Title page

Title: Psychometric Assessment of the Mental Health Continuum-Short Form in Athletes: a Bi-factor Modelling Approach.

Authors: Stephen Shannon[1, 2], Mark Shevlin[3] and Gavin Breslin [2, 3].

[1]Sport and Exercise Sciences Research Institute, Ulster University Jordanstown Campus, Shore Road, Newtownabbey, Northern Ireland, BT370QB

[2]Bamford Centre for Mental Health and Well-Being, Ulster University, Magee Campus, Derry, Northern Ireland, BT487JL.

[3]School of Psychology, Ulster University, Coleraine Campus, Northern Ireland, BT52 1SA.

Corresponding author

Dr. Stephen Shannon (email: s.shannon@ulster.ac.uk).

Keywords: well-being; measurement; mental illness; ill-being; methods.

Abstract

Aim: A recent mental health in sport consensus statement (Breslin et al., 2019) advocates Keyes (2002) two-continua model with an associated Mental Health Continuum (MHC) instrument to assess mental health in athletes. However, there remains statistically inconsistent usage of the MHC in athletes, so further exploration of the MHC's psychometric factors is required.

Methods: Athletes (*N*=1,097) aged 32.63 (SD =11.16) comprising 603 females (55.7%) and 478 males (44.3%), completed the 14-item MHC-short form (MHC-SF), alongside validated measures of anxiety and depression. Five confirmatory factor analytic (CFA) and bi-factor models were developed based on extant research and theory.

Results: Overall, first-order models did not fit the data, but a bi-factor structure with a ‘general’ positive mental health factor, and three specific factors (‘Hedonic well-being’, ‘Social well-being’ and ‘Psychological well-being’) fitted the data well and was deemed the superior model.

Conclusions: A bi-factor model of the MHC-SF is recommended comprising a composite score alongside specific factors of hedonic, social and psychological well-being.

Keywords: Well-being; psychology; confirmatory factor analysis; validity; sport.

Psychometric Assessment of the Mental Health Continuum-Short Form in Athletes: a Bi-factor Modelling Approach.

In response to the preponderance and legacy of the illness-based model of mental health, Keyes (2002) presented a theory to reclaim 'mental health' as a positive construct characterised by 'flourishing'. Keyes (2005) later examined axioms of multidimensional mental health, presenting a two continua model wherein mental health and mental illness coexist as two distinct, but correlated, unipolar dimensions. To this end, Keyes et al. (2008) considered 'flourishing' as a diagnosable presence of positive mental health, with 'languishing' as the absence of positive mental health. While the determinantal societal effects of mental illness (e.g., depression, anxiety) have been publicly understood and of clear significance to policy makers for generations (Jones, 2013), it is only within the last fifteen years that positive mental health (or well-being) has been considered an essential aspect of public health (Huppert, 2009). Indeed, educational success, living in a safe neighbourhood, family support, and economic prosperity correlate with positive mental health (United Nations, 2015).

Within the context of competitive sport, mental health is a rapidly emerging research field, to the extent that global sporting bodies (e.g., The International Olympic Committee [IOC]), national sport organisations, and researchers have recently developed action plans or consensus statements to safeguard athlete mental health (Vella & Swann, 2021). There are an abundance of elite athlete mental health consensus statements (e.g., Henriksen et al., 2020), including by the IOC (Reardon et al., 2019). Mirroring the messages of, and responding to, recommendations among consensus statements for elite athlete's mental health, the IOC recently developed the Sport Mental Health Assessment Tool 1 (SMHAT-1) and Sport

Mental Health Recognition Tool 1 (SMHRT-1) (Gouttebarge et al., 2021). Notably, both measures and the field at large remain focused on mental illness symptoms and concepts. One international consensus statement focused on non-elite athletes (i.e., Breslin et al., 2019), who comprise the vast majority of sporting participants (Vella & Swann, 2021). As such, it was and remains pertinent that Breslin et al. (2019) recommended that all competitive athletes' mental health be viewed from Keyes' (2002) theoretical perspective. Indeed, the view put forward by Breslin et al. (2019) and others (e.g., Uphill, Sly & Swain, 2016) is that Keyes' (2002) model is theoretically robust, and reflective of a multidimensional mental health construct comprising well-being, broadening the existing dominant focus on mental illness.

Indeed, in a review of existing well-being measures in sport Giles et al. (2020) argued that researchers have typically employed proxy indicators of well-being (e.g., life satisfaction, affect, subjective vitality) without sufficient theoretical basis. A lack of theoretically guided research ultimately hinders progress on understanding the correlates that influence an athlete's overall mental health (Lundqvist & Sandin, 2014). As such, there is need for theoretically derived, valid measurement tools to screen athletes' mental health as conceptualised by Keyes (2002; 2005) two continua model (Uphill, Sly & Swain, 2016). Having such instruments is crucial for assessing types of suitable care for athletes, intervention effectiveness, and providing policymakers with valid and reliable data (Breslin et al., 2017; Breslin & Leavey, 2019; Giles et al., 2020).

Keyes' (2002; 2005) Mental Health Continuum (MHC) instrument was constructed via philosophical traditions and contemporary theories (e.g., Diener & Emmons, 1984; Ryan & Deci, 2000). The mental health (or well-being) continua derives its structure and items from hedonic (i.e., Diener's subjective well-being), social (i.e., Keyes' social functioning), and eudemonic (i.e., Ryff, Self-Determination Theory) theories. The mental illness continua include latent measures such as major depressive order, panic, generalized anxiety disorder and alcohol dependence as defined by the Diagnostic and Statistical Manual of Mental Disorders. From Keyes' (2005) perspective, a number of possible mental health profiles emerge, for example an athlete could simultaneously experience positive mental health along with mental illness. Contrastingly, an athlete could be free from mental illness, but experience low levels of mental health (i.e., languishing).

Keyes (2002; 2005) long-form MHC instrument comprised of 42-items measuring three factors of hedonic (i.e., positive affective states, life satisfaction), eudemonic (e.g., psychological functioning, sense of purpose), and social (i.e., relationships, integration) mental health. However, most researchers opt for the Mental Health Continuum-Short Form (MHCSF; Keyes et al., 2008), likely due to its retention of psychometric validity, whilst obtaining practical ease and lessening participant time burden (Jovanović, 2015). The 14-item MHC-SF includes three items (two for positive emotions, and one for life satisfaction) in the hedonic construct; six items for the eudemonic (or psychological) construct; and, five items for the social construct. From its inception, the MHC-SF is a leading mental health instrument in public mental health research (Longo et al., 2020), including more recent epidemiological studies among athletes (McGivern, Shannon & Breslin, 2021).

However, Jovanović (2015) initially questioned the widescale adoption of the default three-dimensional structure of MHC-SF. Indeed, several studies have reported either marginally acceptable (Joshanloo & Jovanović, 2017) or unacceptable (Jovanović, 2015) model fit indices for a first-order three-factor solution. Moreover, among the studies testing the measurement properties of the MHC-SF with athletes, one study among adolescent non-elite athletes found an adequate fit for the three-factor model only following the removal of three items (Salama-Younes, 2011); another study solely among collegiate athletes revealed an unacceptable fit (Foster & Chow, 2019). Indeed, despite any prevailing statistical evidence, several athlete mental studies have treated the instrument as a composite score suggesting a unitary construct (Vella et al., 2020; McGivern, Shannon & Breslin, 2021). Such limited sample compositions and issues of model misfit require solutions and clarity, as an instrument's validity informs clinical practice, research, and policy decisions (Park, Han & Cho, 2011; Fried, 2017).

Confirmatory Factor Analysis (CFA) encompasses specified correlations between observed questionnaire items and latent variable(s). Through inspection of conventional fit statistics, researchers can determine the strength of evidence for a psychometric instrument's ability to capture its underlying 'true' or 'natural' construct(s) (Schreiber et al., 2006). Specifically, using CFA researchers can assess competing CFA models that include unidimensional (i.e., one underlying construct) and first order (i.e., correlated sub-dimensions) structures (Jackson et al., 2009). Furthermore, confirmatory bi-factor modelling (CBFA) permits items to correlate with a general factor (e.g., mental health) alongside sub-dimensions, or specific factors (Reise, 2012), with the caveat that additional bi-factor specific calculations are warranted alongside conventional fit statistics (Rodriguez, Reise & Haviland, 2016). It has been proposed that a sound measure should display nomological validity, which pertains to the correlation between the measured construct and further constructs within the same theory (e.g., Hagger & Chatzsarantis, 2009), for example, Keyes' (2002) hypothesised correlation between mental well-being and mental illness.

