

1 Hill, A. P., Appleton, P. R., & Mallinson, S. H. (in press). Development and Initial
2 Validation of the Performance Perfectionism Scale for Sport (PPS-S). *Journal of*
3 *Psychoeducational Assessment*. Accepted for publication: 29/04/2015.

4
5 To appear in a special issue of the *Journal of Psychoeducational Assessment* titled
6 "Advances in the Assessment of Perfectionism" appearing in print in October 2016.

7
8
9
10
11
12
13 Development and Initial Validation of the Performance Perfectionism Scale for Sport (PPS-
14 S)

15 Andrew P. Hill¹, Paul R. Appleton², & Sarah H. Mallinson¹

16 ¹York St John University, UK

17 ²University of Birmingham, UK

18
19
20
21
22
23 Author Note

24 Andrew P. Hill, Faculty of Health and Life Sciences, York St John University, UK; Paul R
25 Appleton, School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham;
26 Sarah H. Mallinson, Faculty of Health and Life Sciences, York St John University, UK

27
28 Correspondence concerning this article should be addressed to Andrew Hill, Faculty of
29 Health and Life Sciences, York St John University, York, YO31 7EX, United Kingdom. E-
30 mail: a.hill@yorks.ac.uk

Abstract

Valid and reliable instruments are required in order to appropriately study perfectionism. With this in mind, three studies are presented that describe the development and initial validation of a new instrument designed to measure multidimensional performance perfectionism for use in sport (Performance Perfectionism Scale-Sport, PPS-S). The instrument is based on Hewitt and Flett's (1991) model of perfectionism and includes self-oriented, socially prescribed, and other-oriented performance perfectionism. These dimensions encapsulate the features of Hewitt and Flett's dimensions but are focused on athletic performance, rather than life generally. The three studies outline item generation and refinement, exploratory, confirmatory, and exploratory-confirmatory examination of factor structure, and initial assessment of construct validity in multiple samples of adolescent and young adult athletes. Findings suggest that the PPS-S is likely to be reliable and valid measure of performance perfectionism in youth sport. As validation continues, we expect the instrument to have wider applicability for use in adults and other performance contexts (e.g., education and work).

Key Words: Perfectionism, Questionnaire, Survey, Psychometrics

1 Research examining perfectionism in sport extends across 25 years and includes over
2 150 studies (see Hill, 2016). This research has revealed perfectionism to be a complex,
3 multidimensional, personality characteristic with important implications for athletes. On one
4 hand, some dimensions of perfectionism (typically labelled perfectionistic strivings) are
5 associated with desirable correlates, processes, and consequences (e.g., self-confidence,
6 problem-focussed coping, and performance). On the other hand, other dimensions of
7 perfectionism (typically labelled perfectionistic concerns) are associated with undesirable
8 desirable correlates, processes, and consequences (e.g., anxiety, avoidant coping, and
9 burnout). As evidenced by this research, perfectionism has much to say regarding the
10 experiences of athletes.

11 A number of instruments have been used to assess perfectionism in sport (e.g., Frost,
12 Marten, Lahart, & Rosenblate, 1990; Gotwals & Dunn, 2009; Stoeber, Otto, & Stoll, 2006).
13 One popular instrument/model is that developed by Hewitt and Flett (Multidimensional
14 Perfectionism Scale, HF-MPS, 1991, 2004). Hewitt and Flett define perfectionism as a
15 marked need for absolute perfection from self and others. According to their model, trait
16 perfectionism has self-oriented, socially prescribed, and other-oriented dimensions. Self-
17 oriented perfectionism (SOP) is the tendency to set excessively high personal standards, to
18 focus on flaws in personal performance and to respond to substandard performance with
19 harsh self-criticism. Socially prescribed perfectionism (SPP), by contrast, is the belief that
20 significant others impose unrealistic standards on the self and that approval is contingent on
21 their achievement. Finally, other-oriented perfectionism (OOP) is the tendency to impose
22 perfectionistic standards on others.

23 Hewitt and Flett's (1991) model has a number of notable strengths. In particular, the
24 model is grounded in the work of classic clinicians and theorists and arguably offers the most
25 complete theoretical model of perfectionism currently available. Unlike other models, for

1 example, it includes an explanation of the developmental origins of perfectionism, identifies
2 moderating and mediating factors, and outlines the tenets of effective treatment/management
3 of perfectionism. Importantly for us here, research has also found strong support for the
4 predictive ability of this model in a wide range of domains including sport (see Jowett,
5 Mallinson, & Hill, 2016, for a recent review). Research in sport suggests that SOP includes
6 both desirable and undesirable features while SPP is uniformly problematic. Less is known
7 about OOP in sport as studies have typically excluded this dimension in favor of examining
8 the personal (as opposed to interpersonal) influence of perfectionism but recent research
9 suggests it is also likely to be important, particularly in terms of team performance (e.g., Hill,
10 Stoeber, Brown, & Appleton, 2014).

11 When using the instrument developed by Hewitt and Flett (1991) in sport, researchers
12 have typically adapted it in various ways. Most commonly, the instructions given to
13 respondents have been changed to focus their attention on sport when completing the items
14 (e.g., "...in relation to your sport participation...") and/or items have been amended so to
15 focus on sport (e.g., changing "my life" to "my sport"). Adapting instruments in this manner
16 is a common strategy in research and can ensure close correspondence between concepts
17 when measured in different domains. However, even after amending items it is unclear
18 whether all items are best suited, applicable, or readily interpretable in context of sport or
19 whether the instrument captures perfectionism fully in sport. This is because the instrument
20 was not developed with sport, or specific aspects of sport, in mind (see Stoeber & Madigan,
21 2016, for a further discussion of these and other issues pertaining to the measurement of
22 perfectionism in sport).

23 Researchers have sought to address such drawbacks by developing domain-specific
24 measures of perfectionism in sport (e.g., Sport-Multidimensional Perfectionism Scale-2, S-
25 MPS-2, Gotwals & Dunn, 2009). There is strong support for the assessment of personality

1 characteristics when anchored in a specific context or frame-of-reference (e.g., Bing,
2 Whanger, Davison & VanHook, 2004; Hunthausen, Truxillo, Bauer, & Hammer, 2003;
3 Lievens, De Corte & Schollaert, 2008). In addition, in regards to perfectionism in particular,
4 it is also common for individuals to report being more or less perfectionistic depending on the
5 domain. This has been illustrated across multiple life domains (e.g., Stoeber & Stoeber, 2009)
6 and has been illustrated in relation to sport specifically. For example, in comparing the scores
7 of successful intercollegiate athletes in terms of perfectionism in sport, school, and in life in
8 general, Dunn, Gotwals, and Causgrove Dunn (2005) found that the athletes typically
9 reported significantly higher perfectionism in sport than in other domains. One consequence
10 is that domain-specific measurement of perfectionism has been found to have greater
11 predictive ability when compared to general measures of perfectionism in sport (e.g., Dunn,
12 Craft, Causgrove Dunn, & Gotwals, 2011). Therefore, there is a strong case for the
13 availability of instruments that measure domain-specific perfectionism.

14 Against this backdrop, in the current study we sought to develop an instrument to
15 measure the dimensions of perfectionism in Hewitt and Flett's (1991) model as they apply to
16 a specific aspect of sport, namely performance. Performance is one of the defining features of
17 the sport domain and is perhaps the single most important aspect of an athlete's life. In
18 focusing on performance, we provide a domain-specific measure of self-oriented
19 performance perfectionism (SOPP), socially prescribed performance perfectionism (SPPP),
20 and other-oriented performance perfectionism (OOPP). We conceive these dimensions of
21 performance perfectionism to be subordinate to Hewitt and Flett's (1991) trait dimensions
22 and to operate at a more specific, contextual level (i.e., "I expect my performances to be
23 perfect") than the original three traits that one would expect to be evident at multiple levels
24 including a general level (e.g., "I expect to be perfect in everything I do") and a dispositional
25 level (e.g., "I expect to be perfect in sport"). In this sense, dimensions of performance

1 perfectionism are similar to dimensions of perfectionism that manifest in other specific
2 contexts such as in practice and in competition (e.g., Stoeber et al., 2006).

