Personality and Individual Differences 114 (2017) 187–192

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

A simulator of the degree to which random responding leads to biases in the correlations between two individual differences

Nicholas S. Holtzman [a,*], M. Brent Donnellan [b]

[a] Georgia Southern University, United States
[b] Texas A&M University, United States

ARTICLE INFO

Article history:
Received 17 December 2016
Received in revised form 5 April 2017
Accepted 6 April 2017
Available online xxxx

Keywords:
ICRAVUR
Insufficient effort responding
Narcissism
Monte Carlo
Psychopathy
Random responding
Secure attachment
Self-esteem

ABSTRACT

Random responding can inflate Type I and Type II error rates (Huang, Liu, & Bowling, 2015b, *Journal of Applied Psychology*, *100*). Type II error inflation often involves certain variables having Invalid Centered Responses And Valid Uncentered Responses (ICRAVUR; pronunciation: /aikreıvər/). Although Huang et al. (2015b) offer a set of formulas for calculating the expected bias in a correlation when such variables are present, they do not offer a way to simulate the effects. We offer two sets of Monte Carlo simulations of ICRAVUR variables. Study 1 examines the correlation between narcissism and psychopathy—thought to be a large effect. The effect was inflated (by $r = 0.16$), comparable to what the Huang formulas forecast. Study 2 examines the correlation between secure attachment and self-esteem—thought to be a large effect. The effect was inflated (by $r = 0.26$), but this time the simulation result was larger than the forecast from the Huang formulas. Thus, our simulator offers a way to test tailored hypotheses about specific variables—sometimes yielding effects more extreme than the Huang formulas. We guide the readers through software, available at the first author's website, allowing for estimating the impact of ICRAVUR variables on any Pearson correlation.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Random responding can inflate correlations under certain empirical conditions (Huang et al., 2015b). This may seem counterintuitive, as classic texts argue that random responding attenuates relationships (Nunnally & Bernstein, 1994). After all, random responding adds noise to a given variable and this will decrease reliability thereby presumably attenuating associations with external correlates. Thus, the traditional logic is that random responding is little more than a nuisance that ultimately reduces statistical power (i.e., the ability to detect a true effect that exists in nature). Random responding can make correlations more extreme (Huang et al., 2015b) under the following three conditions: (a) There are participants who respond in a valid way in which the mean response falls away from the midpoint of the scale—these are valid uncentered responses; (b) there are participants who respond in a random way thereby yielding averages in the middle of the scale—these are invalid centered responses; and (c) the first two conditions are met for both variables that enter into the correlation. Because the variables involve Invalid Centered Responses And Valid Uncentered Responses, we call these ICRAVUR variables (pronunciation: /aikreıvər/, like the word "eye", the word "crave", and the suffix "er").

We aim to do several things in this paper. First, we reiterate the point made by Huang and colleagues: Random responding is a serious threat to making valid conclusions in research on individual differences. Accordingly, we encourage readers to first consult the work of Huang and colleagues before reading this paper. Second, we demonstrate that ICRAVUR variables may be even more detrimental than Huang and colleagues show. Third, we guide the reader through an easy-to-use Excel spreadsheet that facilitates estimating the impact of ICRAVUR variables. Fourth, we apply this research to two specific widely researched correlations in social-personality psychology—namely, the association between narcissism and psychopathy and the association between self-esteem and secure attachment. This will provide some concrete implications of ICRAVUR variables. Last, in the General Discussion, we provide guidance about how to reduce random responding, given the emerging consensus that it is so insidious (Huang et al., 2015b).

To begin, we will walk the reader through an illustrative case of how ICRAVUR variables manifest in a biased correlation, shown in Fig. 1. The two ICRAVUR variables are measured on Likert-type scales that range from one to five. In this particular situation, there are fourteen total respondents; half are random respondents and half respond in a valid way. The seven respondents who respond in a valid way yield scores clustered around the bottom left part of the figure (near x = 1.5, y = 1.5). This is the type of scenario one might expect on some measures of psychopathology—most people score low; these are the valid

* Corresponding author at: Georgia Southern University, Department of Psychology, Post Office Box 8041, Statesboro, GA 30460-8041, United States.
E-mail address: nholtzman@georgiasouthern.edu (N.S. Holtzman).

http://dx.doi.org/10.1016/j.paid.2017.04.013
0191-8869/© 2017 Elsevier Ltd. All rights reserved.