In view of the above limited evidence for the default three-factor MHC-SF, several authors re-specified the structure among general populations, and tested alternative CFA models including CBFA (Jovanović, 2015). Several studies replicated Jovanović's (2015) methods revealing that a bi-factor model (comprising one general mental health, and three specific factors) to be superior (see, De Bruin and Du Plessis 2015; Hides et al. 2016; Jovanovic´ 2015). It is methodologically advised to test competing psychometric measurement models among diverse, representative, samples (Park, Han & Cho, 2011). Yet, to our knowledge, no such studies have assessed a CBFA of the MHC-SF in athletes despite the measure’s widescale and growing use with athletes. Given Breslin et al.’s (2019) consensus statement advocated Keyes’ (2002) theory, and limited model fit evidence exists for the MHC-SF among narrow athlete samples (i.e., Salama-Younes, 2011; Foster & Chow, 2019), there is a need for a more comprehensive psychometric assessment of the MHC-SF among a diverse athletic sample.

Hence, the aim of this study was to assess competing CFA and CBFA measurement models of the MHC-SF, and its psychometric properties (i.e., nomological validity) across a large, demographically diverse sample of athletes representing a range of competitive sports (e.g., co-active team sports, individual athletic sports) and levels (e.g., elite, semi-athlete, amateur). We specified measurement models as outlined in Figure 1 that were based on extant research (Jovanović, 2015) and consistent with Keyes’ (2002) conceptualisation of mental health and mental illness.

Place Figure 1 here

Methods

Study Design, Recruitment and Participants

Ethical approval was granted by [redacted] Research Ethics Filter Committee. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement was used in the design of the current cross-sectional survey of athletes. Inclusion criteria was based on informed consent, being ≥18 years old, and participants confirming their athlete status using a widely used screening item (e.g., Shannon et al., 2019; Breslin et al., 2019) based on the definition of sport: ‘are you an athlete involved in a structured, competitive physical activity?’ (Rejeski & Brawley, 1996).

Recruitment involved a snowball sampling method wherein an encrypted online questionnaire link produced through SurveyMonkey software was distributed to a range of Twitter and social media outlets, sports club secretaries, and sporting organisations. Several sports organisations responded and distributed online links accordingly to followers and subscribers. Data derived from online psychometric collection methods have been shown to yield sound psychometric reliability and validity estimates in comparison with paper-based surveys, and show an added benefit of reducing attrition and false/missing responses (Lonsdale et al, 2006). Data was collected from January 2019 to March 2021 and took approximately ten minutes to complete. Demographic questions (i.e., gender, age, country), and sporting characteristics (i.e., individual or team sport) were collected.

Subsequently, data was collected from 1,097 participants comprising 603 females (55.7%) and 478 males (44.2%), with one participant (0.1%) indicating ‘other’ for gender. The mean age of participants was 32.63 (SD =11.16) with most identifying as Irish (44.2%), followed by Canadian (27.4%), British (19.3%) and others (e.g., American, Australian). The largest sport represented among the athletes was equestrian (34.3%), followed by rugby (28.7%), hockey (5.3%) and others (e.g., Running, Gaelic sports). Further, 53.3% of the sample participated in individual sports, whereas 46.8% took part in coactive team sports. The vast majority (79.7%) of the participants classified themselves as non-elite (e.g., amateur, local/community leagues) while 13.6% were elite (i.e., professional, international), and 6.7% were semi-elite (e.g., semi-professional). Among those who responded to an item regarding mental illness history (*n*= 891, 81.2%), 51.9% indicated they had not experienced mental illness, 39.6% had experienced mental illness, and 8.5% answered that they did not know or were unsure.

2.2 Outcome Measures

Mental Health Continuum- Short Form (MHC-SF)

Respondents completed the Mental Health Continuum - Short Form (MHC-SF: Keyes et al., 2008), which assesses the positive mental health dimension of Keyes (2005) two-continua model. As described earlier, the 14-item scale is theorised (Keyes, 2002) to derive hedonic (i.e., items 1-3), social (i.e., items 4-8) and psychological (i.e., items 9-14) well-being dimensions. The recall period for the MHC-SF is 'over the past month', wherein respondents rate the frequency of every feeling (e.g., happy) or experience (e.g., that you had warm and trusting relationships) on a 6-point Likert scale ranging from 'Never' (0) to 'Every day' (5). Total scores can range from 0-70, with higher scores indicating positive mental health. High comprehension, internal validity and cross-cultural reliability has been shown for the MHC-SF (Lamers et al., 2011). Consistent with previous research (Lamers et al., 2011; Ferentinos et al., 2019), the scale showed high internal consistency (Cronbach's α=.94),

Depression

Depression symptoms were assessed using the eight-item version of The Patient Health Questionnaire (PHQ-8: Kroenke et al., 2009). The PHQ-8 is a well-established diagnostic and severity measure for major depressive disorders in large clinical and non-clinical samples

(Razykov, Ziegelstein, Whooley & Thombs, 2012), and has demonstrated sound psychometric properties (Wu et al., 2019). Respondents indicated the number of days in the past two weeks in which they experienced a particular depressive symptom (e.g., anhedonia, hopelessness) on a 4-point Likert scale, ranging from ‘Not at all’ (0) to ‘Nearly every day’ (3). Possible scores range from 0-24, with higher scores representing greater severity of depression. Cronbach’s α=.87 in the present sample.

Anxiety

The seven-item Generalized Anxiety Disorder (GAD-7: Spitzer et al., 2006) scale was used as a measure of anxiety. Using a two-week recall period, respondents indicate the degree to which they have been bothered by anxious feelings (e.g., restlessness, afraid as if something might happen) with a 4-point Likert scale, ranging from ‘Not at all’ (0) to ‘Nearly every day’ (3). Sound psychometric properties and diagnostic efficacy have been shown for the GAD-7 among large clinical and non-clinical samples (Löwe et al., 2008), including online study methodologies (Donker et al., 2011). GAD-7 scores range from 0-21, with higher scores representing increased anxiety symptoms. Cronbach’s α=.92 in the present sample.

Resilience

Resilience was measured through the six-item Brief Resilience Scale (BRS) (Smith et al., 2008). Questions were anchored in a 5-point Likert scale (1-strongly disagree to 5- strongly agree) and inquired on “bounce-back-ability” during adversity (e.g., “I tend to bounce back quickly after hard times”). Scores are averaged and range from 0 to 5, with higher scores reflecting stronger resilience. Cross-cultural reliability and validity have been demonstrated for the BRS (Smith et al., 2008; de Holanda Coelho, Hanel, Medeiros Cavalcanti, Teixeira Rezende & Veloso Gouveia, 2016). Cronbach’s α=.57 in the present sample.

Data Analysis

Prior to main analyses, data was inspected for missing responses and outliers. As 4.1% of data was missing on the MHC-SF, Little's MCAR test was calculated and revealed data was missing completely at random (p > .05). Missing data was therefore estimated using multiple imputation function in the SPSS (version 25). Data was fully labelled and exported to AMOS (version 24) to assess the latent structure of the 14-item MHC-SF.

Five competing confirmatory factor analysis (CFA) and confirmatory bifactor analysis (BCFA) models were specified based on model iterations by Jovanović (2015) (see Figure 1 for a visual representation) and estimated using robust maximum likelihood estimation (see Table 1). Models included CFA on a unidimensional structure (Model 1); a first order with two correlated factors (Model 2); a first order model with three correlated factors (Model 3); a BCFA comprising a general factor and two orthogonal specific factors (Model 4), and lastly; a BCFA comprising a general factor and three orthogonal specific factors (Model 5).

The performance of the competing measurement models was assessed through comparison of multiple recommended goodness-of-fit indices (Hu & Bentler, 1999). The Chi-Square [χ^2] goodness-of-fit index was reported, however given that χ^2 is sensitive to large sample sizes (Bentler, 1990) we approached this value with caution. The Normed Fit Index (NFI) Tucker-Lewis Index (TLI) and Comparative Fit Index (CFI) were reported, and values of >.90 or >.95 were considered as acceptable or good-to-excellent model fit, respectively. The root mean square error of approximation (RMSEA) values were calculated, with a cut-off point of .08 or below considered acceptable. Additionally, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were assessed. Improved model performance was observed when AIC and BIC values were lower in comparison to other models. Lastly, the recommendations of Comrey and Lee (1992) were adopted for determining the strength of factor loadings (i.e., <.30 = poor; >.45 = fair; >.55=good; >.63 =

very good, and; >.71 = excellent). Models were tested with 5000 Bollen-Stine bootstraps to improve the accuracy of model parameters (Byrne, 2001).

