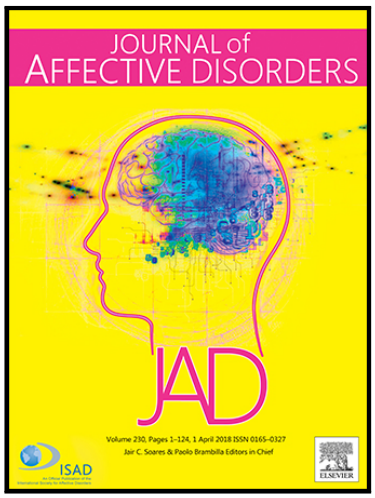
View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Repositorio Institucional de la Universidad de Alicante

Accepted Manuscript



A reliability generalization meta-analysis of the Child and Adolescent Perfectionism Scale

María Vicent , María Rubio-Aparicio , Julio Sánchez-Meca , Carolina Gonzálvez

PII: S0165-0327(18)31781-6
DOI: https://doi.org/10.1016/j.jad.2018.11.049
Reference: JAD 10267

To appear in: *Journal of Affective Disorders*

Received date: 14 August 2018
Revised date: 20 September 2018
Accepted date: 3 November 2018

Please cite this article as: María Vicent , María Rubio-Aparicio , Julio Sánchez-Meca , Carolina Gonzálvez , A reliability generalization meta-analysis of the Child and Adolescent Perfectionism Scale, *Journal of Affective Disorders* (2018), doi: https://doi.org/10.1016/j.jad.2018.11.049

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Highlights:

The average reliability was .87 for the original CAPS total score.

The mean alpha values were .84 and .83, respectively, for SPP and SOP subscales.

The original version of the CAPS can be employed with general research purposes.

The O'Connor's version of the CAPS must be used only for explanatory research.

The reliability induction rate was 29.8%.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

A reliability generalization meta-analysis of the Child and Adolescent Perfectionism Scale

Abstract

Background: Perfectionism is a prevalent disposition of personality involved in the development and maintenance of a wide range of psychological disorders. The Child and Adolescent Perfectionism Scale (CAPS) is the most usually applied test to assess perfectionism in children and adolescents. This study aimed: (a) to conduct a reliability generalization meta-analysis to estimate the average reliability of the CAPS scores and to search for characteristics of the studies that may explain the variability among reliability estimates, and (b) to estimate the reliability induction rate of the CAPS.

Method: An exhaustive search allowed to select 56 studies that reported alpha coefficients with the data at hand for the CAPS.

Results: The average alpha coefficients were .87, .84 and .83, respectively for the CAPS total score and its two subscales, Socially Prescribed Perfectionism (SPP) and Self-Oriented Perfectionism (SOP). Regarding O'Connor's version, the average reliability coefficients were .82, .74 and .73, respectively, for SPP, SOP-Critical and SOP-Strivings. Some study characteristics (ethnicity, language, mean age and standard deviation of the scores, psychometric vs applied) showed a statistical association with the reliability coefficients of SPP and SOP. The reliability induction rate was 29.8%.

Limitations: Due to the scarcity of studies, we could not examine the reliability scores of other versions of the CAPS and test-retest reliability.

Conclusions: In terms of reliability, the original version of the CAPS present better results than O'Connor's version. The original version of the CAPS is a reliable instrument to be employed with general research purposes, but not for clinical practice.

Keywords: Meta-analysis, reliability generalization, Child and Adolescent Perfectionism Scale.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Perfectionism can be defined as "a multidimensional personality disposition characterized by striving for flawlessness and setting exceeding high standards of performance accompanied by overly critical evaluations of one's behavior" (Stoeber, 2018a, p. 3). It is a stormy worldview that constitute a psychological vulnerability factor of clinical relevance, predisposing to the development and maintenance of lot of problems (Hewitt, Flett and Mikail, 2017). Likewise, far from being an exclusive disposition of adulthood, perfectionism is closely related with several disorders, such as anxiety, depression, Obsessive Compulsive Disorder and Eating Behavior Disorders, in child and adolescent population (Morris and Lomax, 2014). In fact, it is deemed that three out of ten young people present maladaptive forms of perfectionism; a rate that increases considerably when other more moderate forms are taken into account (Sironic and Reeve, 2015). On the other hand, Hong et al. (2017) concluded that maladaptive perfectionist trajectories emerge at the beginning of formal education, reflecting children's reactions to a prevalent culture that excessively values academic excellence. It is not surprising, therefore, the growing interest in research about perfectionism in samples of children and adolescents.

Leone and Wade (2017) conducted a systematic review on the psychometric properties of the scales used to measure perfectionism in the population under 15 years old. Concretely, four specific measures of child perfectionism were identified: (a) *The Adaptive-Maladaptive Perfectionism Scale* (AMPS; Rice and Preusser, 2002), (b) *The Children's Disfunctional Attitudes Scale* (CDAS; Allessandro and Abela, unpublished results), (c) *The Perfectionistic Self-Presentation Scale-Junior Form* (PSPS-JR; Hewitt et al., 2011), and (d) *The Child and Adolescent Perfectionism Scale* (CAPS; Flett et al., 2016). Authors concluded that the CAPS was the most advisable scale of the four, not only because it has relatively strong psychometric properties, but also because of its wide use and comparative data availability. In effect, the CAPS is currently the most used instrument of child and adolescent perfectionism (García-Fernández et al., 2016), having being applied in children and adolescents age 8 and over from several countries, mostly English-speaking, such as Canada (Flett et al., 2016), United States (e.g., Affrunti and Woodruff-Borden, 2017), United Kingdom (e.g., Kerr et al., 2016) and Australia (e.g., Ferrari et al., 2018), but also in population from Spain (e.g., Vicent, Inglés, Sanmartín et al., 2017b), Israel (e.g., Freudenstein et al., 2012), Portugal (e.g., Bento et al., 2017), Romania (e.g., Damian et al., 2017), Turkey (e.g., Uz-Bas and Siyez, 2010), France (e.g., Douilliez and Hénot, 2013), China (e.g., Yang et al., 2015),

Ecuador (e.g., Vicent, Inglés, Gonzálvez et al., 2017a), etc. This fact may create some confusion, since the validation of the CAPS was not definitively published until a few years ago, despite having been used for almost two decades since it was cited for the first time as an unpublished manuscript by Hewitt et al. (1997).

The relevance of the CAPS is partly due to the fact that it was developed by one of the research groups with the greatest impact in the field of perfectionism on the bases of the scale for adults of these same authors (i.e., *Multidimensional Perfectionism Scale*, Hewitt and Flett, 2004). The original version of the test consists of a 5-point Likert response scale and 22 items structured around two dimensions: Self-Oriented Perfectionism (SOP; 12 items) which measures the motivation and efforts to be a perfectionist as well as the tendency to self-criticize; and Socially Prescribed Perfectionism (SPP; 10 items) that captures the belief about the perfectionist demands of the environment. The authors also estimated the reliability of the scale across different populations, finding fluctuations between α = .68 and .82 for SOP and between .68 and .89 for SPP. Test-retest reliability was also calculated for intervals of one, three and five years, ranging these values between r = .65 and .40 for SOP and between .35 and .59 for SPP. From our knowledge, seven additional psychometric studies on CAPS have been published (Bento et al., 2014; Douilliez and Hénot, 2013; McCreary et al., 2004; Nobel et al., 2012; O'Connor et al., 2009a; Uz-Baş and Siyez, 2010; Yang et al., 2015).

All of them eliminated some items, with the exception of the Portuguese validation (Bento et al., 2014) that keeps the original scale intact. However, the studies of McCreary et al. (2004), O'Connor et al. (2009a) and Nobel et al. (2012), not only dispense with certain items but they also question the two-dimensional structure of the scale when considering that SOP dimension is better conceptualized by dividing its items into two independent dimensions called Self-Oriented Perfectionism Critical (SOP-C) and Self-Oriented Perfectionism-Striving (SOP-S). These two dimensions refers to self-criticism perfectionism and strivings to reach perfection, respectively. In this way, a new three-dimensional structure of the CAPS is proposed (i.e., SPP, SOP-C and SOP-S). Lastly, there is a Chinese validation of the CAPS consisting of 16 items of the original 22 and three items newly created, structuring all of them in four dimensions: Socially Prescribed Perfectionism Positive, Socially Prescribed Perfectionism Negative, Self-Oriented Perfectionism Positive and Self-Oriented Perfectionism Negative.

In terms of internal consistency, these additional psychometric studies offered good levels of reliability, Cronbach's alpha, for the SPP dimension, ranging between .82 and .86. Nevertheless, taken into account the Nunnally's criterion (1987), who established a minimum value of .70 to consider that a reliability coefficient is acceptable, not all psychometric studies obtained adequate levels of reliability for SOP, SOP-C and SOP-S. Specifically, values ranged from .64 to .83, from .66 to .74, and from .58 to .78, respectively. In contrast, regarding the temporal reliability, those studies that provided data on the test-retest obtained acceptable values, higher than .60 in all cases, with the exception of the Portuguese validation, whose test-retest level was .59 for the SOP dimension. These data show the existence of considerable fluctuations in the reliability levels depending on the characteristics of the employed sample. Meyer defines internal consistency reliability as "the extent to which test scores are consistent with another set of test scores produced from a similar process" (2010, p. 9). It is a psychometric property that must be taken into account in any study because it determines the validity of the conclusions obtained (Nunnally, 1982). However, there is a fairly widespread belief that reliability is an inherent property of an instrument (Sánchez-Meca et al., 2013). Thus, it is common in research to find studies in which either reliability estimates of the measures used are not provided, or the reliability coefficients obtained in previous studies are cited; generally the original validation of the scale (Sánchez-Meca et al., 2013). It has been coined with the name of reliability induction (Vacha-Haase et al., 2000), and it is an erroneous practice because, as mentioned, reliability is a property of the scores of a test for a particular sample of participants. Therefore, it is not an immutable property, but it can vary depending on different factors, such as the characteristics of the sample, the version of the test used, etc. According to Shields and Caruso (2004), and Sánchez-Meca et al. (2017), it is possible to distinguish two types of reliability induction: (a) by omission, that is, when the authors make no reference to the reliability of the test, or (b) by report, when reliability estimates from previous studies are mentioned. In turn, the induction by report may be exact or vague, respectively, depending on whether or not accurate estimates of reliability are provided.