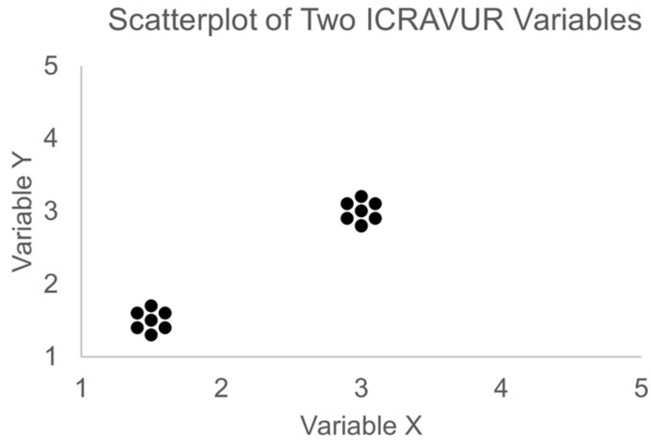

Fig. 1. Scatterplot of two ICRAVUR variables. The correlation between them produces a large positive correlation, despite the fact that in reality there is no correlation between the two variables within valid respondents.

uncentered responses. In addition, there are seven respondents who respond in an invalid (and, in this case, random) way, yielding scores clustered around the middle part of the figure ($x = 3.0$, $y = 3.0$); these are the invalid centered responses. This clustering in the middle is to be expected when participants are responding randomly. For instance, a random respondent may select a "1", a "3", and a "5" on a three item scale—thus averaging to 3.00 overall. The reader will note that each of the variables (X and Y) can be considered an ICRAVUR variable, as each variable has invalid centered responses and valid uncentered responses. It is the process of correlating two ICRAVUR variables that ultimately leads to more extreme correlation magnitudes. In this case, the correlation among only valid responses is essentially 0.00 (and the correlation among only invalid responses is likewise 0.00), but the correlation among all responses is virtually 1.00. This is an extreme case, but it illustrates how much of an impact two ICRAVUR variables might have on one's research conclusions. One must consider the invalid centered responses (ICR) in conjunction with the valid uncentered responses (VUR) in order to see the impact ICRAVUR variables can have on correlation magnitudes.

Thus far, in this paper, we have reiterated arguments by Huang and colleagues (2015b); indeed our research would be impossible without theirs. However, we take their arguments further. In particular, the process outlined in their paper does not capture item-level responses, nor does it capture within-person variation on a particular measure (see their Formula 13). So, although Huang can indeed approximate the bias, confidence in that estimate is lacking. The bias in the correlation could be smaller or larger than what the Huang paper and formulas imply. Thus, it is necessary to create a simulation that takes into account both the item-level responses and the within-person variation on a particular measure (in addition to all of the factors Huang and colleagues account for). Thus, we aim to extend the utility of Huang's argument by using a Monte Carlo simulator for the same purpose (i.e., estimating the magnitude of the bias in r). This will allow us to determine if there are cases in which the inflation of Type I errors might be even more likely than Huang and colleagues suggested (Indeed, we will show that this is true).

In the cases we present in this paper, we will show how ICRAVUR inflates the key correlations toward $+1.00$, although it is possible that a correlation may be biased toward -1.00 under different circumstances. We examine two different, yet widely discussed associations in the social-personality psychology literature: the association between narcissism and psychopathy (Study 1) and the association between secure attachment and self-esteem (Study 2). The first study has implications for the literature on the debate about the extent to which narcissism and psychopathy are separate constructs (Jones & Paulhus, 2011; Muris, Merckelbach, Otgaar, & Meijer, 2017; Vize, Lynam, Collison, & Miller, 2016) and the second study has implications for the literature on the correlates of attachment security and how attachment is related to self-esteem and possibly depression (Hart, Shaver, & Goldenberg, 2005; Roberts, Gotlib, & Kassel, 1996).