3 **Present Research**

4 In summary, the purpose of this research was to develop and begin to validate a
5 domain-specific measure of multidimensional performance perfectionism for use in sport
6 (Performance Perfectionism Scale-Sport, PPS-S). To this end, we provide three studies. The
7 first study describes the process through which items were generated and refined to capture
8 the three performance perfectionism dimensions. The second study provides an exploratory
9 examination of factor structure of the items. The third study provides a further examination of
10 the factor structure of the instrument using confirmatory and exploratory-confirmatory
11 analyses, as well as an initial test of the construct validity of the PPS-S. As much of the
12 research in this area (and much of our own research) has examined perfectionism among
13 youth athletes, we choose to begin the validation of the PPS-S in adolescent and young adult
14 athletes.

15 **Study 1**

16 The purpose of study one was to develop items that assessed the three dimensions of
17 performance perfectionism and were applicable to sport. In addition, items were also assessed
18 in terms of whether they were understandable to adolescent and young adult athletes.

19 *Initial item generation and item refinement*

20 Definitions of SOP, SPP, and OOP provided by Hewitt and Flett (1991, 2004) were
21 adapted to incorporate a focus on perfect athletic performance (“...the demand of *perfect*
22 *athletic performance* from oneself, the tendency to evaluate one’s performance stringently
23 and engage in harsh self-criticism,” “...the perception that others are demanding *perfect*
24 *athletic performance* from the self and that others evaluate one’s attempts to meet these
25 prescribed standards stringently and critically,” and “...the demand of *perfect athletic*

1 *performance* from others and tendency to evaluate other people's performances stringently
2 and criticise others."). The authors then used these definitions along with a list of core
3 characteristics to independently generate items which were thought to capture these
4 dimensions. Following the recommendations of DeVellis (1991), items were generated with
5 the aim of representing all of the core features of each dimension and developing
6 unidimensional subscales. The items were also developed so that they were consistent with
7 the original response format of the HF-MPS (7-point agreement Likert scale) and were
8 appropriate in terms of readability for adolescents and young adults.

9 A number of conceptual issues were also taken into account when constructing items.
10 Firstly, care was taken to refer to flawlessness and perfection, rather than high or
11 exceptionally high standards. This was because there is currently debate regarding the
12 difference between the pursuit of high standards and perfectionistic standards (see Flett &
13 Hewitt, 2006). Secondly, based on the recommendation of Flett and Hewitt (2002), no items
14 made reference to the degree to which standards were attained or unattained and items did not
15 refer to emotional reactions to the failure to meet important standards. In this regards, the
16 intention was to create items that capture perfectionism independent of ability and its
17 consequences. Finally, when constructing items for SPPP and OOPP, no specific other was
18 identified (e.g., coaches, parents, and teammates etc.). Instead, instructions were created so to
19 direct respondents to think of individuals whose "opinions they valued." This decision was
20 made so to balance the desire to capture the concepts as described by Hewitt and Flett (1991,
21 2004) with the need to provide guidance to participants ("Below are statements that reflect
22 beliefs that athletes hold when taking part in sport. Some of the beliefs refer to other people.
23 For these, think about the people involved in your sport participation whose opinion you
24 value. Please read each statement, and then select a number from 1 to 7 to show how much
25 you agree or disagree. There are no right or wrong answers.").

1 This process yielded an initial pool of 196 items. These items were then assessed by
2 the authors for their clarity, readability (assessed using Flesch-Kincaid grade level score;
3 Kincaid, Fishburne, Rogers, & Chissom, 1975), relevance, similarity to other items, and the
4 degree to which they adhered to the criteria outlined above. This review led to a revised pool
5 of 90 items.

6 *External review of items and item refinement*

7 The 90 items were subject to a review conducted by an external panel of five
8 academics with experience of conducting research in the area perfectionism. Each member of
9 this panel had published research in international peer-reviewed journals in this area (2008-
10 onwards). The panel was presented with a definition of each dimension of perfectionism as
11 they manifest in sport, a list of their core features and the proposed items. The expert panel
12 was asked to identify the dimension of perfectionism that each item corresponded with, the
13 content suitability of each item (high, moderate, and low) and the clarity of each item (high,
14 moderate, and low). The external panel were also invited to provide alternative wording and
15 additional items. Based on the feedback from this panel, a second revised pool of 57 items
16 was developed (22 SPPP, 20 SOPP, and 15 OOPP).

17 The revised pool of items was then subject to a second external review by a panel of
18 13 sport coaches (9 males, 4 females, M age = 38.42, s = 8.77 yrs, range 27 to 52 yrs). These
19 coaches were recruited from sport organisations and represented a wide range of sports
20 (football = 3, rugby union = 1, rugby league = 2, netball = 1, cricket = 2, swimming = 1,
21 tennis = 2 and basketball = 1). They had considerable coaching experience (M = 14.31, s =
22 7.04, range 4 to 25 yrs) and coached at a range of levels (recreational = 3, regional = 2,
23 national = 2, international = 3, semi-professional = 1, professional = 2). These coaches were
24 asked to indicate whether they considered the content of each item to be applicable to the
25 sport they coach (applicable versus not applicable) and whether the item was clear (high,

1 moderate, and low clarity). Based on the feedback from this panel, the pool of items was
2 revised further and a set of 61 items was developed (24 SPPP, 20SOPP, and 17 OOPP).

3 The final phase of item refinement was the completion of three focus groups. The aim
4 of the focus groups was to assess the readability, comprehension, and clarity of the items. The
5 focus groups also provided a means of assessing if respondents understood questions in the
6 same manner and whether respondents were willing and able to answer the questions
7 (Collins, 2003). The participants were all adolescent athletes (5 males, 11 females, *M* age =
8 14.15, *s* = 1.31 yrs, range 12 to 16 yrs) from a range of sports (netball = 7, swimming = 2,
9 football = 3, rugby union = 2, gymnastics = 1) and varying levels (club = 8, county = 6,
10 regional = 2). The focus groups were conducted following the recommendations of Morgan
11 (1992). All focus groups included same-sex participants, 5 or 6 members, and lasted between
12 60 and 90 minutes. In each session participants were presented with a written set of the items.
13 A “think-aloud” method was followed whereby athletes were asked to comment on what they
14 believed to be the meaning of each item (Ericsson & Simon, 1998). This procedure was
15 supplemented by the use of predetermined probes that were aimed at exploring
16 comprehension (e.g., “What did you understand by this word/question?” “What/who are you
17 thinking about when answering this question?” and “How would you explain this question to
18 someone else?”; Collins, 2003). Following the focus groups, a final revised pool of 75 items
19 was developed (29 SPPP, 25 SOPP, and 21 OOPP).

20

Study 2

21 The purpose of Study 2 was to reduce the number of items and explore items and
22 factor structure in relation to the original HF-MPS three-factor model. This included
23 examination of the initial pool of items generated in Study 1 in a first sample and a
24 subsequent examination of factor structure in a second sample. The final set of items was also
25 assessed in terms of internal reliability and readability.