The Reliability Generalization (RG) is a meta-analytical approach that emerges as a criticism of the widespread practice of induction of reliability. The purpose of this method is to estimate the average reliability of the scores of a given test, as well as to determine the variability of the reliability coefficients reported by the different studies

that have used this test. Moreover, if the variability is very high, another aim is to explore which characteristics of the studies may be statistically associated to the reliability estimates (Henson and Thompson, 2002; Rodríguez and Maeda, 2006; Sánchez-Meca et al., 2013).

The purpose of this research was to conduct an RG meta-analysis of those empirical studies that have applied the CAPS. The specific aims of this study were: (a) to calculate the average reliability of the CAPS dimensions scores to have an approximate estimate of their overall reliability; (b) to identify which characteristics of the studies may influence the variability of the reliability coefficients; and (c) to propose a predictive model to estimate the expected reliability of the CAPS according to the characteristics of the studies. Likewise, (d) the reliability induction rate of the CAPS was also estimated. Finally, in order to assess the extent to which the results of our RG meta-analysis can be generalized, we compared the characteristics of the studies that induced the reliability with those that provided some reliability coefficient with the data at hand.

Method

Selection criteria

The following criteria were considered to include each study in the meta-analysis: (a) being an empirical research where the original version of the CAPS (Flett et al., 2016) or any of its adaptations or versions were applied; (b) being written in English, Spanish or French; (c) being published and evaluated by experts; (d); employing any type of target population (community or clinical); (e) using a sample of at least 10 participants; (f) and reporting any reliability estimate of the CAPS or any of its subscales (internal consistency, test-retest) with the data at hand. The same criteria were considered for selecting studies that induced reliability, with the exception of (e) and (f).

Searching for the studies and selection process

The following data bases were consulted: Web of Science, Scopus, PsycINFO and ProQuest. The research strategy employed was: *"Child-Adolescent Perfectionism Scale" or "Child and Adolescent Perfectionism Scale" or (CAPS and perfectionism).* The search period covered from 1997 (date of publication of the first study that have used the CAPS) to march 2018.

ACCEPTED MANUSCRIPT

Figure 1 shows a flowchart describing the selection process of the studies. A total of 214 references were obtained, out of which 130 were removed for different reasons. Of the remaining 84 empirical studies, 59 reported some reliability coefficient whereas the other 25 induced the reliability.

INSERT HERE FIGURE 1

Data extraction

The following characteristics of the studies were extracted: (a) mean and standard deviation of CAPS (for total score and subscales), (b) CAPS adaptation (original, O'Connor, Portuguese adaptation), (c) language of the scale/adaptation, (d) study focus (psychometric vs. applied), (e) continent where the study was carried out, (f) target population (community, clinical), (g) type of disorder (in case of clinical sample), (h) mean age of the sample, (i) gender (% male), (j) ethnicity (% Caucasian), (k) financial source of the study, (l) year of the study, and (m) conflict of interest declaration. These characteristics were extracted from studies that reported any reliability estimate with the data at hand. In addition, such characteristics as the target population, mean and standard deviation of the CAPS and subscales, mean age, gender, and ethnicity were also extracted from the studies that induced reliability. This enabled us to compare the characteristics of the studies that induced and reported reliability estimates, with the purpose of examining the extent to which our meta-analytic results could be generalized to the total population of studies that applied the CAPS, regardless of whether they induced or reported reliability estimates.

To assess the reliability of the coding process of the study characteristics, all studies were doubly coded by two independent coders, both psychologists with PhD in psychology. Results were highly satisfactory, with kappa coefficients for qualitative characteristics ranging between .82 and 1 (M = .93), and intra-class correlations for continuous variables yielding values between .88 and 1 (M = .96).

Reliability estimates

In this RG study, the alpha coefficients were taken into account to assess internal consistency of the measures. Although, we intended to include in our meta-analysis test-retest temporal stability coefficients, the scarce references (e.g., Bento et al., 2014; Flett et al., 2016; O'connor et al., 2009a) that reported this type of reliability did not allow us

to carry out this analysis. Therefore, only alpha coefficients were extracted for the CAPS score and for each one of their subscales. In order to normalize their distribution and stabilizing their sampling variances, alpha coefficients, $\hat{\alpha}_i$, were transformed by means of Bonett's (2002) formula: $L_i = Ln(1-|\hat{\alpha}_i|)$, with *Ln* being the natural logarithm. The sampling variances were obtained by (Bonett, 2002):

$$V(L_i) = \frac{2J}{(J-1)(n_i-2)}, \quad (1)$$

with *J* being the number of items of the scale and n_i being the sample size of the study.

Statistical analyses

Separate meta-analyses were conducted for the alpha coefficients obtained from the total scale and for each of the two subscales of the original version of the CAPS.

To obtain summary statistics of alpha coefficients, a random-effects model was assumed (Borenstein et al., 2010). Thus, the alpha coefficients were weighted by the inverse variance, this defined as the sum of the within-study (Equation 1) and the between-studies variance, estimated by restricted maximum likelihood (López-López et al., 2013). In each meta-analysis, an average alpha coefficient and a 95% confidence interval were computed using the method proposed by Hartung (1999; see also Sánchez-Meca and Marín-Martínez, 2008; Sánchez-Meca et al., 2013). The heterogeneity exhibited by the alpha coefficients was assessed by constructing a forest plot and by calculating the *Q* statistic and the I^2 index. The *Q* statistic can be applied to test the homogeneity assumption among the alpha coefficients and I^2 values about 25%, 50%, and 75% can be considered as reflecting low, moderate, and large heterogeneity, respectively (Huedo-Medina et al., 2006).

For meta-analyses with at least 30 coefficients where evidence of heterogeneity was found, moderator analyses were performed through weighted ANOVAs for qualitative variables and meta-regressions for continuous variables. Mixed-effects models were assumed for these analyses, using the improved method proposed by Knapp and Hartung to test the statistical significance of the moderator variable (Knapp and Hartung, 2003; Rubio-Aparicio et al., 2017; Viechtbauer et al., 2015). In addition, the proportion of variance accounted for by the moderator variables was estimated with R^2 (López-López et al., 2014). Q_W and Q_E statistics were applied for testing the model misspecification of ANOVAs and meta-regressions, respectively.

ACCEPTED MANUSCRIPT

To facilitate the interpretation of the results, the average alpha coefficients, their confidence limits, and the slope estimates obtained with Bonett´s transformation were back-transformed to the original metric of alpha coefficient.

Last, the risk of publication bias was assessed applying the Egger test and constructing funnel plots with the trim-and-fill method (Duval & Tweedie, 2000).

All statistical analyses were carried out with *metafor* package in *R* (Viechtbauer, 2010).

Results

Mean reliability and heterogeneity

The present RG study was focused on the 59 studies that reported alpha coefficients with the data at hand. Of the 59 studies, three of them could not be included in our RG meta-analysis because they reported a range of alpha coefficients (Fairweather-Schmidt and Wade, 2015; Flett et al., 2012c; Vekas and Wade, 2017), or they employed other versions of the CAPS with not enough studies to be compared, this is the case of the French (Douilliez & Hénot, 2013) and Chinese (Yang et al., 2015) versions of the scale, or due to other reasons. Thus, the remaining 56 studies that reported alpha coefficients were included in our RG meta-analysis.

As several studies reported alpha coefficients for two or more different samples, the dataset of our RG meta-analysis was composed by a total of 64 independent samples.[1] The total number of participants was N = 28483 (min. = 37; max. = 2142), with a mean of 445 participants per sample (Median = 257; SD = 489). Out of the 64 independent samples, 59 (92.2%) were written in English, and the 5 remaining samples (7.8%) were written in Spanish. Regarding the location of the studies, five continents were represented in our RG study: North America with 26 samples (40.6%), Europe with 23 samples (35.9%), Asia with 8 samples (12.5%), Oceania with 5 samples (7.8%), and South America with 2 samples (3.1%). Finally, we found that 54 samples (84.4%) used the CAPS original version, 8 samples (12.5%) used the O'Connor version, and 2 samples (3.1%) used the Portuguese version. Separate meta-analyses for each one of these versions of the CAPS were carried out.

[1] The database with the 64 independent samples can be obtained from the corresponding author on request.

Table 1 presents the average alpha coefficients obtained for the total scores as well as for the two subscales of the original CAPS version. The 11 samples that reported alpha coefficients for the total score yielded a mean coefficient of .87 (95%CI: .84 and .90; 95% prediction interval (PI): .73 and .94). For the subscales, alpha coefficients were computed in 51 samples, yielding an overall estimate of .84 (95%CI: .82 and .85; 95%PI: .72 and .91) for the SPP subscale, and for the SOP subscale the average coefficient calculated with the 47 samples was of .83 (95%CI: .81 and .84; 95%PI: .66 and .91). The number of samples that applied the subscales was greatly larger than those that applied the total scale. For this reason, forest plots were only constructed for the SPP and SOP subscale scores (see Figures 2 and 3, respectively). Alpha coefficients for the total scale and subscales presented a statistically significant heterogeneity, with I^2 above 90%.

INSERT HERE TABLE 1, FIGURES 2 AND 3

Table 1 also presents the average alpha coefficients obtained for the three subscales of the O´Connor version. The reason for not including the total scale in the analyses was that only one study (Wojtowicz and Von Ranson, 2012) reported an alpha coefficient (α = .91) for that. The 7 estimates reported for SPP yielded a mean coefficient of .82 (95%CI: .76 and .86; 95%PI: .62 and .92). SOP-C and SOP-S showed lower average reliability coefficients than the SPP subscale above described. Concretely, the 6 samples that reported an alpha coefficient for SOP-C yielded an overall estimate of .74 (95%CI: .65 and .80; 95%PI: .52 and .86) and the 6 estimates for SOP-S presented a mean of .73 (95%CI: .67 and .77; 95%PI: .59 and .82).

Finally, only two studies reported reliability coefficients for the total scale of the Portuguese version: α = .81 (Bento et al., 2014) and .88 (Bento et al., 2010).

Analysis of moderator variables

As alpha coefficients for the SPP and SOP subscales of the original version presented more than 30 reliability estimates, the analyses of moderator variables were carried out only for these subscales. Meta-regressions and ANOVAs were conducted for continuous and categorical variables, respectively, on transformed alpha coefficients separately for SPP and SOP.

ACCEPTED MANUSCRIPT

Regarding SPP, Table 2 presents the results of the simple meta-regression analyses for each continuous moderator variable. Out of the different moderators analyzed, only the percentage of Caucasian exhibited a positive, statistically relationship with alpha coefficients (p = .002), with a 20% of variance explained. The positive sign of the regression coefficient of this moderator variable indicated larger alpha coefficients as the proportion of Caucasian participants increased. The standard deviation of SPP scores reached a positive, marginally significant relationship with the reliability coefficients (p = .066), as psychometric theory predicts, indicating that the larger the standard deviation of SPP, the larger the reliability.

Table 3 shows the results of the ANOVAs applied on the alpha coefficients of SPP for each qualitative moderator variable. The language of the SPP version presented a statistically influence on the reliability estimates (p < .001) and a 39% of variance accounted for. However, due to the large number of different adaptations of the original CAPS (in English) to at least six different languages, this variable was dichotomized to "English" vs. "other" languages. In this case, although the proportion of variance explained for the moderator was slightly lower (R^2 = .20), statistically significant differences were also found (p = .003), with a higher overall reliability for the "English language" (mean = .85) than for "other languages" (mean = .81). The remaining qualitative moderator variables analyzed did not reach statistical significant.

INSERT HERE TABLES 2 AND 3

With regard to SOP, Table 4 presents the results of the simple meta-regression analyses for the continuous moderator variables. As psychometric theory predicts, a positive, statistically significant relationship between the standard deviation of SOP and the alpha coefficients was found (p = .001) with a 29% of variance explained. The mean age of the samples also exhibited a positive, statistically significant relationship with the reliability estimates (p = .029), with a 11% variance accounted for. Last, the year of the study showed a statistically significant relationship with the alpha coefficients (p < .001), with a percentage of variance explained of 30%. In particular, the publication year exhibited a negative relationship with the reliability coefficients, so that lower alpha coefficients were obtained in the most recently published studies.

Table 5 shows the results of the ANOVAs applied on the alpha coefficients of SOP for each qualitative moderator variable. Once again, both the language and the

language dichotomized as "English" vs. "others" presented a statistically significant relationship with the alpha coefficients (p = .003 and p = .023, respectively) with percentages of variance accounted for of 33% and 12%, respectively. In particular, when the language was dichotomized larger average reliability was found for "English language" (M = .84) than for "other languages" (M = .80). The study focus showed a statistically significant relationship with the reliability estimates (p = .033). Concretely, psychometric studies showed a larger average reliability (M = .84) than those applied studies with a substantive purpose (M = .78). Finally, the continent where the studies were carried out also exhibited a statistically significant relationship with the alpha coefficients (p = .026), with a 19% of variance accounted for, with larger average coefficients for studies conducted in Oceania, North America and Asia (M = .85, .85, and .83, respectively), and lower averages yielded by those conducted in Europe and South America (means = .79 and .79, respectively).

INSERT HERE TABLES 4 AND 5

Explanatory models

As can be seen in Tables 2-5, all Q_W and Q_E statistics reached statistical significance (with the exception of type of disorder for SPP subscale), indicating that all of the ANOVAs and simple meta-regressions were misspecified. With the purpose of finding a predictive model able to explain, at least, a large part of the variability among the reliability estimates, weighted multiple meta-regression analyses were applied. Separate explanatory models were fitted for the SPP and SOP subscales of the original CAPS version. The predictors included in the model were selected as a function of the results of the ANOVAs and simple meta-regressions previously conducted.

Table 6 presents the results of the explanatory model for SPP including the percentage of Caucasian and the language dichotomized as predictors. Due to missing data in some variables, the number of studies included in the model was k = 43. The full model exhibited a statistically significant relationship with the alpha coefficients (p = .002), with a 30% of variance accounted for. Out of the two predictors included in the model, the percentage of Caucasian exhibited a statistically significant relationship with the alpha coefficients (p = .002), once the influence of the other variable was controlled. However, the language dichotomized presented a marginally significant result (p = .053). Regarding the contribution in terms of proportion of variance increase of each

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

predictor to the multiple meta-regression model, both the percentage of Caucasian and the language dichotomized showed a similar increase in R^2 (ΔR^2 = 11% and ΔR^2 = 10%, respectively), once the remaining predictor was added to the model. As a counterpart, the model was missespecified as the Q_E test was statistically significant (p < .0001), thus suggesting that other study characteristics were affecting the alpha coefficients variability as well.

INSERT HERE TABLE 6

Regarding SOP, five predictors were included in the multiple meta-regression model: the standard deviation of the SOP subscale scores, the mean age, the year of the study, the language dichotomized, and the study focus. The results are shown in Table 7. Due to missing data in some predictors, the number of studies included in the model was k = 28. Once again, the full model reached a statistically significant result (p < .0001), with a 78% of variance accounted for. When testing individually the predictors, and once the influence of the other predictors was controlled, three of them showed a statistically significant relationship with alpha coefficients: the standard deviation of the SOP subscale scores (p < .0001), the mean age (p = .024), and the study focus (p = .001). Last, the increase in percentage of variance accounted for the standard deviation, the mean age, and the study focus (after incorporating the other one to the model) was 27%, 22%, and 24%, respectively. Once again, the model misspecification test was also statistically significant (p < .0001).

INSERT HERE TABLE 7

Publication bias

The publication bias was assessed through funnel plots applying the trim-and-fill method and Egger tests for the SPP and SOP subscales of the original CAPS version. Figures S1 and S2 of the Supplementary file present the funnel plots obtained for the SPP and SOP subscales, respectively. When the trim-and-fill method was applied on each funnel plot, no alpha coefficients were imputed in the left side of the graph. In addition, non-significant results for the interceptions for SPP and SOP subscales were obtained with the Egger test (p = .818 and p = .259, respectively). Thus, the presence of publication bias can be discarded as a threat to the meta-analytic results.

ACCEPTED MANUSCRIPT

Estimating Reliability Induction

Out of the 84 studies that applied the CAPS, 25 induced reliability from other studies, which implies a 29.8% of reliability induction for this scale (see Figure 1). Out of the 25 studies that induced the reliability, 7 (28%) omitted any reference to the CAPS scores reliability, whereas the remaining 18 studies (72%) induced the reliability from previous studies. In particular, of these 18 studies, 15 (60%) induced the reliability accurately (i.e., reporting specific estimates from previous studies), and 3 (12%) induced vaguely the reliability (not reporting specific estimates).

Comparing Studies Inducing and Reporting Reliability

A main purpose in an RG meta-analysis is to generalize their results to the population of studies that have used the CAPS. However, the analyses in an RG meta-analysis are carried out only with the studies that reported the reliability with the data at hand. Thus, the extent to which the results of an RG meta-analysis can be generalized will depend on the similitude between the composition and variability of the samples of the studies that induce and those that report the reliability. To accomplish this objective, a comparison of the characteristics of inducing and reporting studies (e.g., the means and standard deviations of SPP and SOP, the age, the percentage of males, and the percentage of Caucasians) was performed by means of *t*-tests. These comparisons were conducted separately for studies with non-clinical and clinical samples. The results are presented in Tables S1 and S2, respectively, of the Supplementary file.

As can be seen, no statistically significant differences were found between studies inducing and reporting reliability in none of the characteristics studied in both non-clinical and clinical samples.

Discussion

The purpose of the present RG meta-analysis was to estimate the average internal consistency reliability of the CAPS scores and to identify those characteristics of the studies that affect the reliability coefficients obtained in the applications of the scale.

As mentioned in the Introduction, there are different versions of the CAPS validated in different samples (Bento et al., 2014; Douilliez and Hénot, 2013; McCreary et al., 2004; Nobel et al., 2012; O'Connor et al., 2009a; Uz-Baş and Siyez, 2010; Yang et al., 2015). Thus, although initially any version of the CAPS was included in the

search, only studies that used the original version of the scale, composed by 22 items and two dimensions (SPP and SOP) and the version of 14 items and three dimensions (SPP, SOP-C and SOP-C) proposed by O'Connor et al. (2009a) were considered for the average calculation and the analysis of heterogeneity. This is because there were not a sufficient number of studies that applied the other versions of the scale and reported any reliability estimate with the data at hand.

With respect to the original version (Flett et al., 2016), the average reliability coefficients were .87, .84 and .83, respectively, for the total score and for the SPP and SOP subscales. According to Nunnally and Bernstein (1994), alpha coefficients greater than .70 can be considered acceptable for exploratory research. However, for general research purposes, coefficients higher than .80 are recommended, as well as higher than .90 for using a measure in clinical practice. Consequently, results for the internal consistency of the total score and both subscales of the original version of the CAPS showed that the reliability of the test is adequate to be used for research purposes, but not to make decisions in clinical practice.

Regarding the version proposed by O'Connor et al. (2009a), the average reliability coefficients were acceptable (.82) to use SPP in research, but not in the clinical setting. In contrast, the average coefficients for SOP-C (.74) and SOP-S (.73) showed that both subscales do not possess the minimum levels of reliability required to be used for clinical and general research purposes. So, their use should remain limited to exploratory research.

A high heterogeneity was observed in the reliability coefficients reported by the researches applying the original version of the CAPS, both for the total scale and for the SPP and SOP subscales. In the case of the O'Connor's version, a high heterogeneity observed for SPP contrasted with a lower variability for SOP-C and SOP-S. The moderator analysis allowed to know which characteristics of the studies exhibited a statistical relationship with the reliability coefficients for the two dimensions (i.e., PSP and PAO) of the original version of the CAPS. Based on those characteristics that significantly contributed to explain the reliability estimates variability, an explanatory model was proposed for each perfectionist dimension. Regarding PSP, results indicated that alpha coefficients obtained in the studies were larger as the percentage of Caucasian increased and when the language of the CAPS version was the English. This result is not surprising, given that the CAPS was originally validated in Caucasian English-speaking population. In fact, practically all the accumulated knowledge in terms of

ACCEPTED MANUSCRIPT

perfectionism is based on studies conducted in North American or English population, even though there are several investigations that warm about the existence of certain sociocultural factors that could affect the way in which perfectionism is manifested through cultures (DiBartolo and Rendón, 2012). For example, it is conceivable that SPP is less maladaptive in collectivistic cultures than in individualistic cultures (Stoeber, 2018b). Specifically, the role of family has been pointed as one of these possible factors that influence the perfectionistic behavior (Ortega et al., 2014; Vicent, Inglés, Gonzálvez et al., 2017a; Yoon and Lau, 2008).