1.1. Software

We present two simulation studies using a Microsoft Excel file that could generate output given certain input specifications (available at https://nickholtzman.com/publications/). First, the file has in it an option for including up to 1000 participants. It also has an option for the fraction of valid participants. For each survey, the user must specify: (1) the number of items up to 100, (2) the lowest Likert scale response option (e.g., 1 on a 5-point Likert-type scale), (3) the highest Likert scale response option (e.g., 5 on a 5-point Likert-type scale), (4) the mean of valid responses based on prior research, (5) the standard deviation of valid responses across participants from prior research, (6) the within-person standard deviation of responses from prior research, and (7) the mean of invalid responses, which is the center point on the scale (e.g., 3 on a 1-to-5 Likert-type scale). Based on the number of participants specified and the fraction of valid participants, the program creates a list of the participants who are providing valid, non-random data.

If a participant is providing non-random data, then their true score is the mean of valid responses plus some random normal deviate; the standard deviation for the non-random data is the standard deviation of valid responses across participants. If the participant is providing random data, then their theoretical true score is the scale midpoint. The program simply generates a random number given the endpoints of the scale (e.g., 1 to 5 using a 5-point scale). For the non-random data, the program cuts off the scores at the lower limit and upper limit; that is, if a random normal deviate pushes a score outside of the range of possible values on the scale, then the program pulls the score back within range. The true score—definitely within range after applying the algorithm—is provided. Next, the program generates a score for the first item, with the normal deviate, where the deviate is generated based on the within-person standard deviation of responses that the researcher specifies. The same upper limits and lower limits are applied in order to adjust data that is out of the Likert range, and then the final score for that particular item is generated. This process is repeated for all items, up to the number of items specified by the user. The final item scores are averaged for the survey. This process is repeated for the second survey.

The program returns the expected correlation when the correlation in the population is actually zero (it calculates the expected correlation 50 times, simultaneously, labeled "Trials"). This provides 50 estimates of the bias in the observed correlation. Negative expected correlations indicate that the correlation estimate in empirical work is biased toward -1.00. Positive expected correlations indicate that the correlation estimate in empirical work is biased toward $+1.00$. Finally, the program is also designed to calculate Cronbach's alpha for an estimate of internal consistency; it is currently programmed to calculate alpha only on the data shown (i.e., not 50 times over).

2. Study 1

2.1. Introduction to Study 1

In the first study, we estimate the relationship between narcissism and psychopathy. Narcissism involves traits such as arrogance, entitlement, and vanity (Raskin & Terry, 1988); psychopathy involves traits such as callousness, recklessness and antisocial behavior (Neumann & Hare, 2008). Recently, researchers have highlighted the finding that these traits are "substantially intercorrelated" (Muris et al., 2017, p. 183). Indeed, the correlation between narcissism and psychopathy is estimated to be 0.42, with a 95% CI of 0.39 to 0.45 (O'Boyle, Forsyth, Banks,

& McDaniel, 2012); it is 0.51 when corrected for unreliability. The apparent convergence between the two traits has led some researchers to collapse narcissism and psychopathy as being facets of a larger construct (Jonason, Li, & Teicher, 2010). Along with the construct Machiavellianism, these traits are known as the Dark Triad (Paulhus & Williams, 2002). A substantive question in the literature is whether there is anything unique about narcissism apart from psychopathy (or Machiavellianism). The question at hand here is whether the apparent correlation between narcissism and psychopathy is upwardly biased in the literature because both narcissism and psychopathy may involve ICRAVUR variables. If the correlation is not as large as previously thought, then this enables researchers to make additional conceptual and empirical distinctions between the constructs.

The empirical situation for the first study is as follows: Most people score below the midpoint of the scale on both variables because both narcissism and psychopathy are relatively low base rate phenomena. Scores for random respondents, in contrast, tend to collect around the middle of the scale. We expect that random responding will positively bias the observed correlation—that is, that the correlation will appear to be larger in studies than it is in reality. Overall, it is expected that this empirical situation will lead to inflated estimates of the correlation between narcissism and psychopathy.

2.2. Method

2.2.1. Specifications

To simulate the association between narcissism and psychopathy, we used the following specifications: For psychopathy, (1) item number was 64, (2) lowest response option was 1.00, (3) highest response option was 5.00, (4) the mean of valid responses was 2.05, (5) the standard deviation of valid responses across participants was 0.39, (6) the within-person standard deviation of responses was 1.43, and (7) the mean of invalid responses was 3.00. For narcissism, (1) item number was 40, (2) lowest option was 0.00, (3) highest option was 1.00, (4) the mean of valid responses was 0.30, (5) the standard deviation of valid responses across participants was 0.17, (6) the within-person standard deviation of responses was 0.45, and (7) the mean of invalid responses was 0.50. These numbers are based on the use of the commonly used 64-item SRP measure of psychopathy (Paulhus, Neumann, & Hare, in press) and the widely used 40-item Narcissistic Personality Inventory (Raskin & Terry, 1988).