Methods

Participants

Sample one. Three-hundred and twenty-one sports participants completed the pool of items (196 males, 125 females; M age = 14.30 yrs, s = 1.50, range 11 to 18). Participants were recruited from a range of individual and teams sports (e.g., swimming, football, and rugby) and included representatives of a range of competitive levels (recreational/fun = 27, club = 101, county/district = 72, region = 99, country = 20, unspecified = 2). On average, athletes trained and competed 4.17 hrs per week (s = 3.02) and considered participation in their sport very important in comparison to other things in their life (M = 7.85, s = 1.18, range 1 to 9).

Sample two. Two hundred and twenty-nine sports participants completed items derived from the analyses of sample one (102 males, 125 females, 2 non-respondents; M age = 14.96, s = 1.58, range 12 to 18). Again, participants were recruited from a range of individual and team sports and included representatives of a range of competitive levels (recreational/fun = 29, club = 49, county/district = 26, region = 38, country = 40, unspecified = 47). On average, athletes trained and competed 6.21 hrs per week (s = 3.44) and considered participation in their sport very important in comparison to other things in their life (M = 7.63, s = 1.58, range 1 to 9).

Data analysis

Items were assessed in terms of content along with general characteristics (means, variances, and distribution). Following the removal of items based on this assessment, exploratory factor analysis (EFA) was conducted in accordance with common recommendations (e.g., Child, 2006; Tabachnick & Fidell, 2001; Worthington & Whittaker, 2006). Factor solutions/retention was explored using principal components analysis (PCA) and assessed using three common strategies: eigenvalues, screeplot, and parallel analysis

1 (using O'Connor, 2000, with PCA and assessment of 95% percentiles). This was followed by
2 common factor analysis using principal axis factoring extraction (PAF) with oblique rotation
3 ($\delta = 0$) in which items were constrained to load on the number of retained factors. Factor
4 solutions were then assessed based upon interpretability, structural/pattern coefficients ($> .30$
5 was considered meaningful), degree of cross-loading (i.e., the presence of loadings above $.30$
6 on more than one factor), and communalities ($> .20$ was considered meaningful). Internal
7 reliability was assessed using Cronbach's α , inter-item correlations and corrected item-total
8 correlations (Cronbach's $\alpha > .70$, inter-item correlations between $.20$ and $.70$, and item-total
9 correlations $> .30$ were used to guide assessment; Kidder & Judd, 1986). Readability was
10 assessed using Flesch-Kincaid grade level score (Kincaid et al., 1975).

11 **Results**

12 *Exploratory factor analysis and internal reliability*

13 The analyses described above revealed that the most robust and interpretable solution
14 in sample one consisted of 12-items loading on three factors. In arriving at this solution, it is
15 noteworthy that the final PCA on all items provided two eigenvalues (rather than three) that
16 exceeded one and the scree plot and parallel analysis supported the retention of only two
17 factors (actual $\lambda_1 = 5.30$, $\lambda_2 = 1.79$, $\lambda_3 = 0.94$ versus $\lambda_1 = 1.40$, $\lambda_2 = 1.30$, $\lambda_3 = 1.22$ from
18 parallel analysis). However, a three factor solution was retained for a number of reasons.
19 Firstly, in addition to data-derived strategies, factor analysts recommend that relevant theory
20 should also guide decisions regarding the number of factors to retain (Fabrigar, Wegener,
21 MacCallum, & Strahan, 1999). Secondly, the three factor solution provided pattern
22 coefficients that were more interpretable (i.e., all items loaded on the intended subscales).
23 Finally, unlike the two factor solution, the three factor solution displayed simple structure
24 (i.e., there were no cross-loadings that exceeded $.30$).

1 To verify the 12-item three factor solution, the EFA procedure described earlier was
2 also conducted using sample two. On this occasion eigenvalues and parallel analysis
3 supported the three-factor solution (actual $\lambda_1 = 3.78$, $\lambda_2 = 2.02$, $\lambda_3 = 1.32$ versus $\lambda_1 = 1.49$, λ_2
4 $= 1.36$, $\lambda_3 = 1.26$ from parallel analysis). Based on this replication, we concluded that the 12-
5 item three factor solution offered the most robust item/factor structure on which validation of
6 the instrument should proceed. The PAFs for both sample one and two are displayed in Table
7 1 and provide strong support for the 12-item three-factor solution with all items loading
8 meaningfully on factors reflective of the HF-MPS (i.e., the three factors are discernible in
9 terms of being self-oriented, socially prescribed, and other-oriented), minimal cross-loading
10 (only two instances), and all communalities exceeding the minimum threshold.

11 In terms of internal reliability, all subscales displayed acceptable Cronbach's α : SOPP
12 $\alpha = .83/.70$, SPPP $\alpha = .75/.73$ and OOPP $\alpha = .87/.79$ (sample one left and sample two right).
13 In addition, all inter-item correlations were within recommended limits and all corrected
14 item-total correlations were acceptable (i.e., exceeded .30).

15 *Assessment of readability*

16 Flesch-Kincaid grade level scores for the items ranged from 4.7 (4th grade, typically
17 suitable for 9 to 10 year olds) to 10.7 (10th grade, typically suitable for 15 to 16 year olds).
18 Nine of the 12 items scores were within 6th grade to 8th grade reading ability range (i.e.,
19 typically suitable for 11 to 14 year olds). One other item was associated with 4th grade and
20 two items scored higher than 8th grade, both of which were associated with a 10th grade
21 reading ability (SOPP10 and SPPP9). Overall, based on these scores we concluded that the
22 instrument is likely to be appropriate for use among adolescents and young adults (with the
23 caveat that the two items identified above may need further revision to improve readability
24 for younger participants).

25

Study 3

1 The first purpose of Study 3 was to further examine the factor structure of the new
2 instrument using both confirmatory and exploratory-confirmatory analyses. Typically
3 confirmatory factor analysis (CFA) is adopted at this phase of the validation process. CFA is
4 a popular analysis because it allows researchers to test a specified factor structure between
5 indicators (e.g., items) and latent factors (e.g., dimensions of perfectionism), it provides
6 standard errors for parameter estimates, and allows for a vigorous test of factor structure in
7 terms of fit with observed data. As such, it is a valuable analysis when validating
8 psychometric instruments. However, despite its utility, a number of criticisms of CFA have
9 recently emerged. In particular, in CFA each item is permitted to load on only one factor with
10 zero cross-loadings on all others (i.e., perfect simple structure). This specification is
11 considered to be too restrictive and unrealistic for many multidimensional models with more
12 complex structures (i.e., at least one item cross-loads on more than one factor) (Marsh et al.,
13 2009). As a result, this (mis)specification is associated with a number of undesirable
14 consequences including failure to replicate structures using CFA even when based on
15 multiple EFA (Marsh et al., 2009), inflated factor correlations (Marsh, Nagengast, & Morin,
16 2013), and biased estimates in the non-measurement part of a structural equation model
17 (SEM) (Asparouhov & Muthén, 2009).

18 To overcome these limitations, exploratory structural equation modelling (ESEM) can
19 be used. ESEM combines the strengths of CFA and EFA within a SEM framework
20 (Asparouhov & Muthén, 2009). Consistent with EFA, ESEM allows for a complex structure
21 where all indicators are permitted to load on all factors and, consistent with CFA, ESEM
22 provides robust means of evaluating model adequacy (e.g., standard errors for parameter
23 estimates and goodness-of-fit indexes). In summarising the relative strengths (and
24 weaknesses) of CFA and ESEM, Myers, Chase, Pierce, and Martin (2011) suggested that
25 CFA is the preferred technique when a prior measurement theory exists and ESEM is the

1 preferred technique when a prior measurement theory does not exist. In initial validation
2 studies, when it is difficult to conclude that adequate a prior measurement theory exists,
3 Myers et al. (2011) argued that it is advantageous to use both CFA and ESEM. We therefore
4 did so here using three independent samples.

5 The second purpose of Study 3 was to examine the construct validity of the new
6 instrument (i.e., “the degree to which a test measures what it claims, or purports, to be
7 measuring”, Brown, 1996, pp. 231). This is tested here by examining correlations between
8 dimensions of the PPS-S and an established domain-specific measure of perfectionism (i.e.,
9 criterion-related or concurrent validity). The instrument used was the S-MPS-2 (Gotwals &
10 Dunn, 2009). In terms of instruments available to researchers in sport, there is strong
11 evidence to support the S-MPS-2 in terms of its reliability and validity among athletes (see
12 Dunn et al., 2002; Dunn et al., 2006; Gotwals & Dunn, 2009). It is also the most widely used
13 domain-specific measure of multidimensional perfectionism in sport (Stoeber & Madigan,
14 2016). Indeed, when recently reviewing instruments available to researchers in sport, Stoeber
15 and Madigan concluded that the S-MPS-2 is an excellent domain-specific measure of
16 perfectionism and recommended its use when examining perfectionism in athletes.

17 Support for the construct validity of the PPS-S is provided if its subscales
18 demonstrated meaningful relationships with subscales of the S-MPS-2 in a theoretically
19 expected manner. In this case, in keeping with previous research examining the relationships
20 between different measures of multidimensional perfectionism (e.g., Cox, Enns, & Clara,
21 2002; Dunn et al., 2006; Frost, Heimberg, Holt, Mattia, & Neubauer, 1993), it was
22 hypothesised that (i) SOPP would be positively correlated with all dimensions of the S-MPS-
23 2 but most strongly with personal standards and organisation, (ii) SPPP would be positively
24 correlated with all dimensions of the S-MPS-2 but most strongly with concern over mistakes
25 along with perceived coach and parental pressure, and (iii) OPOP would be positively

1 correlated with all dimensions of the S-MPS-2 but to a lesser degree than the two other
2 dimensions and most strongly with personal standards.

3 **Methods**

4 *Participants*

5 Sample three. Two-hundred and forty-one athletes were recruited to sample three (98
6 males, 143 females; M age = 15.11, s = 2.03, range 11 to 19). Participants were recruited
7 from a range of individual and team sports (e.g., netball, football, and tennis) and included
8 representatives of a range of competitive levels (recreational = 27, club = 107, county/district
9 = 65, region = 28, country = 14). On average, athletes trained and competed 4.12 hrs per
10 week (s = 3.62) and considered participation in their sport very important in comparison to
11 other things in their life (mean = 6.93, s = 1.73, range 1 to 9).¹

12 Sample four. Two-hundred and twenty-two athletes were recruited to sample four (65
13 males, 157 females; M age = 13.51, s = 1.53, range 11 to 18). Participants were recruited
14 from a range of individual and team sports (e.g., netball, football, and hockey) and included
15 representatives of a range of competitive levels (recreational/fun = 38, club = 105,
16 county/district = 62, region = 11, country = 4). On average, athletes trained and competed
17 5.09 hrs per week (s = 5.08) and considered participation in their sport very important in
18 comparison to other things in their life (M = 7.27, s = 1.64, range 1 to 9).

19 Sample five. Two-hundred and fifty-two athletes were recruited to sample five (20
20 males, 232 females; M age = 13.65, s = 1.14, range 11 to 16 yrs). Participants were recruited
21 from a range of individual and teams sports (e.g., netball, football, and hockey) and included
22 representatives of a range of competitive levels (recreational/fun = 37, club = 107,
23 county/district = 81, region = 22, country = 2, unspecified = 3). On average, athletes trained

¹ This sample is the same as reported in Mallinson, Hill, Hall, and Gotwals (2014). However, the PPS-S was not examined in Mallinson et al.'s study.

1 and competed 3.00 hrs per week ($s = 2.14$) and considered participation in their sport very
2 important in comparison to other things in their life ($M = 7.22$, $s = 1.69$, range 1 to 9).

3 *Data Analysis*

4 CFA and ESEM were conducted using *Mplus* 5.0 (Muthén & Muthén, 2007) with
5 robust maximum likelihood estimator (MLR). Oblique target rotation was implemented in the
6 ESEM. The same guidelines as presented in study two were followed in terms of interpreting
7 factor loadings (supplemented by tests of statistical significance provided in both CFA and
8 ESEM). Multiple indexes were used to assess model fit in the confirmatory and exploratory-
9 confirmatory analyses: chi-square statistic (χ^2), comparative fit index (CFI), root mean square
10 error of approximation (RMSEA), 90% confidence intervals of the RMSEA, and the
11 standardized root-mean-square residual (SRMR). Conventional criteria were used when
12 interpreting these indexes with values $>.90$ CFI, $< .08$ RMSEA (90% CI $<.05$ to $<.08$) and
13 $<.08$ SRMR providing evidence of adequate model fit (Marsh, Hau, & Wen, 2004). It should
14 be noted that while the use of these indexes are well established in CFA, there adequacy in
15 ESEM is less clear (Marsh, et al., 2010). Therefore, as advised by Morin and Maïano (2011),
16 the criteria for the indexes identified above were used as part of an overall assessment of the
17 features of the models. Cohen's (1992) guidelines of small (.10), medium (.30), and large
18 (.50) were used when interpreting factor correlations and bivariate correlations.

19 **Results**

20 *Assessment of factorial structure*

21 Fit indexes, factor loadings, uniquenesses, and factor correlations for CFAs and
22 ESEMs are reported in Tables 2, 3, and 4. CFAs revealed that the hypothesized model
23 provided an adequate fit, or approached adequate fit, in samples three and four. However, the
24 hypothesized model provided inadequate fit in sample five. Examination of the standardized
25 parameter estimates from the CFAs indicated that all loadings were significant and large.

1 ESEMs provided clearer support for the model in that fit was typically better when using this
2 analysis. Sample four was however an exception in this regard. Across all the samples,
3 almost all items loaded significantly and meaningfully on the expected factors. The only
4 exceptions were SOPP10 in samples three and five. There were a small number of cross-
5 loadings but these were typically not meaningful (i.e., $<.30$). The notable exceptions were
6 SOPP10 (sample five only) and SPPP7 (sample three and four). In the case of SPPP7, cross-
7 loadings were smaller than loadings on the expected factor. Factor correlations in CFAs and
8 ESEMs were typically medium (SOPP-OOPP) and large (SPPP-SOPP and SPPP-OOPP).
9 Collectively, the results from CFAs and ESEMs provided support for the hypothesised three-
10 factor model of the PPS-S.

11 *Construct validity*

12 Bivariate correlations between the subscales of the PPS-S and the S-MPS-2 are
13 reported in Table 5. Across the three samples, athletes reported moderate levels of SOPP,
14 moderate-to-low levels of SPPP, and low levels of OOPP and moderate levels of
15 perfectionism as captured by the S-MPS-2 (based on Likert scales). Examination of the
16 bivariate correlations between subscales of the PPS-S revealed that SOPP had a significant
17 positive relationship with all subscales of the S-MPS-2. These were typically medium and
18 medium-to-large in size with the largest relationship evident with personal standards. SPPP
19 also had a significant positive relationship with all subscales of the S-MPS-2. These were
20 typically medium-to-large or large in size. Notably, its relationships with perceived coach and
21 parental pressure were among its largest relationships and exceeded those associated with the
22 other dimensions of the PPS-S. Finally, OOPP had a significant positive relationship with all
23 subscales. The relationships were typically medium and medium-to-large in size. The
24 relationships were largely consistent across the three samples.

