (*r* = .84; Bradlee & Emmons, 1992), but does not have a strong correlation with the communion axis of the interpersonal circumplex (*r* = .08; Bradlee & Emmons, 1992). Thus, narcissism falls within the interpersonal circumplex quadrant labeled unmitigated agency (Buss, 1990; Helgeson & Fritz, 2000). Unmitigated agency is characterized by ‘‘a focus on the self to the exclusion of others [which] . . . includes being hostile, cynical, greedy, and arrogant’’ (Helgeson & Fritz, 1999, p. 132; see also Rauthmann & Kolar, 2013).”

In an attempt to summarize the agency and communion literature more thoroughly, we have now folded into our original synopsis a description of several additional, and more recent studies that show narcissists value agency over communion, including Findley and Ojanen (2013), Gu, He, and Zhao (2013), and Jones and Brunell (2015).

“As mentioned above, scholars have established that a person predominantly self-enhances attributes that are most central to his or her self-concept (Gaertner et al., 2008; Gebauer et al., 2012; Sedikides et al., 2003), therefore it corresponds that narcissists’ positive illusions would give priority to agentic characteristics based on agency’s alignment with “self-seeking, egocentric motives” (Wiggins, 1991, p. 91). In support of this theoretical rationale, scholars have recently compiled a variety of diverse evidence demonstrating that narcissists value agentic outcomes and attributes more than communal outcomes and attributes. To illustrate, narcissism has been associated with agentic goals (e.g., power and status), but not communal goals (e.g., affiliation and closeness; Findley & Ojanen, 2013), and in a lab study, Besser and Priel (2010) found that narcissistic participants reported being angry after reading a hypothetical vignette about an achievement threat, but did not report being angry after reading a hypothetical vignette about a romantic rejection threat. Finally, in a daily diary study, narcissists’ state self-esteem was decreased by negative achievement events, but was immune to both positive and negative social events that the authors considered to be indicators of communion (Zeigler-Hill, Myers, & Clark, 2010).

Additional research even points to the fact that Narcissists’ preference for agency over communion is apparent at an implicit, unconscious level. Gu, He, and Zhao (2013) discovered that narcissists exhibited attentional biases for performance words such that “they were highly vigilant to failure words and had difficulty disengaging from success words” but that they were not affected by interpersonal words. Similarly, in a surprise recall task, narcissists were more likely to recall agentic trait descriptors than communal trait descriptors, suggesting that narcissism affects information processing in such a way that narcissists are more likely to remember agentic trait content because it is more self-relevant (Jones & Brunell, 2015). Results like these led Paulhus (2001) to propose that narcissism is an extreme form of agency, and more recently, Campbell and colleagues introduced an agency model of narcissism (Campbell, Brunell, & Finkel, 2006; Campbell & Foster, 2007). It appears that agency, but not communion, is consistent with narcissists’ grandiose conception of success.” (p. 11-12).

**Reviewer 2, Comment 16.**

**p. 11 - Naumann et al. (2008) is relevant to accuracy of thin slice judgments of narcissism.**

We have now added a description of Vazire, Naumann, Rentfrow, and Gosling (2008) to our introduction where we discuss the thin slices of behavior paradigm.

“Based on the thin slices of behavior paradigm, after a mere 30 seconds of exposure, participants identified narcissists as being extraverted and likeable (Oltmanns, Friedman, Fiedler, & Turkheimer, 2004). It appears that individuals are able to make snap judgments about whether or not someone is high in narcissism based on a variety of visual cues. For example, observers associate narcissism with wearing stylish and expensive clothing, having a neat and organized appearance that presumably took a long time to put together, being physically attractive, and for women, wearing makeup and having visible cleavage (Vazire, Naumann, Rentfrow, & Gosling, 2008).” (p. x14

**Reviewer 2, Comment 17.**

**p. 12 - When describing hypothesis 3, change 'when based on ratings from close others' to 'when the criterion measure is based on ratings from close others**

We have now made this change in two locations.

First, on page 15, Hypothesis 3 now reads

“Hypothesis 3: Narcissists’ self-enhancement bias will be larger in magnitude when the criterion measure is based on ratings from close others than when based on ratings from new acquaintances.”

Second, we also corrected the wording when we describe our results for the acquaintanceship moderator, on page 28.

“The self-enhancement bias was not larger in magnitude when the criterion measure was based on ratings from family/friends/coworkers who were in the long acquaintance category, as compared to newer acquaintances or strangers (i.e., short acquaintance or zero acquaintance).”

**Reviewer 2, Comment 18.**

**p. 17 - 'between the study moderator variables' should be 'among'**

We have now changed “between the study moderator variables” to “among the study moderator variables” (p. 22).

**Reviewer 2, Comment 19.**

**p. 17 - 'Many of the relationships in Table 1' should be 'Table 2'**

We apologize for this confusing labeling error. We have now changed this to Table 2 (p. 22).

**Reviewer 2, Comment 20.**

**p. 17 - 'percentage of total variance between effect sizes' should be 'percentage of total variance in effect sizes'**

We have now changed “between” to “in” (p. 22).

**Reviewer 2, Comment 21.**

**p. 17 - I'm a little confused about the ICC analyses - isn't MLM necessary if the proportion of variance within samples is high enough (not between samples)?  either way, whether it's 27% or 63%, I think MLM is warranted, but I think justifying the need for MLM by saying that there was 'sufficient between samples variance' is confusing. (Also, be consistent about 'between samples' vs. 'between studies')**

Nonindependence (or the degree to which responses cluster by group) is evidenced by ICC(1), which is the proportion of total variance that can be explained by group membership (like *R2*, adjusted for group size). A large ICC(1) shows that a single rating from an individual provides a reliable rating of the group mean, and addresses the question how much of the total variance is at the group-level? When discussing this group-level variance, sometimes methodologists refer to it as “between group” variance (see Snijders & Bosker, 1999), however “between group” is just another way of saying ‘group-level’ variance.

To simplify, we removed references to “between samples” and “between studies” variance and instead refer to it as group-level variance. In addition, we also removed the statement about there being a “sufficient amount” of group-level variance to justify our analyses, and instead only stated the ICC(1) level (p. 22-23).

References:

Snijders, T. A. B. & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* Thousand Oaks, CA, Sage Publications.

**Reviewer 2, Comment 22.**

**p. 21 - the sentence beginning "The number of effect sizes decreased but..." is confusing.**

We have now rewritten this sentence on page 24 in an attempt to make it less confusing:

“When effect sizes derived from difference scores were removed, 128 effect sizes and 28 independent samples remained in the analysis. In other words, the majority of the correlations in our original data set were based on the self-criterion residual method (i.e., 76%), consequently, even when studies using difference scores were removed, we still had enough remaining data to conduct our analyses.”

**Reviewer 2, Comment 23.**

**p. 22 - 'the competing moderators' model' should be 'the competing moderators model'**

We no longer have a competing moderators model (see Reviewer 2, Comment 9).

**Reviewer 2, Comment 24.**

**p. 22 - 'the extent to which narcissists' self-enhanced' should be 'the extent to which narcissists self-enhanced'**

We have now changed “narcissists’ self-enhanced” to “narcissists self-enhanced” (p. 28).

**Reviewer 2, Comment 25.**

**p. 28 - 'If round-robin data is available' should be 'If round-robin data are available'**

We have now changed “is” to “are” (p. 35).

**This is a signed review.  
-Simine Vazire**