Using a large sample ($N = 1000$), we ran 50 simulations while the fraction of valid participants was set to various levels (0.00 to 1.00) in 0.05 increments (i.e., 50 simulations in which the proportion of valid respondents was 0.00; 50 simulations of 0.05; 50 simulations of 0.10, etc., up to 50 simulations of 1.00), thereby covering the entire range of scenarios.

2.3. Results

First, we calculated the bias based on an R script provided by an anonymous reviewer who graciously instantiated the Huang formulas. Then we conducted the Monte Carlo simulations using the Excel sheet available from the first author's website (https://nickholtzman.com/publications/). The simulations revealed that the published correlation between narcissism and psychopathy is too close to the positive pole by $r = 0.16$ when the fraction of valid respondents was 0.40—this was the maximum bias we observed for this two-variable scenario. That is, the published effect is 0.42, but the association in the population could be as low as 0.26. Fig. 2 displays the bias based on the Huang formulas and based on our simulation (i.e., the Monte Carlo Simulation [MC simulation]) as a function of the proportion of valid respondents. The bias was evident in both the Huang formulas and the MC simulation, albeit with different skew for the two ways of calculating ICRAVUR-based bias in r.

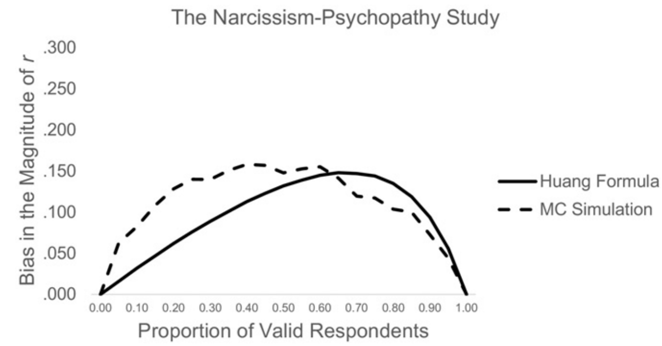

Fig. 2. Biases in the magnitude of r (between narcissism and psychopathy) as a function of the proportion of valid respondents and the two types of estimation procedures: The Huang formulas and the Monte Carlo (MC) simulation.

2.4. Discussion

Under conditions of random responding, the observed correlation between narcissism and psychopathy is biased in favor of the hypothesis that narcissism and psychopathy are strongly positively correlated (using the standards from Hemphill, 2003). Therefore, narcissists are probably slightly less psychopathic than originally presumed and vice versa. This means that theories that collapse the two constructs may have overemphasized the similarities of these traits (because both are ICRAVUR variables). Thus, theories that emphasize the conceptual differences between the two constructs may be more reasonable (Holtzman & Donnellan, 2015; Mealey, 1995; Morf & Rhodewalt, 2001; Paulhus, 2014). At the very least, future studies should attempt to estimate associations between narcissism and psychopathy while taking steps to address random responding (a topic covered in the General Discussion).

It is worth mentioning that the Monte Carlo simulations produced results that more or less mapped onto the forecasts of Huang's formulas. Therefore, we set out to see whether this finding could be true of other cases that involve ICRAVUR variables. We set out to explore another instance in social-personality where ICRAVUR variables may be at play—namely, the correlation between self-esteem and secure attachment.

3. Study 2

3.1. Introduction to Study 2

Self-esteem reflects an overall positive evaluation of the self (Donnellan, Trzesniewski, & Robins, 2011) and it is thought to be a vulnerability factor for depression (Orth, Robins, & Meier, 2009; Sowislo & Orth, 2013). The core idea of attachment theory is that responsive caregivers help individuals develop particular patterns of thinking about the self in the context of close relationships (so-called internal working models) and that these models guide subsequent thoughts, feelings, and behaviors (Hazan & Shaver, 1987). A secure internal working model is thought to help individuals flourish and thrive in life and attachment security is generally correlated with mental health variables such as lower depression and higher well-being (B. C. Feeney & Collins, 2015). Roberts et al. (1996) explicitly suggested that self-esteem was a mediator between attachment security and subsequent depression. Indeed, researchers have long found positive associations between attachment security and self-esteem (Collins & Read, 1990; Feeney & Noller, 1990).