1 Multiple regressions are reported in Table 6. The S-MPS-2 was a significant predictor
2 of all dimensions of the PSS-S in all three samples. For SOPP, 43%, 39%, and 44% of
3 variance was explained ($p < .001$). For SPPP, 43%, 33%, and 33% of variance was explained
4 ($p < .001$). For OOPP, 21%, 21%, and 24% of variance was explained ($p < .001$). SOPP was
5 significantly predicted by personal standards (all samples) and concern over mistakes
6 (samples three and five), and, to a lesser degree, by perceived coach pressure (sample four)
7 and doubts about action (sample five). SPPP was significantly predicted by perceived coach
8 pressure (all samples), concern over mistakes (samples three and five), and perceived parental
9 pressure (sample three). Finally, OOPP was significantly predicted by concern over mistakes
10 (samples four, five, and, marginally, in sample three, $p = .051$), personal standards (sample
11 five), perceived parental pressure (sample four), and perceived coach pressure (sample five).

12 Discussion

13 The purpose of this research was to develop and begin to validate an instrument
14 designed to measure of multidimensional performance perfectionism for use in sport (PPS-S).
15 Three studies were reported here that described item generation and refinement, exploratory,
16 confirmatory, and exploratory-confirmatory analysis of factor structure, and an initial test of
17 construct validity.

18 Item development and refinement was used to provide items that measured
19 performance perfectionism and were interpretable and meaningful in context of sport. The
20 relevance of the items was confirmed by both coaches and athletes. Readability analyses also
21 indicated that generally the items are likely to be suitable for adolescents and young adults.
22 We therefore consider the PPS-S to offer a good means of assessing performance
23 perfectionism in these groups plus, subject to confirmation by future research, in all
24 likelihood adult athletes. Our analyses suggested that two items (SOPP10 and SPPP9) may be
25 more difficult for younger participants. While these two items were not identified as

1 problematic in the focus groups or when assessing internal reliability, these items may
2 therefore need minor revision to improve readability as validation of the PPS-S continues. In
3 the meantime, we recommend that when distributing the instrument to younger athletes,
4 particular attention is given to these items as part of standard procedures for assessing the
5 properties of psychometric instruments (e.g., assessing internal reliability and factor
6 structure).

7 After initial exploratory work, the factor structure of the new instrument was revealed
8 to be sound and in keeping with Hewitt and Flett's (1991) original model. In terms of
9 possible improvement, there were seven (of 72 possible) instances of cross-loading when
10 using ESEM. Three of the cross-loading were large enough to be considered meaningful
11 (SPPP7 on SPPP and SOPP, twice, and SOPP10 on SOPP and SPPP) and, of these three, in
12 the last instance the size of the cross-loading was larger than the loadings of items on
13 expected factors. In considering these instances, we note that Dunn et al (2006) similarly
14 found personal standards items to load on both personal standards and perceived coach and
15 parental pressure factors (it was the most common cross-loading observed in their study).
16 Dunn et al suggested that this may be because some respondents did not differentiate between
17 their own standards/expectations and those set by others. This may also be an issue for the
18 two items involved in the cross-loading here. As such, as validation work continues the cross-
19 loading of items SPPP7 and SOPP10 is another issue that may require scrutiny.

20 Evidence of the construct validity of the PPS-S was provided by correlations and
21 regression analyses using the S-MPS-2. As expected, SOPP was best characterised by
22 personal standards. In two of the three samples, concern over mistakes was also a significant
23 predictor. We consider this to indicate that SOPP adequately captures the duality of the
24 dimension when manifested more generally. That is, SOP is considered to be highly
25 motivating but also a vulnerability factor for motivation, performance, and psychological

1 difficulties (Flett & Hewitt, 2005, 2006). Examining whether demanding perfect performance
2 from oneself does indeed render athletes vulnerable to difficulties is an important avenue for
3 future research in terms of testing the construct validity of this dimension of the PPS-S.

4 Given the specificity of SOPP, we speculate that our narrower conceptualisation of SOP may
5 even be a more potent and proximal predictor of such difficulties in performance contexts.

6 SPPP was revealed to be characterised by concerns over mistakes and a sense of
7 external pressure. Again, this was as expected and can be considered to provide support for
8 the notion that SPPP encapsulates the core features of SPP generally. It is notable that the
9 regressions indicated that across the three samples SPPP was better predicted by perceived
10 coach pressure than parental pressure. The items of SPPP do not direct athletes to either
11 coaches or parents. We did, however, direct athletes to individuals whose “opinions they
12 valued” via the instructions to the items. The finding here suggest that respondents were
13 thinking of coaches more so than parents when responding to the items. We note, however,
14 that SPPP was not so highly correlated with perceptions of coach (or parental) pressure to
15 suggest that SPPP is redundant with these existing measures. Rather, overall, the findings
16 suggest that SPPP in part reflects perceptions of these important others but is sufficiently
17 independent so to reflect others (e.g., friends and family members) and neurotic tendencies
18 indicative of SPP generally.

19 The findings regarding OOPP were a little more mixed. In research outside of sport,
20 OOP tends to be positively correlated to most dimensions of the S-MPS-2 and its predecessor
21 (the Frost Multidimensional Perfectionism Scale, Frost et al., 1990), and is typically most
22 closely related to personal standards (e.g., Cox et al., 2002; Frost et al., 1993; Slaney et al.,
23 2001). There is little research to draw upon in sport regarding OOP. However, in a similar
24 manner, Dunn et al (2006) found OOP to be positively related to all subscales of the S-MPS
25 in one sample of athletes and related to only personal standards in another. The findings

1 regarding OOPP here were similar to previous research in that it was positively associated
2 with all dimensions of perfectionism. However, the prominence of personal standards in
3 relation to other dimensions was not evident. Instead, OOPP appeared to be characterised by
4 a broader array of dimensions and more clearly included perfectionistic concerns (i.e.,
5 concern over mistakes and perceived pressures). This was most apparent in the regressions
6 where concern over mistakes was the only consistent predictor of OOPP. In comparison to
7 OOP then, OOPP may be somewhat more distinct. In this regard, it may be noteworthy that
8 OOPP items focus more on denigration associated with imperfect performance (e.g., “I
9 criticise...” and “I have a lower opinion...”) whereas the original OOP items include a mix of
10 denigration and high standards or expectations (e.g., “I have high expectations for the people
11 who are important to me.”). Further insight into features of OOPP is clearly required and
12 might be provided by focusing on the issue of standards/expectations versus denigration.

13 **Conclusion**

14 The validation of a new instrument designed to measure multidimensional
15 performance perfectionism has begun in earnest. Here, we have reported on the first stage of
16 its validation across three studies involving multiple samples. Following its initial
17 development, exploratory and exploratory-confirmatory examination of its factor structure
18 and initial assessment of construct validity provided support for the instrument. Therefore,
19 early indication is that the PPS-S offers a reliable and valid measure of performance
20 perfectionism that due to its brevity can be easily included in future research. Here we
21 examined the PPS for use in adolescent and young adult athletes. We believe, however, that
22 as validation continues the PPS is likely to prove suitable for use in adults and in other
23 performance contexts (e.g., education and work).