In the case where the bi-factor model was considered the 'best' fit, further assessment of the general and specific factors is required (Rodriguez, Reise & Haviland, 2016) so supplementary BCFA fit statistics were calculated using Dueber's (2017) software. Specifically, omega reliability (ω; i.e., proportion of common item variance explained by the general and specific factors), omega hierarchical (ωH; i.e., proportion of variance within the items attributable to the general or specific factors, controlling for the specific and general factor), relative omega (ωR: i.e., proportion of variance attributable to the general factor independent of the specific factors, and specific factors independent of the general factor), and index *H* (i.e., how a set of items represents a latent variable, and the likelihood of that latent variable replicating across studies) were calculated. Omega coefficients and index *H* values range from 0-1, and values > .80 reflect satisfactory reliability and replicability (Rodriguez et al., 2016). We also reported the item explained common variance (I-ECV), which reflects the extent to which an item's responses are accounted for by variation on the latent general dimension alone. When I-ECV are > .80 or .85, a unidimensional structure for the item is likely (Stucky & Edelen, 2015). A second table comprising the CBFA fit statistics was produced, and a third table for the retained model describing the items and corresponding factor loadings.

Nomological validity assessments were determined using hypotheses from Keyes (2002) two-continua model of mental health. Based on Keyes (2002) theory we hypothesised that the mental ill-being (i.e., depression and anxiety) outcomes would be inversely correlated with the composite mental health score, whereas the BRS would be positively correlated. Additionally, we inspected the correlations between the retained specific sub-factors and the GAD-7, BRS and PHQ-8 whilst controlling for the composite score to determine their relative external validity contribution. Pearson’s Product-Moment Correlation (*r*) with alpha significance set at $p < .05$ was calculated for two tables as per above, considering values of .0 - 0.3 as weak, .31 - .70 as moderate, and .71 and above as strong (Field, 2013).

Results

CFA Fit Statistics

Fit statistics for the competing measurement models are presented in Table 1. The χ^2 values were all significant, likely due to the large sample size, and therefore did not lead to rejection of the models (Tanaka, 1987). In Model 1, factor loadings were good-to-excellent ranging from .57 (item 8) to .82 (item 14), all statistically significant ($p < .05$), and indicated some supporting evidence for an overarching general mental health factor. However, all CFA fit indices were below or above the recommended thresholds by Hu and Bentler (1999). Similarly, in Model 2 all factor loadings were statistically significant, ranging from .58 (item 8) to .87 (item 2), and the covariation pathway between the two correlated factors was .84. Whilst minor fit improvements were observed in comparison to Model 1, all fit statistics were again below or above the recommended thresholds.

The default Model 3 on the other hand, comprising three correlated factors displayed some acceptable fit statistics, namely NFI and CFI values of > .90, and marginally TLI (i.e., 916). However, the RMSEA was above .08 (i.e., .088). AIC and BIC statistics continued to decline, and all factor correlations were statistically significant, and elevated in comparison to Models 1 and 2, ranging from .65 (item 8) to .87 (item 2). By conventional standards, Model 3 with three correlated factors showed marginally acceptable factorial validity.