Gillath, Hart, Noftle, and Stockdale (2009) developed a measure of state adult attachment security (items like "I feel loved" and "I feel secure and close to other people.") and noted that attachment security and self-esteem have long been implicated in the literature. The second author (Donnellan) had data from a large sample of college students that included their attachment security scale along with the widely-used Rosenberg Self-Esteem Scale (Rosenberg, 1965). The correlation between the state security scale and global self-esteem was 0.540 ($n = 658$) and we observed that both self-esteem and state security have sample averages above the midpoint (State Security: $M = 5.84$ on a 7-point scale; Self-Esteem $M = 3.80$ on a 5-point scale; alphas = 0.94 and 0.88, respectively). Indeed, it is widely reported that self-esteem scores are above the scale midpoint (Baumeister, Campbell, Krueger, & Vohs, 2003) and most people report being securely attached, at least in the United States (Mickelson, Kessler, & Shaver, 1997). Thus, we were concerned that the correlation between security and self-esteem could be inflated because of ICRAVUR variables.

3.2. Method

3.2.1. Specifications

To assess the potential for bias in the association between self-esteem and secure attachment, we used the following specifications based on the observed data: For self-esteem, (1) item number was 10, (2) lowest scale option was 1.00, (3) highest scale option was 5.00, (4) the mean of valid responses was 3.80, (5) the standard deviation of valid responses across participants was 0.70, (6) the within-person standard deviation of responses was 0.74, and (7) the mean of invalid responses was 3.00. For secure attachment, (1) item number was 7, (2) lowest scale option was 1.00, (3) highest scale option was 7.00, (4) the mean of valid responses was 5.83, (5) the standard deviation of valid responses across participants was 1.12, (6) the within-person standard deviation of responses was 0.57, and (7) the mean of invalid responses was 4.00.

We ran the simulations in the same way as we did in Study 1: Using a large sample ($N = 1000$), we ran 50 simulations while the fraction of valid participants was set to 0.00 to 1.00 in 0.05 increments (i.e., 50 simulations of 0.00; 50 simulations of 0.05; 50 simulations of 0.10, etc.).

3.3. Results

Fig. 3 displays the results of Study 2. The simulation revealed that the observed correlation between secure attachment and self-esteem is too close to the positive pole by $r = 0.26$ (when the proportion of valid respondents was 0.55). Thus, while the observed effect size hovers around 0.50, our simulation suggests that the association could be as low as 0.24 in the population. The most noteworthy aspect of Fig. 3, however, is that the bias magnitudes for the Monte Carlo simulation are more elevated than the bias magnitudes based on the Huang formulas.

3.4. Discussion

As was the case for psychopathy and narcissism, there are indications that observed correlations between self-esteem and attachment security might be positively biased because of random responding. Valid scale means tended to be lower than the midpoint for narcissism and psychopathy but above the midpoint for attachment security and self-esteem. Indeed, the college students in our example dataset tended to report feeling securely attached and having a positive sense of self-worth. This fact might create problems for considering the association between the two constructs if some subgroup of the total sample includes participants who are exerting insufficient effort and responding randomly. Moreover, in one respect, the simulation results were conservative because we assumed all of the college student responses (from our original sample) were valid; this assumption means that valid respondents had even more extreme responses (i.e., the mean that we used for valid respondents was artificially pulled toward the center point on the Likert scale). Ultimately, this very likely led us to underestimate the size of the bias because the bias tends to go up when the mean for valid responses is more extreme.

In terms of the larger implications for the literature on random responding: The simulation in Study 2 shows that the Huang formulas may have some limitations for forecasting bias; in this case, apparently the Huang formulas underestimated the bias in r. We recommend that researchers concerned about ICRAVUR variables in their work use our Excel spreadsheet for implementing the Monte Carlo simulations as a supplement to the Huang formulas.