On the other hand, as expected from the psychometric theory (Nunnally and Bernstein, 1994), alpha coefficients for SOP (and also marginally for SPP) obtained in the studies were larger as the standard deviation increased. Other relevant predictors of the variability of the reliability coefficients were the mean age, in a positive sense, and the study focus. That is, the coefficients are higher as the participants' average age increases and when the study is psychometric.

Finally, 29.8% of the studies included in this meta-analysis induced reliability, either by reporting a reliability estimate of previous studies, or by omission, i.e. not providing any reference to the reliability of the scale scores (Shields and Caruso, 2004). The number of studies that induced reliability according to our study contrasts with the high rate of reliability induction that usually characterizes research in general and that is around 75% of the research that apply a psychological measurement scale (Sánchez-Meca et al., 2015; Vacha-Haase et al., 2002). Perhaps, the constant criticisms and warnings from the RG approach (Sánchez-Meca et al., 2009) have contributed to make the researchers aware of the fact that reliability is not an inherent characteristic of the instrument, but must be analyzed whenever a test is administered.

With the purpose to generalize the results beyond the studies that reported reliability estimates with the data at hand, and testing the existence of "reporting bias" (Sterne et al., 2011), the composition and variability of the participants employed in the studies that reported reliability was compared separately in both clinical and non-clinical samples with those studies that induced reliability. Thus, taking into account the non-significant results as well as the low reliability induction rate found (29.8%), we can conclude that the results of our RG study can be generalized to all of the studies that have applied the CAPS, regardless of having induced or not reliability.

Limitations and future research

ACCEPTED MANUSCRIPT

This study has various limitations. Fist, due to the lack of empirical studies to be compared, it was not possible to examine the average reliability of the scores of other versions of the CAPS different from the original one and that proposed by O'Connor et al. (2009a). For the same reason, only those characteristics that explained the variability for SPP and SOP of the original version of the CAPS could be examined in the moderator analysis. Moreover, since not all studies provided information about the analyzed characteristics, the sample of studies examined varied depending on the moderator variable examined. Similarly, only the analysis of internal consistency was considered in this work, given the limited number of studies that reported other reliability coefficients, such as test-retest. The explanatory models proposed were misspecified, pointing towards the existence of other relevant moderator variables not taken into account in our meta-analysis. Finally, it must be remembered that although there is an homologous scale of the CAPS destined to adult population (i.e., The Multidimensional Perfectionism Scale, Hewitt and Flett, 2004) which also includes SPP and SOP among its subscales, results obtained in the present meta-analysis are only generalizable to those studies that apply the CAPS.

Conclusions and practical implications

In conclusion, we can affirm that the scores of the original version of the CAPS, for the total scale and for both SPP and SOP subscales, present an acceptable reliability for research purposes but not for making clinical decisions in a professional performance context. Likewise, the reliability coefficients reported for the scores on the SPP, SOP-C and SOP-S subscales of the CAPS version proposed by O'Connor et al. (2009a) show that SPP presents acceptable reliability for use in general research, whereas the use of the SOP-C and SOP-S subscales is only recommended for exploratory research. In the light of the outcomes, it is recommended to those researchers who are planning to apply the CAPS to turn to the original version proposed by Flett et al. (2016). That is because the original version offers better guarantees, in terms of reliability, than the 14-item version of O'Connor et al. (2009a). Nevertheless, researchers should take into account that it is possible that studies carried out in non-Caucasian and non-English speaking population report lower reliability coefficients for SPP. In this sense, it should be noted that it is necessary to perform the cross-cultural validation of a test before applying it to a different sample from the reference population used in the original validation. Keeping in mind that there are validations of the CAPS

ACCEPTED MANUSCRIPT

in North-American, French, Portuguese, Chinese, Turkish and Scottish population, it would be recommendable to carry out psychometric studies of the CAPS in Spanish and Latin American population, since after English, research in the Spanish-speaking population was the most numerous. Similarly, future studies should take into account that the reliability estimates for SOP scores are sensitive to the standard deviation of the subscale scores, to the age of the sample and to the purpose of the study (psychometric or applied).

Finally, although the reliability induction rate found in our study is much lower than that found by previous meta-analytic research, it is also worth noting that reliability induction is an erroneous practice that must be eradicated since it can cause errors in the estimation of the measures used (Sánchez-Meca et al., 2009).

Author statement

This manuscript describes original work and is not under consideration by any other journal.

Corresponding author:
María Rubio-Aparicio, Department of Basic Psychology and Methodology, Faculty of Psychology, University of Murcia, Campus de Espinardo, 30100 Murcia (Spain).
Email: maria.rubio1@um.es

Contributors

Dr. María Vicent conducted the searches, extracted the data, and wrote the manuscript. Dr. María Rubio-Aparicio participated performing statistical analysis and writing this manuscript. Prof. Julio Sánchez-Meca has designed, reviewed and supervised this research. Dr. Carolina Gonzálvez was involved in data collection and reviewed the manuscript.

Conflict of interest: None

Funding sources: Ministerio de Economía y Competitividad of the Spanish Government; Fondo Europeo de Desarrollo Regional (FEDER) Project No. PSI2016-77676- P.

Acknowledgements: None

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

References

*Affrunti, N.W., Gramszlo, C., Woodruff-Borden, J. 2016. Executive function moderates the association between fearful temperament and dimensions of perfectionism.
Pers. Individ. Dif. 89, 117-122. http://dx.doi.org/10.1016/j.paid.2015.10.012

*Affrunti, N.W., Woodreuff-Bordern, J. 2016. Negative affect and child internalizing symptoms: the mediating role of perfectionism. Child Psychiatry Hum. Dev. 47, 358-368. https://doi.org/10.1007/s10578-015-0571-x

*Affrunti, N.W., Woodruff-Borden, J., 2017. Emotional control mediates the association between dimensions of perfectionism and worry in children. Child. Psychiatry Hum. Dev. 48, 73-81. https://doi.org/10.1007/s10578-016-0654-3

*Asseraf, M., Vaillancourt, T. 2015. Longitudinal links between perfectionism and depression in children. J. Abnorm. Child Psychol. 43, 895-908.

*Bachner-Melman, R., Zohar, A.H., Ebstein, R.P., Bachar, E. 2007. The relationship between selflessness levels and the severity of anorexia nervosa symptomatology. Eur. Eat. Disord. Rev. 15(3), 213-220. http://dx.doi.org/10.1002/erv.740

*Bento, C., Pereira, A.T., Maia, B., Marques, M., Soares, M.J., Bos, S., Valente, J., Gomes, A., Azevedo, M.H.P., Macedo, A., 2010. Perfectionism and eating bihaviour in Portuguese adolescents. Eur. Eat. Disord. Rev. 18, 328-337.

Bento, C., Pereira, A.T., Roque, C., Tavares, J.M., Ferreira-Macedo, A.J., 2017. Longitudinal effects of an intervention on perfectionism in adolescents. Psicothema 29(3), 317-322. https://doi.org/10.7334/psicothema2016.223

*Bento, C., Pereira, A.T., Saraiva, J.M., Macedo, A., 2014. Children and adolescent perfectionism scale: validation in a Portuguese adolescent sample. Psicologia: Reflexão e Crítica. 27, 228-232. https://doi.org/10.1590/1678-7153.201427203

Bonett, D.G., 2002. Sample size requirements for testing and estimating coefficient alpha. J. Educ. Behav. Stat. 27, 335-340. https://doi.org/10.3102/10769986027004335

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2010. A basic introduction to fixed-effect and random-effects models for meta-analysis. Research Synthesis Methods 1, 97-111. https://doi.org/10.1002/jrsm.12.

*Castro, J., Gila, A., Gual, P., Lahortiga, F., Saura, B., Toro, J. 2004. Perfectionism dimensions in children and adolescents with anorexia nervosa. J. Adolesc. Health 35, 392-398. https://doi.org/10.1016/j.jadohealth.2003.11.094

ACCEPTED MANUSCRIPT

*Damian, L.E., Stoeber, J., Negru, O., Baban, A. 2013. On the development of perfectionism in adolescence: perceived parental expectations predict longitudinal increases in socially prescribed perfectionism. Pers. Individ. Dif. 55, 688-693. http://dx.doi.org/10.1016/j.paid.2013.05.021

*Damian, L.E., Stoeber, J., Negru, O., Baban, A. 2014a. Perfectionism and achievement goal orientations in adolescent school students. Psychol. Sch. 51(9), 960-971. http://dx.doi.org/10.1002/pits.21794

*Damian, L.E., Stoeber, J., Negru, O., Baban, A. 2014b. Positive and negative affect in adolescents: an investigation of the 2 × 2 model of perfectionism. cognition, brain, behavior. An Interdisciplinary Journal 18, 1-16.

Damian, L.E., Stoeber, J., Negru-Subtirica, O., Baban, A., 2017. On the development of perfectionism: the longitudinal role of academic achievement and academic efficacy. Journal of Personality 85(4), 565-577. https://doi.org/10.1111/jopy.12261

DiBartolo, P.M., Rendón, M.J., 2012. A critical examination of the construct of perfectionism and its relationship to mental health in Asian and African Americans using a cross-cultural framework. Clin. Psychol. 32(3), 139-152. https://doi.org/10.1016/j.cpr.2011.09.007

*DiBartolo, P.M., Varner, S.P. 2012. How children's cognitive and affective responses to a novel task relate to the dimensions of perfectionism. J. Rat-Emo. Cognitive-Behav. Ther. 30, 62-76. https://doi.org/10.1007/s10942-011-0130-8

Douilliez, C., Hénot, E., 2013. Mesures du perfectionnisme chez l'adolescent: validation des versions francophones de deux questionnaires. [Measures of perfectionism in adolescents: validation of the French versions of two questionnaires]. Can. J. Behav. Sci. 45, 64-71. https://doi.org/10.1037/a0022686

Duval, S. J., Tweedie, R. L., 2000. Trim and fill: A simple funnel-plot based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Fairweather-Schmidt, A.K., Wade, T.D., 2015. Piloting a perfectionism intervention for pre-adolescent children. Behav. Res. Ther. 73, 67-73. https://doi.org/10.1016/j.brat.2015.07.004

*Ferrari, M., Yap, K., Scott, N., Einstein, D.A., Ciarrochi, J., 2018. Self-compassion moderates the perfectionism and depression link in both adolescence and adulthood. Plos One 21, 1-19. https://doi.org/10.1371/journal.pone.0192022

*Flett, G.L., Coulter, L.M., Hewitt, P.L. 2012a. The perfectionistic self-presentation scale–junior form: psychometric properties and association with social anxiety in early adolescents. Can. J. Sch. Psychol. 27(2), 136-149.

*Flett, G.L., Coulter, L.M., Hewitt, P.L., Nepon, T. 2011a. Perfectionism, rumination, worry, and depressive symptoms in early adolescents. Can. J. Sch. Psychol. 26(3), 159-176. https://doi.org/10.1177/0829573511422039

*Flett, G.L., Druckman, T., Hewitt, G.L., Hewitt, P.L., Wekerle, C. 2012b. Perfectionism, coping, social support, and depression in maltreated adolescents. J. Rat-Emo. Cognitive-Behav. Ther. 30, 118-131. https://doi.org/10.1007/s10942-011-0132-6

*Flett, G.L., Hewitt, P.L., Besser, A., Su, C., Vaillancourt, T, Boucher, D., Munro, Y., Davidson, L.A., Gale, O., 2016. The Child-Adolescent Perfectionism Scale: development, psychometric properties, and associations with stress, distress, and psychiatric symptoms. J. Psychoeduc. Assess. 34(7), 634-652. https://doi.org/10.1177/0734282916651381

*Flett, G.L., Hewitt, P.L. Cheng, W.M.W. 2008. Perfectionism, distress, and irrational beliefs in high school students: analyses with an abbreviated survey of personal beliefs for adolescents. J. Rat-Emo. Cognitive-Behav. Ther. 26, 194-205. https://doi.org/10.1007/s10942-007-0066-1

*Flett, G.L., Panico, T., Hewitt, P.L. 2011b. Perfectionism, type A behavior, and self-efficacy in depression and health symptoms among adolescents. Curr. Psychol. 30, 105-116. https://doi.org/10.1007/s12144-011-9103-4

*Freudenstein, O., Valevski, A., Apter, A., Zohar, A., Shoval, G., Nahshoni, E., Weizman, A., Zalsman, G., 2012. Perfectionism, narcissism, and depression in suicidal and nonsuicidal adolescent inpatients. Compr. Psychiatry. 53, 746-752. https://doi.org/10.1016/j.comppsych.2011.08.011

*García-Fernández, J.M., Inglés, C.J., Vicent, M., Gonzálvez, C., Gómez-Núñez, M.I. Poveda-Serra, P., 2016. Perfeccionismo durante la infancia y la adolescencia. Análisis bibliométrico y temático (2004-2014). [Perfectionism during childhood and adolescence. Bibliometric and thematic analysis (2004-2014)]. Revista

Iberoamericana de Psicología y Salud 7(2), 79-88. https://doi.org/10.1016/j.rips.2016.02.001

*Goodwin, H., Haycraft, E., Meyer, C. 2014. Psychological risk factors for compulsive exercise: a longitudinal investigation of adolescent boys and girls. Pers. Individ. Dif. 68, 83-86. http://dx.doi.org/10.1016/j.paid.2014.03.048

*Goodwin, H., Haycraft, E., Willis, A.M., Meyer, C. 2011. Compulsive exercise: the role of personality, psychological morbidity, and disordered eating. Int. J. Eat. Disord. 44, 655-660. http://dx.doi.org/10.1002/eat.20902

*Greenaway, R., Howlin, P. 2010. Dysfunctional attitudes and perfectionism and their relationship to anxious and depressive symptoms in boys with autism spectrum disorders. J. Autism. Dev. Disord. 40, 1179-1187. http://dx.doi.org/10.1007/s10803-010-0977-z

*Guignard, J.H., Jacquet, A.Y., Lubart, T.I. 2012. Perfectionism and anxiety: a paradox in intellectual giftedness? PLoS One 7(7), e41043. https://doi.org/10.1371/journal.pone.0041043

Hartung, J., 1999. An alternative method for meta-analysis. Biom. J. 41, 901-916. https://doi.org/10.1002/(SICI)1521-4036(199912)41:8<901::AID-BIMJ901>3.0.CO;2-W

Henson, R.K., Thompson, B., 2002. Characterizing measurement error in scores across studies: Some recommendations for conducting «reliability generalization» studies. Meas. Eval. Couns. Dev. 35, 113-126.

Hewitt, P.L., Blasberg, J. S., Flett, G. L., Besser, A., Sherry, S. B., Caelian, C., Papsdorf, M., Cassels, T.G., Birch, S., 2011. Perfectionistic self-presentation in children and adolescents: development and validation of the Perfectionistic Self-Presentation Scale-Junior Form. Psychol. Assess. 23(1), 125-142. https://doi.org/10.1037/a0021147

*Hewitt, P.L., Caelian, C.F., Chen, C., Flett, G.L. 2014. Perfectionism, stress, daily hassles, hopelessness, and suicide potential in depressed psychiatric adolescents. J. Psychopathol. Behav. Assess. 36, 663-674. https://doi.org/10.1007/s10862-014-9427-0

Hewitt, P.L., Flett, G.L., 2004. Multidimensional Perfectionism Scale (MPS): Technical manual. Multi-Health Systems, Toronto.

Hewitt, P.L., Flett, G.L., Mikail, S.F., 2017. Perfectionism. A relational approach to conceptualization, assessment and treatment. Guildford, New York.

*Hewitt, P.L., Newton, J., Flett, G.L., Callander, L., 1997. Perfectionism and suicide ideation in adolescent psychiatric patients. J. Abnorm. Child. Psychol. 25(2), 95-101.

*Hill, A.P., Appleton, P.A. 2012. Perfectionism and athlete burnout in junior elite athletes: the mediating role of motivation regulations. J. Clin. Sport. Psychol. 6(2), 129-146.

Hong, R.Y., Lee, S.S.M., Chng, R.Z., Zhou, Y., Tsai, F., Tan, S.H., 2017. Developmental trajectories of maladaptive perfectionism in middle childhood. Journal of Personality 85(3), 409-422. https://doi.org/10.1111/jopy.12249

Huedo-Medina, T.B., Sánchez-Meca, J., Marín-Martínez, F., Botella, J., 2006. Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? Psychol. Methods 11, 193-206. https://doi.org/10.1037/1082-989X.11.2.193

*Huggins, L., Davis, M.C., Rooney, R., Kane, R. 2008. Socially prescribed and self-oriented perfectionism as predictors of depressive diagnosis in preadolescents. Australian Journal of Guidance & Counselling 2, 182-194. https://doi.org/10.1375/ajgc.18.2.182

*Inglés, C.J., García-Fernández, J.M., Vicent, M., Gonzálvez, C., Sanmartín, R. 2016. Profiles of perfectionism and school anxiety: a review of the 2 x 2 model of dispositional perfectionism in child population. Front. Psychol. 7, 1403. https://doi.org/10.3389/fpsyg.2016.01403

*Kerr, C., Watkins, B., Jones, F.W., 2016. Inflated responsibility and perfectionism in child and adolescent anorexia. Adv. Eat. Disord. 4(3), 309-314. https://doi.org/10.1080/21662630.2016.1217494

Knapp, G., Hartung, J., 2003. Improved tests for a random effects meta-regression with a single covariate. Statistics in Medicine, 22, 2693–2710. https://doi.org/10.1002/sim.1482

Leone, E.M., Wade, T.D., 2017. Measuring perfectionism in children: a systematic review of the mental health literature. Eur. Child Adolesc. Psychiatry 1-15. https://doi.org/10.1007/s00787-017-1078-8

López-López, J.A., Botella, J., Sánchez-Meca, J., Marín-Martínez, F., 2013. Alternatives for mixed-effects meta-regression models in the reliability generalization approach: A simulation study. J. Educ. Behav. Stat. 38, 443-469. https://doi.org/10.3102/1076998612466142

López-López, J.A, Marín-Martínez, F., Sánchez-Meca, J., van den Noortgate, W., Viechtbauer, W., 2014. Estimation of the predictive power of the model in mixed-effects meta-regression: A simulation study. Br. J. Math. Stat. Psychol. 67, 30-48. https://doi.org/10.1111/bmsp.12002

McCreary, B.T., Joiner, T.E., Schmidt, N.B., Ialongo, N.S., 2004. The structure and correlates of perfectionism in African American children. J. Clin. Child. Adolesc. Psychol. 33(2), 313-324. https://doi.org/10.1207/s15374424jccp3302_13

*McVey, G.L., Davis, R., Tweed, S., Shaw, B.F. 2004. Evaluation of a school-based program designed to improve body image satisfaction, global self-esteem, and eating attitudes and behaviors: a replication study. Int. J. Eat. Disord. 36, 1-11. http://dx.doi.org/10.1002/eat.20006

*McVey, G.L., Pepler, D., Davis, R., Flett, G.L., Abdolell, M. 2002. Risk and protective factors associated with disordered eating during early adolescence. J. Early Adolesc. 22(1), 75-95.

Meyer, J.M., 2010. Reliability. Oxford University Press, New York.

*Mitchell, J.H., Broeren, S., Newall, C., Hudson, J.L. 2013a. An experimental manipulation of maternal perfectionistic anxious rearing behaviors with anxious and non-anxious children. J. Exp. Child Psychol. 116(1), 1-18. http://dx.doi.org/10.1016/j.jecp.2012.12.006

*Mitchell, J.H., Newall, C., Broeren, S., Hudson, J.L. 2013b. The role of perfectionism in cognitive behaviour therapy outcomes for clinically anxious children. Behav. Res. Ther. 51, 547-554. http://dx.doi.org/10.1016/j.brat.2013.05.015

Morris, L., Lomax, C., 2014. Review: assessment, development, and treatment of childhood perfectionism: a systematic review. Child and Adolescent Mental Health 19(4), 225-234. https://doi.org/10.1111/camh.12067

*Nobel, R. J., Manassis, K., Wilansky-Traynor, P., 2012. The role of perfectionism in relation to an intervention to reduce anxious and depressive symptoms in children. J. Ration Emot. Cogn. Behav. Ther. 30, 77-90. https://doi.org/10.1007/s10942-011-0133-5

Nunnally, J.C., 1978. Psychometric Theory. McGraw-Hill, New York.