24

25

References

- 1
- 2 Asparouhov, T. & Muthen, B. (2009). Exploratory structural equation modeling. *Structural*
- 3 *Equation Modeling, 16*, 397-438.
- 4 Bing, M. N., Whanger, J.C., Davison, H.K., & VanHook, J.B. (2004). Incremental validity of
- 5 the frame-of-reference effect in personality scale scores: A replication and extension.
- 6 *Journal of Applied Psychology, 89*, 150-157.
- 7 Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall
- 8 Regents.
- 9 Child, D. (2006). *The essentials of factor analysis*. New York, NY: Continuum International
- 10 Publishing Group.
- 11 Cohen, J. (1992). A power primer. *Psychological bulletin, 112*, 155-159.
- 12 Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality*
- 13 *of Life Research, 12*, 229–238.
- 14 Cox, B.J., Enns, M.W., & Clara, I.P. (2002). The multidimensional structure of
- 15 perfectionism in clinically distressed and college student samples. *Psychological*
- 16 *Assessment, 14*, 365-373.
- 17 DeVellis, R.F. (1991). *Scale development: Theory and applications*. Newbury Park: Sage
- 18 Publications, Inc.
- 19 Dunn, J.G.H., Causgrove Dunn, J., Gotwals, J.K., Vallance, J.K.H., Craft, J.M., & Syrotuik,
- 20 D.G. (2006). Establishing construct validity evidence for the Sport Multidimensional
- 21 Perfectionism Scale. *Psychology of Sport and Exercise, 7*, 57-79.
- 22 Dunn, J.G.H., Causgrove Dunn, J., & Syrotuik, D.G. (2002). Relationship between
- 23 multidimensional perfectionism and goal orientations in sport. *Journal of Sport and*
- 24 *Exercise Psychology, 24*, 376-395.

- 1 Dunn, J.G.H., Craft, J.M., Causgrove Dunn, J., & Gotwals, J.K. (2011). Comparing a
2 domain-specific and global measure of perfectionism in competitive female figure
3 skaters. *Journal of Sport Behavior*, 34, 25-46.
- 4 Dunn, J.G.H., Gotwals, J.K., & Causgrove Dunn, J. (2005). An examination of the domain
5 specificity of perfectionism among intercollegiate student-athletes. *Personality and*
6 *Individual Differences*, 38, 1439-1448.
- 7 Ericsson, K.A. & Simon, H.A. (1998). How to study thinking in everyday life: Contrasting
8 think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, &*
9 *Activity*, 5, 178-86.
- 10 Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use
11 of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-
12 299.
- 13 Flett, G.L., & Hewitt, P.L. (2002). Perfectionism and maladjustment: An overview of
14 theoretical, definitional, and treatment issues. In G. L. Flett & P. L. Hewitt (Eds.),
15 *Perfectionism: Theory, research, and treatment* (pp. 5-31). Washington, DC: American
16 Psychological Association.
- 17 Flett, G.L., & Hewitt, P.L. (2005). The perils of perfectionism in sports and exercise. *Current*
18 *Directions in Psychological Science*, 14, 14-18.
- 19 Flett, G. L., & Hewitt, P. L. (2006). Positive Versus negative perfectionism in
20 psychopathology A comment on Slade and Owens's dual process model. *Behavior*
21 *Modification*, 30, 472-495.
- 22 Flett, G.L., Hewitt, P.L., Blankstein, K.R., & Gray, L. (1998). Psychological distress and the
23 frequency of perfectionistic thinking. *Journal of Personality and Social Psychology*, 75,
24 1363-1381.

- 1 Frost, R.O., Heimberg, R.G., Holt, C.S., Mattia, J.I., & Neubauer, A.L. (1993). A comparison
2 of two measures of perfectionism. *Personality and Individual Differences, 14*, 119-126.
- 3 Frost, R.O., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of
4 perfectionism. *Cognitive Therapy and Research, 5*, 449-468.
- 5 Gotwals, J.K. & Dunn, J.G.H. (2009). A multi-method multi-analytic approach to establish
6 internal construct validity evidence: The Sport Multidimensional Perfectionism Scale 2.
7 *Measurement in Physical Education and Exercise Science, 13*, 71-92.
- 8 Hewitt, P.L. & Flett, G.L. (1991). Perfectionism in the self and social contexts:
9 Conceptualization, assessment and association with psychopathology. *Journal of*
10 *Personality and Social Psychology, 60*, 456-470.
- 11 Hewitt, P.L., & Flett, G.L. (2004). *Multidimensional Perfectionism Scale (MPS): Technical*
12 *manual*. Toronto, Canada: Multi-Health Systems.
- 13 Hewitt, P.L., Flett, G.L., Sherry, S.B., Habke, M., Parkin, M., Lam, R.W., ... Stein, M.B.
14 (2003). The interpersonal expression of perfection: Perfectionistic self-presentation and
15 psychological distress. *Journal of Personality and Social Psychology, 84*, 1303-1325.
- 16 Hill, A. P. (2016). Conceptualizing perfectionism: An overview and unresolved issues. In A.
17 P. Hill (Ed.), *The psychology of perfectionism in sport, dance, and exercise* (pp. 3-30).
18 London: Routledge.
- 19 Hunthausen, J., Truxillo, D.M., Bauer, T.N., & Hammer, L. (2003). A field study of frame-
20 of-reference effects on personality test validity. *Journal of Applied Psychology, 88*, 545-
21 551.
- 22 Jowett, G. E., Mallinson, S. H., & Hill, A. P. (2016). An independent effects approach to
23 perfectionism in sport, dance, and exercise. In A. P. Hill (Ed.), *The psychology of*
24 *perfectionism in sport, dance, and exercise* (pp. 85-149). London: Routledge.

- 1 Kidder, L., & Judd, C. (1986). *Research methods in social science*. New York, NY: CBS
2 College Publishing.
- 3 Kincaid, J. P., Fishburne, R. P., Rogers, R. L. & Chissom, B. S. (1975). *Derivation of new*
4 *readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease*
5 *Formula) for Navy enlisted personnel*. (No. RBR-8-75). Millington, TN: Naval
6 Technical Training Command Research Branch.
- 7 Lievens, F., De Corte, W., & Schollaert, E. (2008). A closer look at the frame-of-reference
8 effect in personality scale scores and validity. *Journal of Applied Psychology*, 93, 268-
9 279.
- 10 Mallinson, S.H., Hill, A.P., Hall, H.K., & Gotwals, J.K. (2014). The 2 × 2 model of
11 perfectionism and school- and community-based sport participation. *Psychology in the*
12 *Schools*, 9, 972-985.
- 13 Marsh, H. W., Hau, K. T., & Wen, Z. L. (2004). In search of golden rules: comment on
14 approaches to setting cutoff values for fit indexes and dangers in overgeneralising Hu &
15 Bentler (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, 11,
16 320-341.
- 17 Marsh, H.W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A.J.S., Trautwein, U., &
18 Nagengast, B. (2010). A new look at the Big-Five factor structure through exploratory
19 structural equation modeling. *Psychological Assessment*, 22, 471–491.
- 20 Marsh, H.W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A.J.S., &
21 Trautwein, U. (2009). Exploratory Structural equation modeling, integrating CFA and
22 EFA: Application to students' evaluations of university teaching. *Structural Equation*
23 *Modeling*, 16, 439-476.