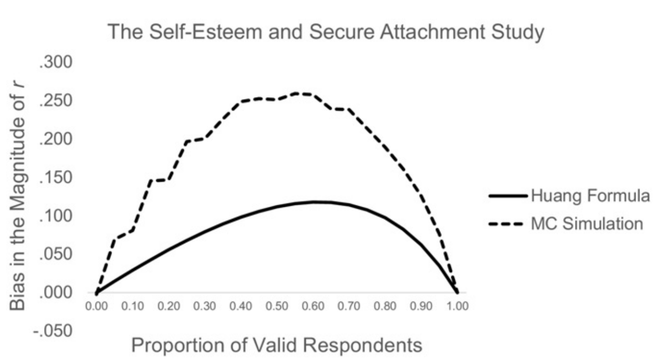

Fig. 3. Biases in the magnitude of r (between self-esteem and secure attachment) as a function of proportion of valid respondents and the two types of estimation procedures: The Huang formulas and the Monte Carlo (MC) Simulation.

4. General discussion

In this paper, we have illustrated two scenarios in which ICRAVUR variables can lead to biased effect size estimates. The contribution of this paper is methodological but these results have clear substantive implications for interpreting correlations between variables when both attributes are pulled toward either the low end of the scales (as in Study 1) or the high end of the scales (as in Study 2). The spreadsheet used to run the Monte Carlo simulations is freely available from the first author (it is posted on https://nickholtzman.com/publications/) and it may aid in the discovery of possible biases in reported associations because of how random responding can influence the associations between two variables. As we stated in the Discussion for Study 2, the biases in effect size magnitudes are not fully captured by the formulas offered by Huang and colleagues (2015b). This is visible in Fig. 3 in the elevation for MC simulation relative to the effects for the Huang formulas in that same figure.

Bias is likely to be washed out if researchers select scales that cluster around the midpoint because the true scores are clustered at this point as well. The problem is that this is not always easy to do because many of the constructs of interest to psychologists are pulled in one direction or the other. Psychopathic tendencies are probably relatively low in the population whereas attachment security is probably relatively high. These possibilities point to the wide applicability of the method contained in this paper. This paper echoes the concern voiced by Huang et al. (2015b) about how random responding can also make effects more extreme and thereby contribute to Type M errors whereby the observed effect size estimates are biased (Gelman & Carlin, 2014).

In other words, ICRAVUR variables can contribute to inferential errors about the magnitude of the association between two variables. To the extent that researchers are chasing subtle effects, random responding becomes even more problematic, as the real effect may be dwarfed by the impact of random responding. If one calculates the variance explained in Study 2, the variance explained without accounting for random responding is 25%, but the variance explained after accounting for random responding is actually less than 5%, revealing a 20% change in the R^2 value. This phenomenon is also an issue for meta-analyses given that researchers are often working with published data of unknown quality. Meta-analysts who do not clean the data themselves are very likely analyzing data in which effect sizes have been contaminated by ICRAVUR variables.

Fortunately the concerns identified in this paper can be addressed moving forward. In particular, we recommend that researchers take steps to identify careless responding in their own datasets and perhaps more importantly, take steps to minimize careless responding in the future. In this vein, we reiterate pleas to clean data (Desimone, Harms, & Desimone, 2015; Osborne, 2013; Osborne & Blanchard, 2011). Some recommendations for data cleaning include using items that catch invalid respondents (i.e., "catch items", AKA "bogus items"). Example items are: "I have landed on Mars" or simply "Please mark strongly agree." Pre-existing scales are available as well (see Table 1 in Huang, Bowling, Liu, & Li, 2015a; Marjanovic, Struthers, Cribbie, & Greenglass, 2014). Respondents who select the wrong responses would be excluded from analyses, as it can be inferred that they were not paying attention. Additionally, Curran (2016) offers a number of post-hoc analyses one can do to clean one's data, even if catch items are not included when designing the survey. Examples include calculating the inter-item standard deviations (Marjanovic, Holden, Struthers, Cribbie, & Greenglass, 2015) or excluding participants who endorse antonyms (Goldberg & Kilkowski, 1985), such as the two antonyms "I am very happy" and "I am very sad." We also like the idea of including a conspicuous opt-out question at the end of the survey so participants can self-identify as providing invalid information on the survey; their responses can be discarded without losing incentives (Aust, Diedenhofen, Ullrich, & Musch, 2013).

When using data cleaning procedures, we have two additional recommendations. First, pre-register those data cleaning procedures to reduce the possibility the decisions will be results-contingent (i.e., researchers may select results that are consistent with their underlying hypotheses). There are very few hard-and-fast rules about the exclusion of careless respondents, and so the process can add to researcher degrees of freedom (Simmons, Nelson, & Simonsohn, 2011; Wicherts et al., 2016). Second, consider reporting results both using the final cleaned dataset and with the complete dataset before cleaning. This can provide researchers with insights about the potential degree of bias introduced by careless responding. That is, seeing how correlations change across datasets is informative and will help provide the field with additional examples of when careless responding attenuates correlations and when it pushes correlations toward the extremes.