Table 1: Fit Statistics for the Competing CFA and CBFA Models Tested

Model		χ^2	df	CFI	TLI	NFI	RMSEA	BIC	AIC
1	Unidimensional	1424.342	77	.852	.825	.845	.126	1620.351	1480.342
2	First-order (2-factor)	1000.864	76	.899	.879	.891	.105	1203.874	1058.864
3	First-order (3-factor)	695.751	74	.932	.916	.924	.088	912.761	757.751
4	Bi-factor (2-specific)	453.519	63	.957	.938	.951	.075	747.533	537.519
5*	**Bifactor (3-specific)**	**336.829**	**63**	**.970**	**. 957**	**.963**	**.063**	**630.843**	**420.829**

Note: *Chosen as best fitting model

However, inspection of BCFA Models 4 and 5 showed further improvements regarding CFA fit statistics and prediction of item variance. Model 4 comprising a general positive mental health factor, and two specific factors of eudemonic and hedonic well-being, yielded excellent NFI and CFI values of >.95, and an RMSEA value of .075 (see Table 1). Additionally, AIC and BIC values continued to reduce, and all but one (i.e., item 4) of the specific factor loadings were statistically significant, alongside all general factor loadings that were statistically significant. Notably, however, the loadings of items 4-8 on the eudemonic well-being (EW) specific factor were in a negative direction, some as large as -.50. Such associations suggest that those items have a negative contribution to the factor and would thus require subtraction in any model calculations. Significantly, items 4-8 constituted the specific social well-being factor specified in Model 5, and given the inconsistencies in the direction of associations with the general and specific factor items, suggested the possibility of a distinct specific factor, as identified in Model 5.

To this end, the superior performance of Model 5 was evident in excellent fit statistics (CFI = .97, TLI = .96, NFI = .96) values that outperformed the aforesaid models, as did the RMSEA value at .063. AIC and BIC were at their lowest observed levels across all models. Aside from item 4 (i.e., ‘that you had something important to contribute to society’) all specific factor loadings were statistically significant, and in a positive direction, ranging from .10 (item 5) to .57 (item 8) (see Table 3). All general factor loadings were statistically significant, good-to-excellent, and ranged from .59 (item 10) to .82 (item 14).

Bi-factor fit statistics

When calculated in Dueber's (2017) bi-factor software, Model 5 showed, general and specific factor ω values of >.80 (see Table 2), and the majority (i.e., 9/ 14) of the I-ECV item values were <.80 rather than >.80 (see Table 3), suggesting some contribution of a multi-dimensional structure. However, ωH and ωR remained relatively low in relation to Rodriquez et al.'s (2017) benchmarks, as did *H*, suggesting a need for caution.

Table 2: Bi-factor indices calculator (Dueber, 2017) indicating reliability and construct replicability for the competing bi-factor models 4 and 5.

Factor	**ω**	**ωH**	**ωR**	***H***
General Factor	.947	.864	.913	.930
Hedonic well-being	.887	.227	.256	.416
Social well-being	865	.184	.213	.532
Psychological well-being	.887	.177	.200	.457

Note: GF= general factor; HW=hedonic well-being; EW=eudemonic well-being; SW=social well-being; PW=psychological well-being; ECV and PUC values are calculated at the model-level, rather than construct-level.

Taken collectively, a somewhat contradictory picture emerged from the CFA and BCFA model analyses. That is, by conventional standards the only marginally acceptable CFA model included the three correlated factors (i.e., Model 3). Yet, despite the BCFA Model 5 outperforming all models (see Table 1), the bifactor fit statistics (see Table 2) showed a fairly strong overarching general mental health factor with relatively weak specific factors. Hence, we propose the retention of Model 5, using a cautious approach in the calculation of both general and subfactor scores.

In doing so, and applying Keyes (2002) figurative labels to the factors, Model 5 comprised a strong general ‘positive mental health’ factor, and three specific factors labelled: ‘Hedonic’ (items 1-3), ‘Social (items 4-8) and ‘Psychological’ (items 9-14) well-being. As visually illustrated in Figure 1, paths between items and the ‘GF’ symbol refer to loadings on the general positive mental health construct, whereas item loadings onto HW (Hedonic Well-Being), SW (Social Well-Being), and PW (Psychological Well-Being) represent the specific factors.

Nomological validity

The correlation matrix for the retained model with the study outcomes is detailed in Table 4. All correlations were statistically significant at $p < .001$. Relating specifically to the correlations between MHC-SF factors (general and specific) and the study outcomes, *r* ranged from .17 to -.57. The composite score representing the general well-being factor showed moderate inverse correlations with depression (r = -.57) and anxiety (r = -.31), and a weak positive correlation with resilience (r = .22). Table 4 also illustrates significant weak-to-moderate correlations between specific subfactors and study outcomes with *r* ranging from .17 to -.56.