- 1 Marsh, H. W., Nagengast, B., & Morin, A. J. (2013). Measurement invariance of big-five
2 factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce
3 vita effects. *Developmental psychology*, *49*, 1194.
- 4 Morgan, D.L. (1992). Designing focus group research. In M. Stewart, F. Tudiver, M. J. Bass,
5 E. V. Dunn, & P. G. Norton (Eds.), *Tools for primary care research* (pp. 177-193).
6 Newbury Park, CA: Sage.
- 7 Morin, A. J., & Maïano, C. (2011). Cross-validation of the very short form of the Physical
8 Self-Inventory (PSI-VS): Invariance across genders, age groups, ethnicities and weight
9 statuses. *Body Image*, *8*, 404-410.
- 10 Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide*. Los Angeles, CA: Muthén &
11 Muthén.
- 12 Myers, N.D., Chase, M.A., Pierce, S.W., Martin, E. (2011). Coaching efficacy and
13 exploratory structural equation modeling: A substantive-methodological synergy.
14 *Journal of Sport & Exercise Psychology*, *33*, 779-806.
- 15 O'Connor, B.P. (2000). SPSS and SAS programs for determining the number of components
16 using parallel analysis and Velicer's MAP test. *Behavior Research Methods*,
17 *Instruments, & Computers*, *32*, 396-402.
- 18 Slaney, R. B., Rice, K. G., Mobley, M., Trippi, J., & Ashby, J. S. (2001). The revised Almost
19 Perfect Scale. *Measurement and Evaluation in Counseling and Development*, *34*, 130-
20 145.
- 21 Stoeber, J., & Madigan, D.J. (2016). Measuring perfectionism in sport, dance, and exercise:
22 Review, critique, recommendations. In A. P. Hill (Ed.), *The psychology of perfectionism*
23 *in sport, dance, and exercise* (pp. 31-56). London: Routledge.

- 1 Stoeber, J., Otto, K., & Stoll, O. (2006). Multidimensional Inventory of Perfectionism
2 (MIPS): English version. *Unpublished manuscript*, School of Psychology, University of
3 Kent, UK. Available from <https://kar.kent.ac.uk/41560/>
- 4 Stoeber, J., & Stoeber, F.S. (2009). Domains of perfectionism: Prevalence and relationships
5 with perfectionism, gender, age, and satisfaction with life. *Personality and Individual*
6 *Differences, 46*, 530-535.
- 7 Tabachnick, B.G., & Fidell, L.S. (2001). *Using multivariate statistics*. Boston, MA: Allyn &
8 Bacon.
- 9 Worthington, R.L., & Whittaker, T.A. (2006). Scale development research: A content
10 analysis and recommendations for best practices. *The Counseling Psychologist, 34*, 806-
11 838.

12

13

14

15

16

17

18

19

20

21

22

23

24

25

Table 1. *Factor Solution for Final Exploratory Factor Analyses (Sample One and Two)*

Pattern coefficients from Oblmin (delta 0) Rotation	F1	F2	F3	h^2
1. I am tough on myself when I do not perform perfectly.	.053 / -.057	.060 / -.003	.745 / .650	.573 / .398
4. I put pressure on myself to perform perfectly.	-.070 / .001	-.033 / .004	.874 / .822	.719 / .676
10. I only think positively about myself when I perform perfectly. ¹	-.071 / .070	-.141 / -.084	.529 / .432	.414 / .234
11. To achieve the standards I have for myself I need to perform perfectly.	.136 / <u>.433</u>	-.065 / .060	.641 / .395	.577 / .455
2. People always expect more, no matter how well I perform.	.576 / .650	.105 / .167	.236 / .087	.482 / .430
7. People always expect my performances to be perfect.	.622 / .631	.006 / -.120	.171 / .146	.538 / .556
9. People view even my best performances negatively.	.539 / .566	-.169 / -.060	-.089 / -.110	.351 / .308
12. People criticise me if I do not perform perfectly.	.308 / .565	-.249 / <u>-.324</u>	.290 / -.026	.477 / .527
3. I have a lower opinion of others when they do not perform perfectly.	-.114 / -.061	-.864 / -.676	.074 / .118	.712 / .463
6. I am never satisfied with the performances of others.	.151 / -.005	-.765 / -.621	-.009 / -.129	.716 / .382
8. I criticise people if they do not perform perfectly.	.150 / .032	-.729 / -.744	-.101 / .096	.601 / .599
5. I think negatively of people when they do not perform perfectly.	-.066 / .096	-.737 / -.681	.123 / .030	.570 / .522
Eigenvalue	3.64 / 2.42	3.62 / 2.39	3.36 / 2.02	
Inter-factor correlation F1		-.367 / -.316	.593 / .382	
F2			-.499 / -.123	

1 *Note.* Sample one (n=321) to left. Sample two (n=229) to the right. Bold typeface denotes loadings above .30 on expected factors. Underlined typeface
2 denotes cross-loadings above .30. ¹In sample one this item was “I only think positively about myself when I meet the standards I have set for myself as an
3 athlete.”

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

1 Table 2. *Goodness of Fit Statistics and Information Criteria for CFA and ESEM (Samples Three, Four and Five)*

	χ^2	<i>df</i>	CFI	RMSEA	RMSEA 90% CI	SRMR	AIC	BIC	ABIC
Sample three									
CFA	114.987***	51	0.910	0.074	[.056, .092]	0.062	9679.282	9813.536	9689.928
ESEM	64.835***	33	0.955	0.065	[.041, .088]	0.031	9654.515	9850.733	9670.075
Sample four									
CFA	110.591***	51	0.899	0.075	[.056, .094]	0.058	8878.201	9008.177	8884.608
ESEM	95.909***	33	0.893	0.096	[.074, .119]	0.042	8877.823	9067.788	8887.187
Samples five									
CFA	127.805***	51	0.871	0.082	[.064, .100]	.070	9496.094	9629.148	9505.551
ESEM	72.546***	33	0.934	0.073	[.050, .096]	.037	9454.935	9659.399	9468.757

Note. CFA= Confirmatory factor analysis; ESEM = Exploratory structural equation modeling; *df* = Degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = Standardised Root Mean Square Residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; ABIC = Sample size adjusted BIC; ESEM were estimated with target oblique rotation; * $p < .001$. ** $p < .01$. *** $p < .05$.

5

6

7

1 Table 3. Standardized Factor Loadings for CFA and ESEM Solutions (Samples Three, Four, and Five)

Item	CFA			ESEM		
	Factor Loading	Uniquenesses	SOPP Factor Loading	SPPP Factor Loading	OOPP Factor Loading	Uniquenesses
1	.520*** / .512*** / .552***	.729*** / .738*** / .695***	.691*** / .512*** / .623***	-.143 / -.048 / -.117	.009 / .007 / .026	.591*** / .757*** / .649***
4	.661*** / .627*** / .833***	.563*** / .606*** / .306**	.754*** / .605*** / .909**	.029 / -.040 / -.016	-.056 / .107 / .007	.430*** / .603* / .181*
10	.540*** / .586*** / .514***	.708*** / .657*** / .736***	.285*** / .490*** / .277***	.212 / .146 / <u>.332**</u>	.136 / .021 / .069	.749*** / .660*** / .697***
11	.737*** / .826*** / .668***	.456*** / .317*** / .553***	.543*** / .855*** / .555***	.189 / .055 / .194	.027 / -.068 / -.051	.562*** / .260 / .595***
2	.710*** / .577*** / .474***	.496*** / .667*** / .775***	.202 / -.045 / .255**	.646*** / .632*** / .322***	-.079 / -.061 / -.016	.480*** / .656*** / .774***
7	.757*** / .648*** / .717***	.427*** / .581*** / .485***	.255 / <u>.336**</u> / <u>.320***</u>	.491*** / .372*** / .354***	.123 / .085 / .188	.481*** / .573*** / .535***
9	.524*** / .599*** / .536***	.725*** / .641*** / .712***	-.197* / -.166 / -.116	.765*** / .692*** / .669***	-.060 / .046 / .053	.556*** / .556*** / .561**
12	.611*** / .720*** / .666***	.626*** / .481*** / .557***	-.034 / .021 / .017	.562*** / .734*** / .681***	.170 / .007 / .065	.569*** / .440*** / .471***
3	.725*** / .789*** / .768***	.475*** / .378*** / .411***	.049 / .036 / .062	-.102 / .101 / -.215*	.634*** / .853*** / .945***	.494*** / .325*** / .267
6	.705*** / .657*** / .660***	.503*** / .569*** / .565***	-.023 / -.023 / .003	.125 / .098 / .083	.613*** / .608*** / .595***	.533*** / .576*** / .581***
8	.782*** / .777*** / .771***	.389*** / .396*** / .406***	.054 / -.009 / .010	-.181 / .071 / .194	.918*** / .728*** / .617***	.290*** / .420** / .438***
5	.745*** / .620*** / .591***	.444*** / .616*** / .651***	-.094 / -.041 / -.172*	.049 / -.009 / .090	.743*** / .641*** / .596***	.435*** / .611*** / .620***

2
3 Note. Sample three (n= 241) left. Sample four (n= 222) middle. Sample five (n = 252) right. Underlined typeface denotes meaningful cross-loadings (>.30).