In conclusion, situations that involve some random responding under conditions when true scores are not centered in the middle of the Likert scale are likely to involve biased estimates of effect size (Huang et al., 2015b). For analyses that involve Pearson product moment correlations, this bias can be estimated using the Excel spreadsheet we developed; this spreadsheet is more precise than the formulas offered in Huang et al. (2015b), mainly because it takes into account the number of items on a measure and the within-person variability in that measure; indeed, we demonstrated that the two estimation approaches sometimes differ in the results they yield. It is our hope that this simulation will elucidate when and to what extent random responding is influential. In the focal cases here, it seems that the narcissism-psychopathy correlation is inflated and the correlation between self-esteem and secure attachment is inflated. If attempts are taken to minimize careless responding, we suspect that researchers will get a solid understanding of the distinctiveness of these constructs.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.paid.2017.04.013.

References

Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods*, *45*, 527–535. http://dx.doi.org/10.3758/s13428-012-0265-2.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science*, 1–44.

Collins, N. L., & Read, S. J. (1990). Adult attachment, working models, and relationship quality in dating couples. *Journal of Personality and Social Psychology*, *58*, 644–663. http://dx.doi.org/10.1037/0022-3514.58.4.644.

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology*, *66*, 4–19. http://dx.doi.org/10.1016/j.jesp.2015.07.006.

Desimone, J. A., Harms, P. D., & Desimone, A. J. (2015). Best practice recommendations for data screening. *Journal of Organizational Behavior*, *36*, 171–181. http://dx.doi.org/10.1002/job.1962.

Donnellan, M. B., Trzesniewski, K. H., & Robins, R. W. (2011). Self esteem: Enduring issues and controversies. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (pp. 718–746). New York: Wiley-Blackwell.

Feeney, B. C., & Collins, N. L. (2015). Thriving through relationships. *Current Opinion in Psychology*, *1*, 22–28. http://dx.doi.org/10.1016/j.copsyc.2014.11.001.

Feeney, J. A., & Noller, P. (1990). Attachment style as a predictor of adult romantic relationships. *Journal of Personality and Social Psychology*, *58*, 281–291. http://dx.doi.org/10.1037/0022-3514.58.2.281.

Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, *9*, 641–651. http://dx.doi.org/10.1177/1745691614551642.

Gillath, O., Hart, J., Noftle, E. E., & Stockdale, G. D. (2009). Development and validation of a state adult attachment measure (SAAM). *Journal of Research in Personality*, *43*, 362–373. http://dx.doi.org/10.1016/j.jrp.2008.12.009.

Goldberg, L. R., & Kilkowski, J. M. (1985). The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and antonym pairs. *Journal of Personality and Social Psychology*, *48*, 82–98.

Hart, J., Shaver, P. R., & Goldenberg, J. L. (2005). Attachment, self-esteem, worldviews, and terror management: Evidence for a tripartite security system. *Journal of Personality and Social Psychology*, *88*, 999–1013. http://dx.doi.org/10.1037/0022-3514.88.6.999.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology*, *52*, 511–524. http://dx.doi.org/10.1037//0022-3514.52.3.511.

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*, 78–79. http://dx.doi.org/10.1037/0003-066x.58.1.78.

Holtzman, N. S., & Donnellan, M. B. (2015). The roots of Narcissus: Old and new models of the evolution of narcissism. In V. Zeigler-Hill, L. L. M. Welling, & T. K. Shackelford (Eds.), *Evolutionary perspectives on social psychology* (pp. 479–489). New York: Springer.

Huang, J. L., Bowling, N. A., Liu, M., & Li, Y. (2015a). Detecting insufficient effort responding with an infrequency scale: Evaluating validity and participant reactions. *Journal of Business and Psychology*, *30*, 299–311. http://dx.doi.org/10.1007/s10869-014-9357-6.

Huang, J. L., Liu, M., & Bowling, N. A. (2015b). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology*, *100*, 828–845. http://dx.doi.org/10.1037/a0038510.

Jonason, P. K., Li, N. P., & Teicher, E. A. (2010). Who is James Bond? The dark triad as an agentic social style. *Individual Differences Research*, *8*, 111–120.

Jones, D. N., & Paulhus, D. L. (2011). Differentiating the dark triad within the interpersonal circumplex. In L. M. Horowitz, & S. Strack (Eds.), *Handbook of interpersonal psychology* (pp. 249–268). New York, NY: Wiley.

Marjanovic, Z., Holden, R., Struthers, W., Cribbie, R., & Greenglass, E. (2015). The inter-item standard deviation (ISD): An index that discriminates between conscientious and random responders. *Personality and Individual Differences*, *84*, 79–83.

Marjanovic, Z., Struthers, C. W., Cribbie, R., & Greenglass, E. R. (2014). The conscientious responders scale: A new tool for discriminating between conscientious and random responders. *SAGE Open*, *4*, 1–10. http://dx.doi.org/10.1177/2158244014545964.

Mealey, L. (1995). The sociobiology of sociopathy: An integrative evolutionary model. *Behavioral and Brain Sciences*, *18*, 523–541.

Mickelson, K. D., Kessler, R. C., & Shaver, P. R. (1997). Adult attachment in a nationally representative sample. *Journal of Personality and Social Psychology*, *73*, 1092–1106. http://dx.doi.org/10.1037/0022-3514.73.5.1092.

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry*, *12*, 177–196. http://dx.doi.org/10.1207/s15327965pli1204_1.

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (Narcissism, Machiavellianism, and Psychopathy). *Perspectives on Psychological Science*, *12*, 183–204. http://dx.doi.org/10.1177/1745691616666070.

Neumann, C. S., & Hare, R. D. (2008). Psychopathic traits in a large community sample: Links to violence, alcohol use, and intelligence. *Journal of Consulting and Clinical Psychology*, *76*, 893–899. http://dx.doi.org/10.1037/0022-006x.76.5.893.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'Boyle, E. H., Jr., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the dark triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, *97*, 557–579. http://dx.doi.org/10.1037/a0025679.

Orth, U., Robins, R. W., & Meier, L. L. (2009). Disentangling the effects of low self-esteem and stressful events on depression: Findings from three longitudinal studies. *Journal of Personality and Social Psychology*, *97*, 307–321. http://dx.doi.org/10.1037/a0015645.

Osborne, J. W. (2013). Is data cleaning and the testing of assumptions relevant in the 21st century? *Frontiers in Psychology*, *4*. http://dx.doi.org/10.3389/fpsyg.2013.00370.

Osborne, J. W., & Blanchard, M. R. (2011). Random responding from participants is a threat to the validity of social science research results. *Frontiers in Psychology*, *2*. http://dx.doi.org/10.3389/fpsyg.2010.00220.

Paulhus, D. L. (2014). Toward a taxonomy of dark personalities. *Current Directions in Psychological Science*, *23*, 421–426. http://dx.doi.org/10.1177/0963721414547737.

Paulhus, D. L., Neumann, C. F., & Hare, R. D. (2017). *Manual for the self-report psychopathy scale*. Toronto: Multi-Health Systems (in press).

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*, 556–563 (Pii s0092-6566(02)00505-6) 10.1016/s0092-6566(02)00505-6

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*, 890–902. http://dx.doi.org/10.1037/0022-3514.54.5.890.

Roberts, J. E., Gotlib, I. H., & Kassel, J. D. (1996). Adult attachment security and symptoms of depression: The mediating roles of dysfunctional attitudes and low self-esteem. *Journal of Personality and Social Psychology*, *70*, 310–320. http://dx.doi.org/10.1037/0022-3514.70.2.310.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366. http://dx.doi.org/10.1177/0956797611417632.

Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin*, *139*, 213–240. http://dx.doi.org/10.1037/a0028931.

Vize, C. E., Lynam, D. R., Collison, K. L., & Miller, J. D. (2016). Differences among dark triad components: A meta-analytic investigation. *Personality Disorders*. http://dx.doi.org/10.1037/per0000222.

Wicherts, J., Veldkamp, C., Augusteijn, H., Bakker, M., van Aert, R., & van Assen, M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies. A checklist to avoid p-hacking. *Frontiers in Psychology*, *7*. http://dx.doi.org/10.3389/fpsyg.2016.01832.