Reviewer's report

Title: Psychometric evaluation and normative data of the Pure Procrastination Scale, the Irrational Procrastination Scale, and the Susceptibility to Temptation Scale

Version: 3  Date: 16 September 2014

Reviewer: Piers Steel

Reviewer's report:

This is a good paper. It is well contextualized and the description of methods is particularly adept. The sample size is good and there is well chosen collection of measures. You have a good selection of statistical measures, with MDC being notable in this context.

Major Compulsory Revisions

1. Some new publications came out in 2014 that would be helpful, especially these two:


It establishes that impulsiveness is core to procrastination better than some of the cites you are presently using.


They did a similar study to yours, except with a substantively smaller French sample and somewhat different array of measures (e.g., using the UPPS to assess impulsiveness). It is worth noting and contrasting.

2. You used a restricted sample by selecting participants based on a minimum of 32 points on the IPS. This is appropriate for clinical inferences, which is your goal, but causes range restriction, which reduces the strength of your variance reliant statistics. For example, for the full population, Steel (2010) reports that the PPS has a standard deviation of 10 (i.e., 12 times .83), which becomes 3.62 with your sample restricted to pretty much the upper half of the scale. If you correct for range restriction, the IPS and PPS become again pretty much equivalent. A caveat noting that when dealing with a clinical sample, and consequently a much reduced range, it is harder to get large coefficients. You might give an example, correcting for range restriction, to underscore this, as right now you almost apologize for it (e.g., “albeit all rather small”).
3. Elsewhere, you need to also incorporate range restriction into your understanding. During discussion, when you compare your results with Steel (2010), you compare your reliability corrected correlations with his raw correlations. The correct comparison would be with a range restriction correction, which would be .90, slightly better that Steel (2010) originally reported. Similar adjustments should be made for the STS comparison. Because correcting for range restriction isn’t quite as good as simply having a full range, you get a less biased but also a less accurate estimate. Again, this isn’t a problem using a restricted range as you want to apply your results to a clinical sample, but it fully explains the reduced correlations.

4. Factor structures are always somewhat difficult to be definitive. As you note, reverse score items are often an instrumentation and routinely load lower or extract to an oblique dimension. Despite this, test developers tend to leave this to help identify sloppy respondents, and minimize the errors from a directional bias. I concur with your interpretation here. Here’s an article which reviews the issue, as per: “Researchers have shown that they may result in an artifactual response factor consisting of all negatively-worded items (Harvey, Billings & Nilan, 1985; Schmitt & Stultz, 1985).”


You may want to use this to better inform your discussion of the matter.

5. On the other hand, you also found that the PPS had two factors, as did Rebetez et al. (2014) factor analysis of the PPS. The only problem is that the two of you have entirely different items loading on the second factor. You note that Factor 1 correlated higher than Factor 2 with your other scales, but the overall PPS (in Table 6) has typically even better correlations. This was the same with Rebetez et al., where the overall PPS typically had the strongest correlations. Given that you thought your findings were preliminary and need to be repeated, having not been repeated, please give interpretation.

6. My understanding is that disorders tend to group in impulsive related and anxiety related, as per the references below. Though externalizing can lead to anxiety, and a relationship (as per below again), they are not internalizing. As per temporal motivation theory, though depressive like symptoms of low self-efficacy and diminished mood predict procrastination, the dominant feature is impulsiveness. I would like more nuanced thinking from you on this points. For example, future research directions may choose to see if procrastination is associated with other externalizing more than internalizing disorders. Likely, we will see a range of underlying reasons for procrastination among people; while there may be a majority who are impulsiveness dominant, there are some who predominantly require addressing their anxiety issues. A diagnostic tool identifying the most promising venues for therapy may be beneficial. This is in line with your notion of the typical and atypical procrastinator.

connections to personality. Personality and Individual Differences, 30(7), 1245-1259.


7. Discussion is the weakest part. Your thoughts here should be more focused and supported. For example, there is very little research backing up any major differences between paper and pencil versus internet administration, such as:


A quick scholar search of internet versus paper will confirm.

Minor Essential Revisions

1. Table 6 is misaligned.
2. Keep the order of your scales consistent among Tables (i.e., tables 5 and 6 have different ordering).
3. You write “the procrastination scales did not correlate highly, r = -.17 to -.35...” It should be -.19 to .35.
4. Forth versus Fourth.
5. Some minor referencing issues (e.g., capitalization)

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests