Further study of the measurement of procrastination: Using item response theory on the pure procrastination scale

Joel Gagnon[1] · Evandro Morais Peixoto[2] · Frédérick Dionne[1]

Abstract

In recent years, researchers have brought into question the construct validity of the available scales to measure procrastination. Among the instruments assessing procrastination, the Pure Procrastination Scale (PPS) has gained considerable attention from the research community. However, diverging results from past research on the PPS have raised several unanswered questions that are critical to the understanding, operationalization, and assessment of procrastination. This study draws from past research and aims to further investigate the psychometric properties of the PPS. First, this study tests the hypothesis that item responses can be explained by a general factor using bifactor modeling. Second, this study uses the Rasch model to study the psychometric properties of each item of the PPS. Finally, this study sought to create cutoff scores to discriminate between low, medium, and high levels of procrastination on the PPS. The sample was comprised of 934 French-speaking university students. Results showed that the bifactor model had better fit statistics across all indices. At the item level, results from the Rasch model showed that the PPS provides relatively little information for participants with low and high attribute levels whereas the PPS provides great levels of information for participants in the middle range of the attribute. Finally, cutoff scores were created and converted into raw scores to facilitate their use among researchers and clinicians.

Keywords Pure procrastination scale · Procrastination · Bifactor model · Item response theory · University students

Procrastination is defined as the voluntary delay of an intended course of action despite expecting negative consequences that outweigh the positive consequences of the delay (Klingsieck 2013b; Sirois and Pychyl 2013; Steel 2007). Procrastination differentiates itself from other forms of delay (for ex., strategic delay) and is characterized by a dysfunctional and irrational delay leading to feelings of discomfort, stress, and guilt (Klingsieck, 2013). Procrastination is often thought as a self-regulatory failure or a mood repair strategy where one's priority is to feel good in the short term often at the cost of one's long-term goals (Pychyl and Flett 2012; Sirois and Pychyl 2013). Past research has found that procrastination exists in different life domains but is more typical for the academic and work domains (Klingsieck 2013a). Researchers estimate the prevalence of students engaging in procrastinatory behaviors to varying from 70 (Schouwenburg et al. 2004) to 95% (Ellis and Knaus 2002). For 20 to 30% of university students, procrastination is a chronic problem that affects their academic performance (Kim and Seo 2015), and their quality of life (Steel and Ferrari 2013). Regarding gender differences, in their epidemiological study, Steel and Ferrari (2013) found that men tend to procrastinate more than women.

A review by Steel (2010) has brought into question the construct validity of the available scales to measure procrastination; criticizing their theoretical basis specifically in terms of being able to differentiate various types of procrastination (i.e., arousal, avoidant, and decisional). This criticism has inspired efforts to create a refined scale to assess the core component of procrastination (i.e., dysfunctional delay) which resulted in the Pure Procrastination Scale (PPS; Steel 2010). The PPS evaluates the prevalence of procrastinatory behaviors conceptualized as a dysfunctional delay. The PPS is based on items from the Decisional Procrastination Scale (DP; Mann 1982) a five-item scale assessing delay in planning and decision-making, the General Procrastination Scale (GPS; Lay 1986) evaluating implemental or behavioral delay and finally, the Adult Inventory of Procrastination Scale (AIP; McCown et al. 1989) assessing promptness, meeting

✉ Joel Gagnon
   Joel.Gagnon2@uqtr.ca

1   Department of Psychology, Université du Québec à Trois-Rivières, Trois-Rivières, C.P. 500, Quebec G9A 5H7, Canada

2   Universidade de Pernambuco, Recife, Brazil

deadlines, and timeliness. Using Exploratory and Confirmatory Factor Analyses (CFA) on a sample of more than 4000 respondents of the general population, items from these three scales were reduced to 12 items forming the final version of the PPS.

The PPS has been translated into several languages. Results from the original English and Norwegian versions showed that the items were loading on a single core factor labeled dysfunctional delay explaining most of the variance (Steel 2010; Svartdal 2017). In the validation article of the French version, Rebetez et al. (2014) found support for a higher-order factor labelled "dysfunctional delay" with two first-order factors "voluntary delay" and "observed delay." In their validation study of the Swedish version, Rozental et al. (2014) have found the presence of two factors both associated with the notion of voluntary delay potentially suggesting the presence of a single higher-order factor being measured. Svartdal et al. (2016) translated the PPS in Finnish, German, Italian, and Polish and tested four competing CFA models: a one-factor model (Steel 2010), the two-factor model from Rebetez et al. (2014), the two-factor model from Rozental et al. (2014), and a three-factor model in which items from the three original scales used to create the PPS were modeled as separate factors labeled "decisional delay," "implemental delay" and "lateness/timeliness." The authors found that the three-factor model was the best fitting solution across six European countries. This result was later replicated in an independent sample of more than 4100 adults from the general population further suggesting that the PPS assesses three distinct components of procrastination (Svartdal and Steel 2017).

Diverging results from past research raised several unanswered questions that are critical to the understanding, operationalization, and assessment of procrastination. First, it is not clear whether the additional specific factors found in the translated versions of the PPS represent substantive factors that could be used to add unique information to predict outcomes in research and clinical settings. In fact, as Steel (2010) and later Rebetez et al. (2014) and Rozental et al. (2014) argued, there could be an overarching core factor encompassing the distinct factors found in previous research that, when accounted for, would explain most of the variance of the construct over and above the specific factors. The bifactor model from the structural equation modeling framework allows testing for such a research question. The bifactor model is constructed by specifying paths from the general factor to all indicators. In the bifactor model, the general factor represents the source of common variance among all the items. It is interpreted as representing the core psychological construct that the instrument was created to measure. If the specific factors do not explain significant variance over and above the general factor, it is evidence that the specific factors cannot be separated from the general core factor (Chen et al.

2012). These results would bring information on the necessity of scoring subscales or a total score (Reise 2012). The bifactor model offers some advantages over the higher-order model; notably it allows for an easier interpretation of the general factor and it allows for the simultaneous examination of both specific and general influences on the indicators.

While past research has focused on studying the factor structure of the PPS in different cultures, not much is known about the psychometric properties at the item level using Item Response Theory (IRT; Embretson and Reise 2000). The IRT analyses systematically analyze the item information function which provides evidence on the ability of each item to discriminate among people relative to their specific level on the attribute (i.e., theta). As pointed by Rozental and Carlbring (2014) and later by Svartdal and Steel (2017), little is known about the capacity of the self-reported measures of procrastination to differentiate between low and high levels of procrastination. In this regard, IRT models can bring valuable information by creating cutoff scores that can be used in clinical and research settings to discriminate participants with low, moderate and high levels of procrastination. Establishing cutoff scores is a critical step in the development and use of self-reported instruments (American Educational Research Association, American Psychology Association, and National Council on Measurement in Education 2014).

The Present Study

This study aims to further investigate the psychometric properties of the PPS among university students using modern statistical modeling methods. More specifically, this study first tests the hypothesis that item responses can be explained by a general factor using bifactor modeling. Second, if the bifactor model proves acceptable, the unidimensionality of the PPS would be supported thus allowing for the use of IRT models to study the psychometric properties of each item of the PPS. Third, this study aims to provide cutoff scores based on the IRT item-mapping method. This method predicts the level of theta expected to be observed for each total score points, which allows the establishment of cutoff based on the parameters estimated in a reference group.

Method

Participants

This study used an integrative data analytic approach to answer the research questions. The sample used in this study was an aggregation of samples from three different research projects that have taken place from Fall 2015 to Fall 2018 among university students in the province of Quebec (Canada). The

first sample was comprised of 370 university students (78% female). The second sample was comprised of 476 university students (84% female). Finally, the third sample was comprised of 88 university students (82% female). In all the recruitment, participants were contacted via the Internet (e.g., social media) and asked to take part in research on procrastination. Participants completed questionnaires via the Internet. Before completing questionnaires, participants were informed of the nature of the study and signed informed consent. All of the research projects have been previously approved by a Research Ethics and Integrity Committee. The aggregated sample was comprised of 934 French-speaking university students with a mean age of 26.59 years old ($SD = 7.82$). The sample was predominantly female (81%) and Caucasian (84%), 67% were undergraduates and the majority were studying full time (86%).

Instrument

The English version of the PPS consists of 12 items (Steel 2010) rated on a 5-point Likert scale (1 = *very seldom or not true of me* to 5 = *very often true or true of me*) with higher scores indicating higher levels of procrastination. In the validation article of the French PPS (Rebetez et al. 2014), the authors removed an item based on its poor functioning in the model and its content creating an 11-item scale for the French version. Examples of items are: "I am continually saying I'll do it tomorrow" and "I delay making decisions until it's too late."

Data Analyses

To test for the unidimensionality of the PPS, CFAs were conducted. The fits of four competing CFA models taken from previous research on the factorial structure of the PPS (e.g., Svartdal et al. 2016) were tested against the fits of a bifactor model. The bifactor model was an extension of the three-factor model found by Svartdal et al. (2016) with the addition of an overall core factor with direct paths on all the items.

The models were computed using the robust maximum likelihood method. The adjustment of the models was compared based on the chi-squared and several fit indices to provide a more conservative and reliable evaluation of the solution (Jaccard et al. 1996). The Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) were first evaluated as measures of goodness-of-fit. For these indices, values above .95 are considered favorable (Hu and Bentler 1999). Two badness-of-fit indices were also evaluated: the Root Mean Square Error of Approximation (RMSEA) and the Standardized Root Mean Square Residual (SRMR). For the RMSEA, values between .05 and .08 represent reasonable errors of adjustment (Browne and Cudeck 1993). For the SRMR, values below .05 are considered favorable (Hu and

Bentler 1999). Finally, the Akaike Information Criterion (AIC) was used. The AIC is a fit index used to select among competing non-nested models. Typically, the model with the smallest AIC is preferred (Kline 2015). These analyses were conducted using the "lavaan" package (Rosseel 2012) available in the R software (R Core Team 2016).

Assuming the unidimensionality of the PPS, IRT analyses were then conducted using the rating scale model. Item and subject model parameters were calibrated using the joint maximum likelihood estimation method. To identify the scale metrics, item difficulties (i.e., average thresholds) were set at zero. Subject parameters (i.e., theta), item parameters, model fit indices (i.e., infit and outfit), reliability indices, category probability curves, item threshold values, and item map qualitative analysis were used to analyze the model fits. Differential item functioning (DIF) was computed to assess whether the item functioning was similar across genders. DIF occurs when people with equal levels of the latent construct respond differentially to an item. DIF contrast values of $\geq .64$ indicate moderate to large DIF (Linacre 2014). IRT analyses were conducted using Winsteps software (Linacre 2014). Finally, three different cutoffs were created based on the results of the IRT model. The 25th and 75th percentiles from the theta scores were used as low and high levels of procrastination respectively based on past research that has reported that 20 to 30% of students are chronic procrastinators (Steel and Ferrari 2013).

Results

Confirmatory Factor Analyses

Model fits from competing CFA models are presented in Table 1. The fit indices of the one-factor and the two-factor models were all below the acceptable values. The adjustment of the three-factor model to the data was within the acceptable range. The second-order factor model displayed similar fits to the three-factor model although the AIC index was smaller indicating that this model should be preferred over the three-factor model. With regard to the bifactor model, results showed that it was the best fitting model across all indices.

Item Response Theory

Table 2 presents the latent trait level (theta) of individual items, standard error, fit indices (infit and outfit) of the model, DIF for gender, the DIF contrast statistic, and the Mantel-Haenzel probabilities of the DIF contrast. The two most challenging items were items 10 and 11 whereas the least challenging items were items 4 and 6. The latent trait level ranged from −.85 to 1.47 suggesting that overall, items appear to be assessing a narrow range of procrastination levels. Participants' levels on the attribute and other descriptive

Table 1 Results from competing confirmatory factor analysis models

Model MLR	χ^2 (df)	CFI	TLI	SRMR	RMSEA [95% CI]	AIC
One factor	969.26*** (44)	.823	.779	.087	.168 [.159, .176]	24,717.02
Two-factor	511.73*** (43)	.914	.890	.077	.119 [.110, .128]	24,170.44
Three-factor	261.99*** (41)	.960	.946	.056	.083 [.074, .092]	23,893.46
2nd order factor	253.45*** (41)	.960	.946	.056	.083 [.074, .093]	23,188.13
Bifactor	169.72*** (33)	.975	.959	.032	.073 [.063, .083]	23,105.15

MLR, Maximum Likelihood Robust; *CFI*, Comparative Fit Index; *TLI*, Tucker Lewis Index; *SRMR*, Standardized Root Mean Square Residual; *RMSEA*, Root Mean Square Error of Approximation; *AIC*, Akaike Information Criterion ***$p < .001$

statistics are presented at the bottom of Table 2. The mean theta at the person-level was .07 ($SD = 1.48$). A graphical representation of the relation between latent trait level at the individual item and attribute at the person level shows that most participants reported levels of procrastination for which the scale contains high information (see Fig. 1).

Infit and outfit values express the correspondence between observed and expected theta values for the participants. The infit statistic is more sensitive to respondents with theta values that are close to the item's difficulty level, whereas the outfit statistic is more sensitive to theta values that are far from the item's difficulty level (e.g., when people with high ability miss easy items or people with low ability get hard items). Values for infit and outfit should be between .70 and 1.30 (Bond and Fox 2007). All the items fit statistics were within the acceptable range indicating that the expected model parameters adequately represent the observed responses. The DIF contrast revealed that the items were functioning similarly for both men and women (DIF contrast < .64 and Mantel-Haenzel probabilities all above than $p > .05$).

Figure 2 presents the item map with rating scale model half-point thresholds which are interpreted as the thresholds from which a person with a certain theta score will be more likely to choose one category of response over another. In Fig. 2, the line named "Person" represents the frequency of participants for different points of theta. The line "Statistics" represents the mean and standard deviations of theta scores. The line "Percentile" shows the percentiles of persons located across the theta scale. Results from Fig. 2 showed that the mean attribute level of the participants was .07 ($SD = 1.49$), indicating that the majority of the participants reported medium levels of theta. Cutoff scores were created based on the 25th and the 75th percentiles on the theta scale which resulted in cutoff scores placed at theta −.87 and .98 for low and high levels of procrastination respectively. We draw vertical lines in Fig. 2 to illustrate the relation between the cutoff scores and the item thresholds. These cutoff scores can be converted into raw scores for convenience of use in future research. Raw scores of 27 or less represent low levels of procrastination. Raw scores between 28 and 39 represent average levels of procrastination and raw scores of 40 and higher represent high levels of procrastination.

Discussion

The goal of this study was to further examine the psychometric properties of the PPS in a large population of French-speaking university students. More specifically, this study had three main objectives. Firstly, to study competing CFA models to test whether the factorial structure of the PPS is better represented by a bifactor model rather than by a multi-factorial structure. Secondly, to study the psychometric

Table 2 Results from item response theory model for individual items and descriptive statistics at the person-level
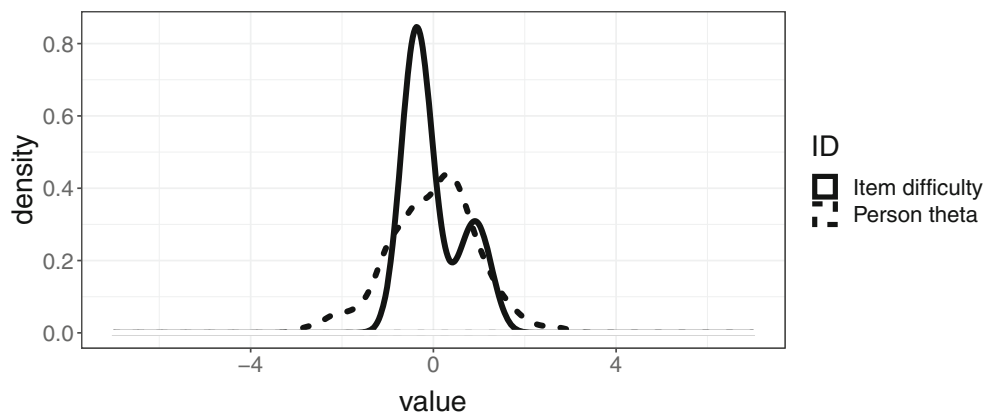
Item	θ	S.E.	Infit	Outfit	r	DiF Men	DiF Women	DiF contrast	Prob.
1. I delay making decisions until it's too late	1.01	.05	1.05	1.06	.71	1.12	.99	−.13	.51
2. Even after I make a decision I delay acting upon it	.04	.05	.95	.95	.75	−.05	.06	.11	.34
3. I waste a lot of time on trivial matters before getting to the final decisions	−.35	.05	1.24	1.25	.68	−.28	−.35	−.06	.62
4. In preparation for some deadlines, I often waste time by doing other things	−.67	.05	.95	1.03	.74	−.58	−.69	−.11	.23
5. Even jobs that require little else except sitting down and doing them, I find that they seldom get done for days	−.47	.04	.84	.85	.78	−.59	−.44	.16	.09
6. I often find myself performing tasks that I had intended to do days before	−.85	.05	.70	.72	.80	−.89	−.85	.03	.73
7. I am continually saying "I'll do it tomorrow"	−.45	.05	.91	.94	.76	−.45	−.45	.00	.78
8. I generally delay before starting on work I have to do	−.46	.05	.86	.88	.77	−.51	−.46	.05	.48
9. I find myself running out of time	−.57	.04	1.17	1.17	.70	−.45	−.59	−.14	.12
10. I don't get things done on time	1.29	.04	1.06	1.13	.69	1.32	1.29	−.03	.77
11. I am not very good at meeting deadlines	1.47	.04	1.14	1.18	.67	1.37	1.50	.13	.28
Person									
Mean	.07	.41	1.00	1.02					
S.D.	1.48	.09	.58	.65					
Min.	−4.66	1.04	3.81	7.10					
Max.	4.75	.37	.07	.06					

properties at the item level of the PPS using an IRT framework. Finally, to use the results from IRT analyses to create cutoff scores that could be used by clinicians and researchers to discriminate between low, medium, and high levels of procrastination.

Regarding the first objective, results showed that among the competing CFA models, the bifactor model had better fit statistics across all indices. This result indicates that a significant amount of unique variance that is not explained by the specific factors (i.e., "decisional delay," "implemental delay" and "lateness/timeliness") can be accounted for by a general core factor (i.e., dysfunctional delay) of procrastination (Steel 2010). Similar results have been found by Rebetez et al. (2014) and Rozental et al. (2014) where the authors have found that the specific factors depend on a common general construct consistent with dysfunctional delay. In this study, the three-factor model better fits the data than the two-factor

model. These results are in line with those found in the validation article of the PPS in six European countries (Svartdal et al. 2016) and later replicated in an independent sample of adults from the general population (Svartdal and Steel 2017). However, in the latter two studies, researchers have not tested for a higher-order model nor a bifactor model. In the three-factor model, the items are grouped based on the original scale (Steel 2010) from which the PPS was created, which could be partly attributed to similar wording of the items (i.e., method effect). Future studies should focus on testing the incremental values of these three factors to add further evidence on the validity of this model.

Regarding the second objective, results showed that at the person level, the majority of the participants were within the range of latent trait level assessed by the scale. At the item level, results showed that the PPS provides relatively little information for participants with low and high attribute levels