Nunnally, J.C., 1982. Reliability of measurement. In H.E. Mitzel (Ed.), Encyclopedia of educational research (pp.1589-1601). NewYork: Free Press.

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

*O'Connor, R.C., Dixon, D., Rasmussen, S., 2009a. The structure and temporal stability of the Child and Adolescent Perfectionism Scale. Psychol. Assess. 21(3), 437-443. https://doi.org/10.1037/a0016264

*O'Connor, R.C., Rasmussen, S., Hawton, K. 2009b. Predicting deliberate self-harm in adolescents: a six month prospective study. Suicide Life Threat. Behav. 39(4), 364-375.

*O'Connor, R.C., Rasmussen, S., Hawton, K. 2010. Predicting depression, anxiety and self-harm in adolescents: the role of perfectionism and acute life stress. Behav. Res. Ther. 48, 52-59. https://doi.org/1 0.1016/j.brat.2009.09.008

Ortega, N.E., Wang, K.T., Slaney, R.B., Hayes, J.A., Morales, A., 2014. Personal and familial aspects of perfectionism in Latino/a students. Counseling Psychologist, 42(3), 406-427.

*Pamies, L., Quiles, Y. 2014. Perfeccionismo y factores de riesgo para el desarrollo de trastornos alimentarios en adolescentes españoles de ambos géneros. Anales de Psicología, 30(2), 620-626. http://dx.doi.org/10.6018/analesps.30.2.158441

Rice, K., Preusser, K., 2002. The adaptive/maladaptive perfectionism scale. Meas. Eval. Couns. Dev. 34, 210-222

Rodríguez, M.C., Maeda, Y., 2006. Meta-analysis of coefficient alpha. Psychol. Methods 11, 306-322.

*Rodríguez-Jiménez, T., Blasco, L., Piqueras, J.A. 2014. Relación entre subtipos de sintomatología obsesivo compulsiva y de perfeccionismo en adolescentes. Psicol. Conductual 22(2), 203-217.

*Roxborough, H.M., Hewitt, P.L., Kaldas, J., Flett, G.L., Caelian, C.M., Sherry, S., Sherry, D.L. 2012. Perfectionistic self-presentation, socially prescribed perfectionism, and suicide in youth: a test of the perfectionism social disconnection model. Suicide Life Threat Behav. 42(2), 217-233. https://doi.org/10.1111/j.1943-278X.2012.00084.x

Rubio-Aparicio, M., Sánchez-Meca, J., López-López, J.A., Marín-Martínez, F., Botella, J., 2017. Analysis of categorical moderators in mixed-effects meta-analysis: Consequences of using pooled vs. separate estimates of the residual between-studies variances. Br. J. Math. Stat. Psychol. 70, 439-456. https://doi.org/10.1111/bmsp.12092

*Sailing, M., Ricciardelli, L.A., McCabe, M.P. 2005. A prospective study of individual factors in the development of weight and muscle concerns among preadolescent

children. J. Youth Adolesc. 34(6), 651-661. https://doi.org/10.1007/s10964-005-8953-1

Sánchez-Meca, J., López-López, J. A., López-Pina, J. A., 2013. Some recommended statistical analytic practices when reliability generalization (RG) studies are conducted. Br. J. Math. Stat. Psychol. 66, 402-425. https://doi.org/10.1111/j.2044-8317.2012.02057.x

Sánchez-Meca J., Marín-Martínez, F., 2008. Confidence intervals for the overall effect size in random-effects meta-analysis. Psychol. Methods 13, 31-48. https://doi.org/10.1037/1082-989X.13.1.31.

Sánchez-Meca J., Rubio-Aparicio M., Marín-Martínez F., López-Pina J.A., 2015. The phenomenon of reliability induction in social and health sciences. Paper presented at the XIV Congress of Methodology of Behavioral Science. Palma de Mallora, Spain.

Sánchez-Meca, J., Rubio-Aparicio, M., Núñez-Núñez, R.M., López-Pina, J., Martín-Martínez, F., López-López, J.A., 2017. A reliability generalization meta-analysis of the Padua Inventory of Obsessions and compulsions. Span. J. Psychol. 20(e70), 1-15. https://doi.org/10.1017/sjp.2017.65

Shields, A.L., Caruso, J.C., 2004. A reliability induction and reliability generalization study of the Cage Questionnaire. Educ. Psychol. Meas. 64, 254-270. https://doi.org/10.1177/0013164403261814

*Sironic, A., Reeve, R.A., 2015. A combined analysis of the Frost Multidimensional Perfectionism Scale (FMPS), Child and Adolescent Perfectionism Scale (CAPS), and Almost Perfect Scale-Revised (APSR): Different perfectionist profiles in adolescent high school students. Psychol. Assess. 27, 1471-1483.

*Soreni, N., Streiner, D., McCabe, R., Bullard, C., Swinson, R., Greco, A., Pires, P., Szatmari, P. 2014. Dimensions of perfectionism in children and adolescents with obsessive-compulsive disorder. J. Can. Acad. Child Adolesc. Psychiatry 23(2), 136-143.

Sterne, J.A.C., Sutton, A.J., Ioannidis, J.P.A., Terrin, N., Jones, D.R., Lau, J., Carpenter, J., Rücker, G., Harbord, R.M., Schmid, C.H., Tetzlaff, J., Deeks, J.J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D.G., Moher, D., Higgins, J.P., 2011. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. British Medical Journal, 343, d4002. https://doi.org/10.1136/bmj.d4002

Stoeber, J., 2018a. The psychology of perfectionism. An introduction, in: Stoeber J. (Ed.), The psychology of perfectionism: theory, research, applications. Routledge, London, pp. 3-16.

Stoeber, J., 2018b. The psychology of perfectionism: critical issues, open questions, and future directions, in: Stoeber J. (Ed.), The psychology of perfectionism: theory, research, applications. Routledge, London, pp. 333-352.

*Stoeber, J., Roche, D.L. 2014. Affect intensity contributes to perfectionistic self-presentation in adolescents beyond perfectionism. J. Rat-Emo. Cognitive-Behav. Ther. 32, 164- http://dx.doi.org/180. 10.1007/s10942-013-0176-x

*Stornelli, D., Flett, G.L., Hewitt, P.L. 2009. Perfectionism, achievement, and affect in children: a comparison of students from gifted, arts, and regular programs. Can. J. Sch. Psychol. 24(4), 267-283. https://doi.org/10.1177/0829573509342392

*Tong, Y., Lam, S.F. 2011. The cost of being mother's ideal child: the role of internalization in the development of perfectionism and depression. Soc. Dev. 20(3), 504-516. https://doi.org/10.1111/j.1467-9507.2010.00599.x

*Travers, L.V., Randall, E.T., Bryant, F.B., Conley, C.S., Bohnert, A.M. 2015. The cost of perfection with apparent ease: theoretical foundations and development of the effortless perfectionism scale. Psychol. Assess. 27(4), 1147-1159. http://dx.doi.org/10.1037/pas0000109

Uz-Bas, A., Siyez, D., 2010. Adaptation of the child and adolescent perfectionism scale to Turkish: the validity and reliability study. Elementary Education Online, 9(3), 898-909.

Vacha-Haase, T., Henson, R.K., Caruso, J.C., 2002. Reliability generalization: moving toward improved understanding and use of scores reliability. Educ. Psychol. Meas. 62(4), 562-569. https://doi.org/10.1177/0013164402062004002

Vacha-Haase, T., Kogan, L.R., Thompson, B., 2000. Sample compositions and variabilities in published studies versus those in test manuals: validity of score reliability inductions. Educ. Psychol. Meas. 60, 509-22.

Vekas, E.J., Wade, T.D., 2017. The impact of a universal intervention targeting perfectionism in children: An exploratory controlled trial. Br. J. Clin. Psychol. 56, 458-473. https://doi.org/10.1111/bjc.12152

*Vicent, M., Inglés, C.J., Gonzálvez, C., Sanmartín, R., García-Fernández, J.M. 2016. Perfeccionismo socialmente prescrito y los cinco grandes rasgos de la personalidad en niños españoles. EJIHPE 6(2), 107-118.

ACCEPTED MANUSCRIPT

*Vicent, M., Inglés, C.J., Gonzálvez, C., Sanmartín, R., Ortega-Sandoval, V.N., García-Fernández, J.M., 2017a. Testing the 2 x 2 model of perfectionism in Ecuadorian adolescent population. J. Health Psychol. 1-7. https://doi.org/10.1177/1359105317733536

*Vicent, M., Inglés, C.J., Sanmartín, R., Gonzálvez, C., García-Fernández, J.M., 2017b. Perfectionism and aggression: identifying risk profiles in children. Pers. Individ. Dif. 112, 106-112. doi: 10.1016/j.paid.2017.02.061

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36, 1-48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., López-López, J.A., Sánchez-Meca, J., Marín-Martínez, F., 2015. A comparison of procedures to test for moderators in mixed-effects meta-regression models. Psychol. Methods 20, 360-374. https://doi.org/10.1037/met0000023

*Vuyk, M.A. 2015. Perfeccionismo, autopresentación perfeccionista, y síntomas depresivos en adolescentes paraguayos según género y grado. Interam. J. Psychol. 49(2), 153-163.

*Wojtowicz, A.E., Von Ranson, K.M., 2012. Weighing in on risk factors for body dissatisfaction: A one-year prospective study of middle-adolescent girls. Body Image 9, 20-30. https://doi.org/10.1016/j.bodyim.2011.07.004

Yang, H., Hong, C., Tao, X., Zhu, L., 2015. Revising the Child and Adolescent Perfectionism Scale: a test of the four-factor structure in a Chinese sample. Meas. Eval. Couns. Dev. 30, 1-12. https://doi.org/10.1177/0748175615578733

Yoon J., Lau A.S., 2008. Maladaptive perfectionism and depressive symptoms among Asian American college students: Contributions of interdependence and parental relations. Cultur. Divers. Ethnic. Minor. Psychol. 14(2), 92-101.

*Zeifman, R.J., Atkey, S.K., Young, R.E., Flett, G.L., Hewitt, P.L., Golgberg, J.O. 2015. When ideals get in the way of self-care: perfectionism and self-stigma for seeking psychological help among high school students. Can. J. Sch. Psychol. 30(4), 273-287. https://doi.org/10.1177/0829573515594372

*Zohar, A.H., Goldman E., Calamary, R., Mashiah, M. 2005. Religiosity and obsessive–compulsive behavior in Israeli jews. Behav. Res. Ther. 43, 857-868. https://doi.org/10.1016/j.brat.2004.06.009

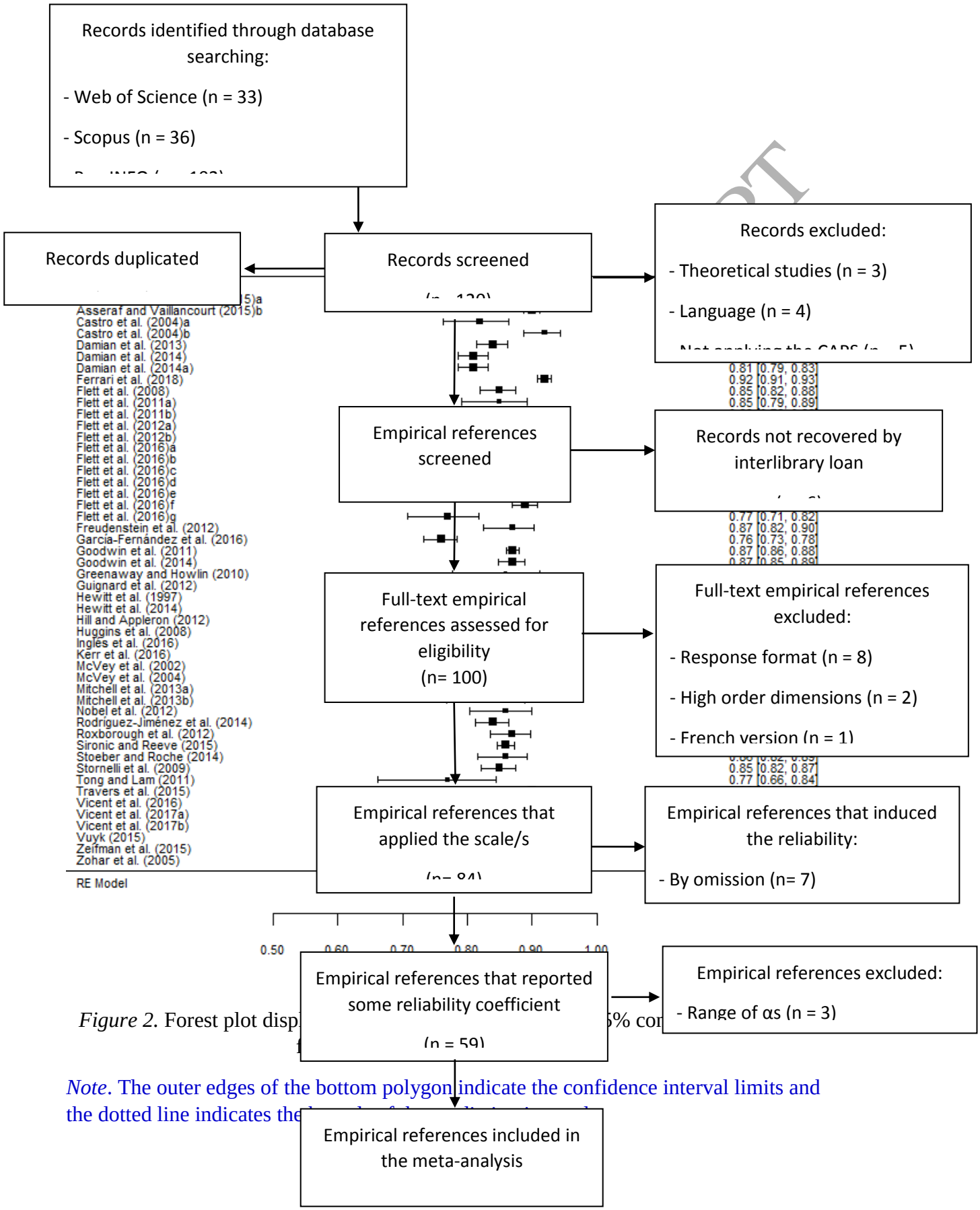


Figure 2. Forest plot displ… 5% con…

Note. The outer edges of the bottom polygon indicate the confidence interval limits and the dotted line indicates the…

Figure 1. REGEMA flow diagram of study selecting process.

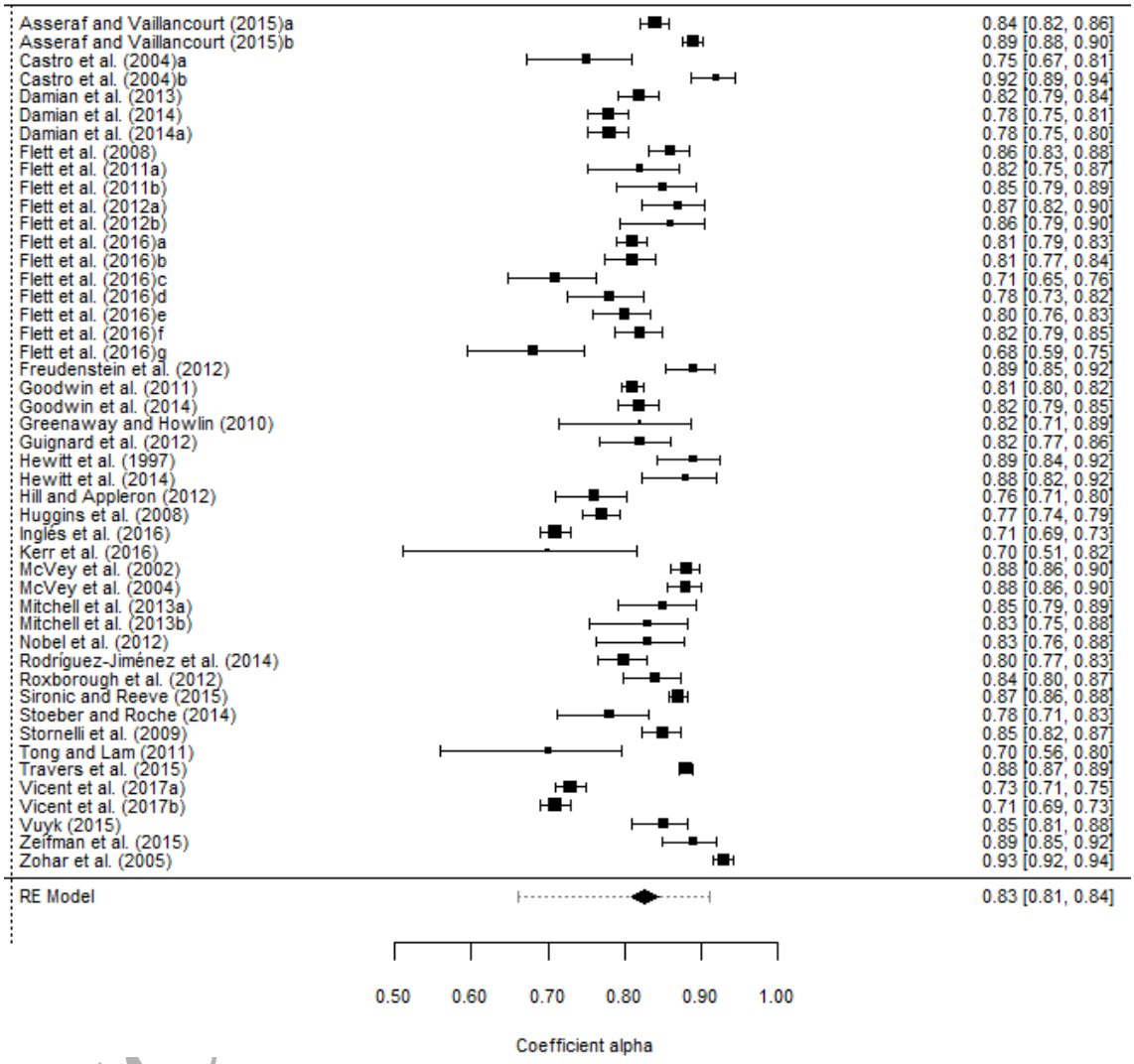


Figure 3. Forest plot displaying the alpha coefficients (and 95% confidence intervals) for the SOP original CAPS version.

Note. The outer edges of the bottom polygon indicate the confidence interval limits and the dotted line indicates the bounds of the prediction interval.

ACCEPTED MANUSCRIPT

Table 1

Average alpha coefficients, 95% confidence and prediction intervals, and heterogeneity statistics for the original CAPS and O'Connor version

			95% CI	95% PI				
	k	α_+	LL ; UL	LL ; LU	*Q*	I^2		$\hat{\tau}^2_{REML}$
Original CAPS:								
Total	11	.87	.84 ; .90	.73 ; .94	174.970****	96.8		.108
SPP	51	.84	.82 ; .85	.72 ; .91	851.738****	93.4		.079
SOP	47	.83	.81 ; .84	.66 ; .91	1010.134****	95.0		.109
O'Connor vers.:								
SPP	7	.82	.76 ; .86	.62 ; .92	34.585****	91.1		.080
SOP_C	6	.74	.65 ; .80	.52 ; .86	14.554**	78.7		.045
SOP_S	6	.73	.67 ; .77	.59 ; .82	12.078*	57.5		.018

SPP = Socially Prescribed Perfectionism. SOP = Self-Oriented Perfectionism. SOP_C = Self-Oriented Perfectionism-Critical. SOP_S = Self-Oriented Perfectionism–Striving. *k* = number of studies. α_+ = mean coefficient alpha. CI = confidence interval. PI = prediction interval. LL and UL= lower and upper limits of the 95% confidence and prediction intervals for α_+. *Q* = Cochran's heterogeneity *Q* statistic; *Q* statistic has *k* – 1 degrees of freedom. I^2 = heterogeneity index. $\hat{\tau}^2_{REML}$ = between-studies variance estimated using restricted maximum likelihood. *p < .05. **p < .01. ****p < .0001.

Table 2

Results of the simple meta-regressions applied on alpha coefficients for the SPP original CAPS version, taking continuous moderator variables as predictors

Predictor variable	*k*	b_j	*F*	*p*	Q_E	R^2
Mean SPP score	40	-0.001	0.03	.971	403.99****	0.0
SD of SPP score	37	0.029	3.61	.066	323.96****	.08

Mean age (years)	43	0.019	1.37	.249	700.26****	0.0
Gender (% male)	46	-0.002	1.28	.264	689.69****	0.0
Ethnicity (% Caucasian)	43	0.005	10.68	.002	666.18****	.20
Year of the study	51	-0.012	1.40	.243	829.20****	0.0

k = number of studies. b_j = regression coefficient of each predictor. *F* = Knapp-Hartung's statistic for testing the significance of the predictor (the degrees of freedom for this statistic are 1 for the numerator and *k* – 2 for the denominator). *p* = probability level for the *F* statistic. Q_E = statistic for testing the model misspecification. R^2 = proportion of variance accounted for by the predictor. ****p* < .0001.

Table 3

Results of the weighted ANOVAs applied on alpha coefficients for the SPP original CAPS version, taking qualitative moderator variables as independent variables

Variable	*k*	α_+	95% CI		ANOVA results
			LL	LU	
Language:					
English	30	.86	.84	.87	$F(6,44) = 5.20$, $p < .001$
Hebrew	3	.87	.83	.90	$R^2 = .39$
Spanish	9	.82	.79	.85	$Q_W(44) = 408.31$, $p < .0001$
French	1	.84	.73	.90	
Romanian	3	.82	.77	.86	
Chinese	4	.72	.64	.78	
Russian	1	.77	.62	.86	

ACCEPTED MANUSCRIPT

Language (dich.):					$F(1,49) = 9.58$, $p = .003$
English	30	.85	.84	.87	$R^2 = .20$
Other	21	.81	.79	.84	$Q_W(49) = 509.29$, $p < .0001$
Study focus:					$F(1,49) = 2.52$, $p = .119$
Applied	9	.81	.77	.85	$R^2 = .03$
Psychometric	42	.84	.83	.86	$Q_W(49) = 842.01$, $p < .0001$
Continent:					
Europe	17	.83	.81	.85	$F(4,46) = 2.32$, $p = .071$
N. America	21	.85	.83	.87	$R^2 = .13$
Asia	7	.80	.75	.84	$Q_W(46) = 590.70$, $p < .0001$
Oceania	4	.87	.82	.90	
S. America	2	.80	.70	.86	
Target population:					
Community	36	.83	.82	.85	$F(3,47) = 1.47$, $p = .239$
Clinical	10	.87	.84	.89	$R^2 = .02$
Comm.+Clinical	4	.83	.78	.88	$Q_W(47) = 828.52$, $p < .0001$
Athletes	1	.80	.64	.89	
Type of disorder:					$F(3,6) = 3.71$, $p = .081$
Anxiety/depression	4	.86	.83	.89	$R^2 = .99$
Eating disorder	2	.90	.86	.93	$Q_W(6) = 5.32$, $p = .514$

ACCEPTED MANUSCRIPT

Mixed psychiatric sample	3	.87	.85	.89	
Other	1	.79	.67	.87	
Financial source:					$F(1,49) = 1.54$, $p = .220$
Public funding	34	.83	.81	.85	$R^2 = .02$
No funding	17	.85	.83	.87	$Q_W(49) = 789.00$, $p < .0001$
Conflict of interest:					$F(1,49) = 0.16$, $p = .689$
No reported	44	.84	.82	.85	$R^2 = 0.0$
No conflict	7	.85	.81	.88	$Q_W(49) = 801.20$, $p < .0001$

k = number of studies. α_+ = mean coefficient alpha. LL and LU = lower and upper 95% confidence limits for α_+. *F* = Knapp-Hartung's statistic for testing the significance of the moderator variable. Q_W = statistic for testing the model misspecification. R^2 = proportion of variance accounted for by the moderator.

Table 4

Results of the simple meta-regressions applied on alpha coefficients for the SOP original CAPS version, taking continuous moderator variables as predictors

Predictor variable	*k*	b_j	*F*	*p*	Q_E	R^2
Mean SOP score	38	0.007	2.46	.125	362.45****	.04
SD of SOP score	35	0.052	13.27	.001	283.17****	.29
Mean age (years)	39	0.045	5.10	.029	627.83****	.11
Gender (% male)	42	-0.004	1.75	.193	820.08****	.03

Ethnicity (% Caucasian)	39	0.003	1.71	.199	802.20****	.02
Year of the study	47	-0.042	16.97	<.001	708.37****	.30

k = number of studies. b_j = regression coefficient of each predictor. *F* = Knapp-Hartung's statistic for testing the significance of the predictor (the degrees of freedom for this statistic are 1 for the numerator and *k* – 2 for the denominator). *p* = probability level for the *F* statistic. Q_E = statistic for testing the model misspecification. R^2 = proportion of variance accounted for by the predictor. ****$p < .0001$.

Table 5

Results of the weighted ANOVAs applied on alpha coefficients for the SOP original CAPS version, taking qualitative moderator variables as independent variables

Variable	*k*	α_+	95% CI LL	95% CI LU	ANOVA results
Language:					
English	28	.84	.82	.86	$F(6,40) = 3.94$, $p = .003$
Hebrew	3	.89	.84	.92	$R^2 = .33$
Spanish	7	.79	.74	.83	$Q_W(40) = 416.26$, $p < .0001$
French	1	.82	.67	.90	
Romanian	3	.79	.71	.85	
Chinese	4	.75	.67	.82	
Russian	1	.68	.42	.82	
Language (dich.):					$F(1,45) = 5.56$, $p = .023$
English	28	.84	.82	.86	$R^2 = .12$
Other	19	.80	.77	.83	$Q_W(45) = 634.80$, $p < .0001$

ACCEPTED MANUSCRIPT

Study focus:					$F(1,45) = 4.84$, $p = .033$
Applied	9	.78	.73	.83	$R^2 = .08$
Psychometric	38	.84	.82	.85	$Q_W(45) = 1004.07$, $p < .0001$
Continent:					
Europe	15	.79	.75	.82	$F(4,42) = 3.09$, $p = .026$
N. America	20	.85	.83	.87	$R^2 = .19$
Asia	7	.83	.78	.86	$Q_W(42) = 533.98$, $p < .0001$
Oceania	3	.85	.78	.90	
S. America	2	.79	.68	.87	
Target population:					
Community	32	.82	.80	.84	$F(3,47) = 1.46$, $p = .239$
Clinical	10	.85	.82	.89	$R^2 = .02$
Comm.+Clinical	4	.81	.73	.87	$Q_W(47) = 828.52$, $p < .0001$
Athletes	1	.76	.52	.88	
Type of disorder:					$F(3,6) = 0.07$, $p = .976$
Anxiety/depression	4	.85	.74	.91	$R^2 = 0.0$
Eating disorder	2	.85	.68	.93	$Q_W(6) = 34.66$, $p < .0001$
Mixed psychiatric sample	3	.87	.76	.92	
Other	1	.86	.59	.95	
Financial sources:					$F(1,45) = 0.04$, $p = .842$

ACCEPTED MANUSCRIPT

Public funding	31	.83	.80	.85	R^2 = 0.0
No funding	16	.83	.80	.86	$Q_W(45)$ = 965.99, p <.0001
Conflict of interests:					
No reported	40	.83	.81	.85	$F(1,45)$ = 0.77, p = .384
No conflict	7	.81	.75	.85	R^2 = 0.0
					$Q_W(45)$ = 878.47, p <.0001

k = number of studies. α_+ = mean coefficient alpha. LL and LU = lower and upper 95% confidence limits for α_+. *F* = Knapp-Hartung's statistic for testing the significance of the moderator variable. Q_W = statistic for testing the model misspecification. R^2 = proportion of variance accounted for by the moderator.

Table 6

Results of the multiple meta-regression applied on alpha coefficients for the SPP original CAPS version, taking as predictors the percentage of Caucasian and the language dichotomized (k = 43)

Predictor variable	b_j	*t*	*p*	Model fit
Intercept	1.330	9.56	< .0001	$F(2, 40)$ = 7.69, p = .002
Ethnicity (% Caucasian)	0.005	3.02	.004	R^2 = .30
Language (dich.)	0.165	1.99	.053	$Q_E(40)$ = 431.62, p < .0001

Model	*F*	*p*	R^2	ΔR^2
Full model	7.69	.002	.30	-
Ethnicity (% Caucasian)	10.68	.002	.20	.11
Language (dich.)	9.58	.003	.19	.10

ACCEPTED MANUSCRIPT

b_j = regression coefficient of each predictor. *t* = statistic for testing the significance of the predictor (with 40 degrees of freedom). *p* = probability level for the *t* statistic. *F* = Knapp-Hartung's statistic for testing the significance of the full model. Q_E = statistic for testing the model misspecification. R^2 = proportion of variance accounted for by the predictors. ΔR^2 = increase in R^2 as consequence of including in the model a predictor once the other predictors had already been introduced.

Table 7

Results of the multiple meta-regression applied on alpha coefficients for the SOP original CAPS version, taking as predictors the SD of PAO scores, the mean age, the year of the study, the language dichotomized and the study focus (k = 28)

Predictor variable	b_j	*t*	*p*	Model fit
Intercept	31.017	1.07	.295	$F(5, 22) = 10.94$, $p < .0001$
SD of PAO score	0.063	5.02	<.0001	$R^2 = .78$
Mean age (years)	0.036	2.43	.024	$Q_E(22) = 61.75$, $p < .0001$
Year of the study	-0.015	-1.07	.296	
Language (dich.)	0.083	1.02	.319	
Study focus	0.347	3.67	.001	

Model	*F*	*p*	R^2	ΔR^2
Full model	10.94	<.0001	.78	-
SD of PAO score	13.27	.001	.29	.27
Mean age (years)	5.10	.029	.11	.22
Year of the study	16.97	.000	.30	0
Language (dich.)	5.56	.023	.12	0

ACCEPTED MANUSCRIPT

Study focus	4.84	.033	.08	.24

b_j = regression coefficient of each predictor. t = statistic for testing the significance of the predictor (with 22 degrees of freedom). p = probability level for the t statistic. F = Knapp-Hartung's statistic for testing the significance of the full model. Q_E = statistic for testing the model misspecification. R^2 = proportion of variance accounted for by the predictors. ΔR^2 = increase in R^2 as consequence of including in the model a predictor once the other predictors had already been introduced.

ACCEPTED MANUSCRIPT