4 SOPP = Self-oriented performance perfectionism; SPPP = Socially prescribed performance perfectionism; OOPP = Other-oriented performance perfectionism.

5 *** $p < .001$. ** $p < .01$. * $p < .05$.

1
2
3
4
5
6
7
8
9
10
11

Table 4. *Standardized Factor Correlations for the CFA and ESEM (Samples Three, Four, and Five)*

	SOPP	SPPP	OOPP
SOPP		.679*** / .597*** / .660***	.326*** / .361*** / .364***
SPPP	.458*** / .478** / .395***		.605*** / .522*** / .704***
OOPP	.242*** / .349*** / .300***	.555*** / .475*** / .578***	

Note. Confirmatory Factor Analysis (CFA) correlations (above the diagonal) and Exploratory Structural Equation Modeling (ESEM) correlations (below the diagonal). Correlations for Sample three (n= 241) left, Sample four (n= 222) middle, and Sample five (n = 252) right. SOPP = Self-oriented performance perfectionism; SPPP = Socially prescribed performance perfectionism; OOPP = Other-oriented performance perfectionism. *** $p < .001$. ** $p < .01$. * $p < .05$.

Table 15. Descriptive Statistics and Bivariate Correlations for PPS-S and S-MPS-2 (Samples Three, Four, and Five)

Subscale	Sample three											Sample four											Sample five												
	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9		
1 SOPP	4.61	1.16	.70									4.63	1.25	.75									4.58	1.23	.77										
2 SPPP	3.47	1.22	.46	.75								3.27	1.25	.45	.75								3.23	1.23	.44	.73									
3 OOPP	2.56	1.26	.28	.50	.83							2.34	1.24	.35	.43	.84							2.20	1.12	.29	.51	.81								
4 PS	3.02	0.82	.62	.42	.33	.84						2.92	0.87	.58	.49	.40	.87						2.80	0.75	.63	.27	.26	.79							
5 COM	2.73	0.86	.54	.57	.43	.66	.87					2.59	0.88	.48	.45	.42	.70	.86					2.49	0.80	.56	.45	.40	.66	.85						
6 PPP	2.38	0.95	.33	.54	.42	.55	.61	.92				2.19	0.88	.39	.46	.30	.61	.61	.92				2.05	0.79	.36	.37	.36	.58	.59	.89					
7 PCP	2.68	0.81	.38	.56	.34	.57	.65	.65	.82			2.42	0.92	.32	.51	.38	.65	.66	.68	.88			2.42	0.77	.33	.48	.37	.48	.54	.54	.82				
8 DAA	2.67	0.79	.28	.42	.35	.42	.60	.52	.57	.84		2.47	0.83	.35	.42	.27	.52	.63	.57	.66	.83		2.37	0.80	.31	.38	.35	.49	.59	.46	.48	.84			
9 ORG	2.89	0.99	.39	.25	.21	.59	.42	.39	.39	.32	.92	2.59	1.02	.36	.33	<u>.18</u>	.56	.41	.39	.43	.35	.91	2.49	0.97	.34	<u>.14</u>	<u>.14</u>	.50	.41	.46	.30	.35	.92		

Note. All bivariate correlations were significant, $p < .01$, except those underlined which were significant at $p < .05$; internal reliability (α) is displayed on the diagonal; SOPP = self-oriented performance perfectionism; SPPP = socially prescribed performance perfectionism; OOPP = other-oriented performance perfectionism; PS = personal standards; COM = concern over mistakes; PPP = perceived parental pressure; PCP = perceived coach pressure; DAA = doubts about actions; ORG = organization.

6

7

8

Table 6. Multiple Regressions of PPS-S Subscales on S-MPS-2 Subscales

PPS-S	S-MPS-2 subscale	Sample three				Sample four				Sample five			
		β	<i>B</i>	<i>S.E</i>	<i>p</i>	β	<i>B</i>	<i>S.E</i>	<i>p</i>	β	<i>B</i>	<i>S.E</i>	<i>p</i>
SOPP		$F(6, 200) = 25.16, p < .001, R^2 = .43$				$F(6, 171) = 17.85, p < .001; R^2 = .39$				$F(6, 195) = 25.52, p < .001; R^2 = .44$			
	Personal standards	.466***	.678	.121	.000	.493***	.688	.137	.000	.443***	.716	.127	.000
	Concern over mistakes	.360***	.493	.117	.000	.122	.170	.138	.222	.346***	.518	.124	.000
	Perceived parental pressure	-.074	-.093	.095	.329	.114	.160	.122	.193	-.058	-.091	.118	.443
	Perceived coach pressure	-.042	-.060	.118	.609	-.206*	-.273	.130	.036	.078	.122	.109	.265
	Doubts about action	-.074	-.111	.106	.299	.073	.108	.127	.397	-.145*	-.219	.104	.037
	Organisation	.013	.015	.078	.845	.074	.090	.090	.317	.018	.023	.079	.775
SPPP		$F(6, 198) = 24.43, p < .001; R^2 = .43$				$F(6, 169) = 13.96, p < .001; R^2 = .33$				$F(6, 192) = 15.96, p < .001; R^2 = .33$			
	Personal standards	-.026	-.039	.125	.756	.185	.256	.141	.072	-.225	-.225	.143	.117
	Concern over mistakes	.295**	.416	.123	.001	.015	.021	.143	.882	.390**	.390	.136	.005
	Perceived parental pressure	.242**	.313	.100	.002	.106	.147	.126	.246	.163	.163	.134	.225
	Perceived coach pressure	.261**	.387	.125	.002	.246*	.322	.133	.017	.591***	.591	.123	.000

Doubts about action	-.003	-.005	.111	.963	.087	.126	.132	.342	.137	.137	.116	.241
Organisation	-.059	-.073	.084	.387	.053	.063	.093	.495	-.143	-.143	.088	.106
OOPP	$F(6, 196) = 8.45, p < .001; R^2 = .21$				$F(6, 171) = 7.75, p < .001; R^2 = .21$				$F(6, 187) = 9.74, p < .001; R^2 = .24$			
Personal standards	.057	.087	.151	.566	.224*	.317	.158	.046	-.048	-.072	.141	.609
Concern over mistakes	.200	.286	.146	.051	.271*	.384	.160	.017	.229*	.328	.141	.021
Perceived parental pressure	.235*	.307	.119	.010	-.073	-.104	.141	.464	.160	.236	.132	.076
Perceived coach pressure	.009	.014	.151	.925	.154	.209	.150	.165	.167*	.244	.123	.048
Doubts about action	.053	.083	.135	.538	-.034	-.051	.148	.732	.130	.186	.118	.117
Organisation	-.043	-.053	.098	.589	-.096	-.119	.104	.253	-.118	-.136	.088	.122

Note. SOPP = self-oriented performance perfectionism; SPPP = socially prescribed performance perfectionism; OOPP = other-oriented performance perfectionism;

* $p < .05$. ** $p < .01$. *** $p < .001$.