Fig. 1 Densities of person theta and item difficulty

```
     -5    -4    -3    -2    -1     0     1     2     3     4     5
     |-----+-----+-----+-----+-----+-----+-----+-----+-----+-----|   NUM   ITEM
     1              1  :     2  :  3  :  4  :     5        5    |  11    Not good deadlines
     1          1     :    2   :  3  :    4   :        5  5    |  10    Don't get things done on time
     |                                                          |
     1          1    :     2    :   B  :      4     :    5  5  |   1    Delay making decisions
     |                                                          |
     |                                                          |
     1      1    :     2    :    3   :       4     :      5    5  |   2    Delay acting upon decision
     |                                                          |
     1     1   :      2       3   :    4     :    5         5  |   3    Waste time on trivial matters
     1    1   :       2         3   :   4    :      5         5  |   7    I'll do it tomorrow
     1      1   :       2        3  :   4       :    5          5  |   8    Delay starting work
     1       1   :    2         3  :   4     :    5           5  |   5    Even little things don't get done
     1    1    :      2        3  :   4     :    5            5  |   9    Running out of time
     1      1   :     2   :   3   :    4     :      5          5  |   4    Waste time doing other things
     1    1   :     2  :    3   :    4     :      5            5  |   6    Perform tasks intended days before
     |-----+-----+-----+-----+-----+-----+-----+-----+-----+-----|
     -5    -4    -3    -2    -1     0     1     2     3     4     5

                 1 1112122223522437333333222111111
     6 4       6 9 6 9  262163791070725011157770139505052 9 7 5  2     55  PERSON
                 T         S        M      S         T                    STATISTICS
     0              10  20 30 40 50 70 80 90                        99    PERCENTILE
```
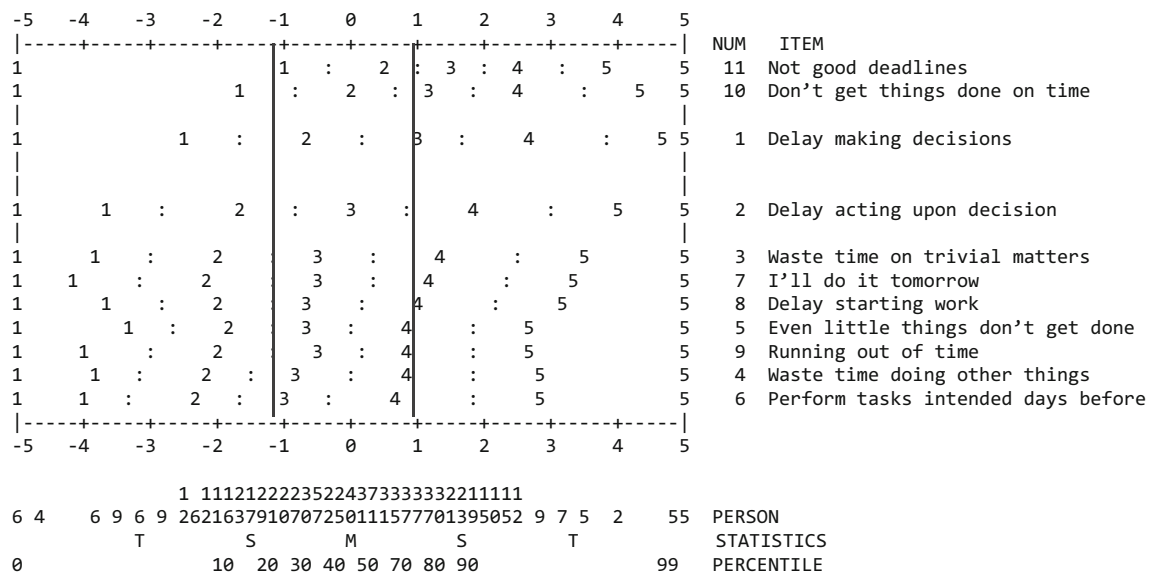
Fig. 2 Item map with rating scale model half-point threshold represented by " : "

(theta) whereas the scale provides great levels of information for participants in the middle range of the attribute. These results are in line with IRT analyses conducted among an international sample of adults from the general population across different translations of the PPS (Svartdal and Steel 2017). Items 11 and 10 were the two most difficult to endorse by the participants, indicating that these items have the most discriminating power. Interestingly, invariance analyses consistently revealed that these items were the least invariant across six European cultures and occupations (Svartdal et al. 2016). Based on these results, it appears that the PPS does not discriminate well between people with low and high levels on the attribute. Items selected to construct the PPS tend to cluster in a way that reflects the middle range of scores which can limit the utility of the PPS for example in clinical contexts where having more precise information for people that display higher prevalence of procrastinatory behaviors is often needed (e.g., Rozental and Carlbring 2014; Svartdal and Steel 2017). Finally, DIF analyses between men and women showed no significant difference that the pattern of responses across the range of theta was similar between gender. This result is in line with those of measurement invariance from Svartdal et al. (2016) who found support for the invariance of the PPS across gender. Based on these results, it appears that the scores on the PPS are not gender-dependent.

Concerning the third objective, based on the results from IRT analyses, we were able to establish cutoff scores to differentiate between low, medium, and high levels of procrastination. These cutoffs were also converted into raw scores to facilitate their use among researchers and clinicians. These raw cutoff scores are 27 or below and 40 or higher for low and high levels of procrastination respectively. Everything in the middle is considered a medium level of procrastination

and should encompass the majority of the student population. We decided to use the 25th and 75th percentiles as cutoff points based on past research that has reported that 20 to 30% of students are chronic procrastinators (Steel and Ferrari 2013). Cutoffs can notably be used to screen for participants' levels of procrastination before entering a clinical trial or be used to create groups when conducting analyses to study the efficacy of interventions that aim to reduce procrastinatory behaviors. When interpreting the cutoff scores, it should be noted, however, that a higher score on the PPS does not indicate the presence of chronic procrastination (i.e., a maladaptive lifestyle of chronic behavioral patterns that result in negative consequences; Ferrari 2010) as the PPS does not assess chronic procrastination but rather the prevalence of dysfunctional delay (for ex., "I delay making decisions until it's too late").

This study contains some limitations that need to be acknowledged to fully appreciate its results. First, results were obtained from a student population, thus further research is needed to try to replicate these results among the general population especially regarding the use of cutoff scores. Second, this study used the French version of the PPS. Readers should, therefore, use caution when extending the present results to other languages and cultures. The third limitation pertains to the fact that although the sample used in this study was quite large, it is not possible to know if it is representative of the general body of students. Notably, the sample was predominantly female which may indicate a slight recruitment bias toward female students in the research projects from which this sample was created. Although no DIF difference was found in this study, large-scale studies have reported lower procrastination scores for female relative to men (for ex., Steel and Ferrari 2013). It is thus possible that the cutoff scores

may have been higher with more male in the sample. Moreover, no data were available regarding the presence or not of a medical diagnostic such as a learning disability. Past research has shown that students with a learning disability reported higher levels of procrastination than their peers with no learning disability (Hen and Goroshit 2014). Notably, it would have been interesting to test DIF between students with and without learning disabilities.

In conclusion, it has been suggested that the PPS assesses distinct components related to procrastinatory behaviors (i.e., arousal, avoidant, and decisional). Results from this study showed that there is an overarching general core factor of procrastination (i.e., dysfunctional delay) encompassing those different components. The PPS has been supported by several studies as a valid measure of procrastinatory behaviors across different cultures and contexts. Using IRT analyses, this study further supports the validity of the PPS as a measure of procrastination most precise among people with average levels of procrastination. Finally, this study used IRT to create cutoff scores to discriminate between low, medium, and high levels of procrastination. Cutoff scores were converted into raw scores to facilitate their use among researchers and clinicians working with the PPS.

Compliance with Ethical Standards

Conflict of Interest The authors declared no conflict of interest.

References

American Educational Research Association, American Psychology Association & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ, US.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions, 154*, 136.

Chen, F. F., Hayes, A., Carver, C. S., Laurenceau, J., & Zhang, Z. (2012). Modeling general and specific variance in multifaceted constructs: A comparison of the bifactor model to other approaches. *Journal of Personality, 80*(1), 219–251.

Ellis, A., & Knaus, W. J. (2002). *Overcoming procrastination* (Revised ed.). New York, NY, US: New American Library.

Embretson, S. E., & Reise, S. P. (2000). *Multivariate applications books series. Item Response Theory for Psychologists*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Ferrari, J. R. (2010). *Still procrastinating: The no regrets guide to getting it done*. John Wiley & Sons.

Hen, M., & Goroshit, M. (2014). Academic procrastination, emotional intelligence, academic self-efficacy, and GPA: A comparison between students with and without learning disabilities. *Journal of Learning Disabilities, 47*(2), 116–124.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new

alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Jaccard, J., Wan, C. K., & Jaccard, J. (1996). *LISREL approaches to interaction effects in multiple regression*. Sage.

Kim, K. R., & Seo, E. H. (2015). The relationship between procrastination and academic performance: A meta-analysis. Personality and Individual Differences, 82(0), 26–33. https://doi.org/10.1016/j.paid.2015.02.038.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Klingsieck, K. B. (2013a). Procrastination in different life-domains: Is procrastination domain specific? *Current Psychology, 32*(2), 175–185.

Klingsieck, K. B. (2013b). Procrastination: When good things don't come to those who wait. *European Psychologist, 18*(1), 24–34. https://doi.org/10.1027/1016-9040/a000138.

Lay, C. H. (1986). At last, my research article on procrastination. *Journal of Research in Personality, 20*(4), 474–495. https://doi.org/10.1016/0092-6566(86)90127-3.

Linacre, J. M. (2014). *Winsteps (version 3.81. 0)*. Beaverton, Oregon: Winsteps. Com.

Mann, L. (1982). *Decision-making questionnaire*. Flinders University of South Australia: Unpublished manuscript.

McCown, W., Johnson, J., & Petzel, T. (1989). Procrastination, a principal components analysis. *Personality and Individual Differences, 10*(2), 197–202.

Pychyl, T. A., & Flett, G. L. (2012). Procrastination and self-regulatory failure: An introduction to the special issue. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 30*(4), 203–212.

Rebetez, M. M. L., Rochat, L., Gay, P., & Van der Linden, M. (2014). Validation of a french version of the pure procrastination scale (pps). *Comprehensive Psychiatry, 55*, 1442–1447. https://doi.org/10.1016/j.comppsych.2014.04.024.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667–696.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software, 48*(2), 1–36.

Rozental, A., & Carlbring, P. (2014). Understanding and treating procrastination: A review of a common self-regulatory failure. *Psychology, 5*(13), 1488–1502.

Rozental, A., Forsell, E., Svensson, A., Forsström, D., Andersson, G., & Carlbring, P. (2014). Psychometric evaluation of the Swedish version of the pure procrastination scale, the irrational procrastination scale, and the susceptibility to temptation scale in a clinical population. *BMC Psychology, 2*(1), 54.

Schouwenburg, H. C., Lay, C. H., Pychyl, T. A., & Ferrari, J. R. (2004). *Counseling the procrastinator in academic settings*. Washington, DC, US: American Psychological Association. https://doi.org/10.1037/10808-000.

Sirois, F., & Pychyl, T. (2013). Procrastination and the priority of short-term mood regulation: Consequences for future self. *Social and Personality Psychology Compass, 7*(2), 115–127.

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin, 133*(1), 65–94. https://doi.org/10.1037/0033-2909.133.1.65.

Steel, P. (2010). Arousal, avoidant and decisional procrastinators: Do they exist? *Personality and Individual Differences, 48*(8), 926–934. https://doi.org/10.1016/j.paid.2010.02.025.

Steel, P., & Ferrari, J. (2013). Sex, education and procrastination: An epidemiological study of procrastinators' characteristics from a global sample. *European Journal of Personality, 27*(1), 51–58. https://doi.org/10.1002/per.1851.

Svartdal, F. (2017). Measuring procrastination: Psychometric properties of the Norwegian versions of the irrational procrastination scale

(IPS) and the pure procrastination scale (PPS). *Scandinavian Journal of Educational Research, 61*(1), 18–30.

Svartdal, F., & Steel, P. (2017). Irrational delay revisited: Examining five procrastination scales in a global sample. *Frontiers in Psychology, 8*, 1927.

Svartdal, F., Pfuhl, G., Nordby, K., Foschi, G., Klingsieck, K. B., Rozental, A., et al. (2016). On the measurement of procrastination: Comparing two scales in six European countries. *Frontiers in Psychology, 7*, 1307.

Team, R. C. (2016). *R: A language and environment for statistical computing* (p. 2014). Vienna: R Foundation for Statistical Computing.