Demonstrating some added contribution of retaining the bifactor model, correlations between the specific sub-factors and study outcomes, independent of the controlled association between the study outcomes and composite MHC-SF score, and while weak, showed several incidences of statistical significance. Namely, and as identified in Table 5, the HW factor was negatively associated with depression (r = -.17) and anxiety (r = -.11); SW was surprisingly positively associated with depression (r = .10), and negatively associated with resilience (r = -.07); and PW was positively associated with anxiety, albeit weakly (r = .08).

Table 3: Retained bifactor model for MHC-SF, including instrument items, factor labels, and loadings with I-ECV values.

Item number and description	**General factor loading**	**I-ECV**	**Specific factor**	**Specific factor loading**	***Item R^2***
1. happy	.678*	.640$^{\vee}$	HW	.509*	.718
2. interested in life	.755*	.762$^{\vee}$	HW	.422*	.748
3. satisfied with life	.763*	.820$^{\wedge}$	HW	.358*	.710
4. that you had something important to contribute to society	.803*	.00$^{\wedge}$	SW	.014	.645
5. that you belonged to a community (like a social group, or your neighbourhood)	.705*	.979$^{\wedge}$	SW	.104*	.509
6. that our society is a good place, or is becoming a better place, for all people	.602*	.603$^{\vee}$	SW	.488*	.600
7. that people are basically good	.575*	.568$^{\vee}$	SW	.501*	.582
8. that the way our society works makes sense to you	536*	.470$^{\vee}$	SW	.569*	.611
9. that you liked most parts of your personality	684*	.726$^{\vee}$	PW	.420*	.644
10. good at managing the responsibilities of your daily life	.591*	.643$^{\vee}$	PW	.440*	.543
11. that you had warm and trusting relationships with others	.653*	.799$^{\vee}$	PW	.328*	.535
12. that you had experiences that challenged you to grow and become a better person	.603*	.866$^{\wedge}$	PW	.237*	.420
13. confident to think or express your own ideas and opinions	.667*	.762$^{\vee}$	PW	.373*	.584
14. that your life has a sense of direction or meaning to it	.815*	.939$^{\wedge}$	PW	.208*	.707

Note: *= statistically significant ($p < .05$); all R^2 values were statistically significant; HW = hedonic well-being specific factor; SW= social well-being specific factor; PW = psychological well-being specific factor; I-ECV=item-level explained common variance via the general factor; $^{\wedge}$ = where I-ECV of >0.80 suggesting a reliable unidimensional structure for item; $^{\vee}$= where I-ECV of <0.80 suggesting some contribution of a multidimensional structure for item.

Table 4: Correlation matrix for the retained model factors and study outcomes.

Variables	1	2	3	4	5	6	7
1. HW	1.000						
2. SW	.664*	1.000					
3. PW	.715*	.703*	1.000				
4. MHC_t	.836*	.899*	.927*	1.000			
5. Resilience	.203*	.166*	.212*	.215*	1.000		
6. Depression	-557*	-.477*	-.528*	-.573*	-.188*	1.000	
7. Anxiety	-.317*	-.277*	-.261*	-.311*	-.171*	.651*	1.000

Note: HW = hedonic well-being specific factor; SW= social well-being; PW = psychological well-being specific factor; MHC_t= Mental health continuum total score; *= $p < 0.01$

Table 5: Correlation matrix for the retained specific subfactors and study outcomes, whilst controlling for the correlation between the composite MHC-SF score and study outcomes.

Variables	1	2	3	4	5
1. HW	1.000				
2. SW	-.361**	1.000			
3. PW	-.289**	-.782**	1.000		
4. Resilience	0.044	-.065*	0.035	1.000	
5. Depression	-.172**	.104**	0.01	-.081**	1.000
6. Anxiety	-.109**	0.007	.078*	-.112**	.607**

Note: HW = hedonic well-being specific factor; SW= social well-being; PW = psychological well-being specific factor; *= $p < 0.05$; **= $p < 0.01$

Discussion

Competing psychometric models of the MHC-SF were assessed in the current study comprising a large, demographically diverse, representative sample of athletes, that included the novel specification of bi-factors models. While the default first-order three factor solution showed marginally acceptable fit statistics, the best representation of the MHC-SF pertained to a bi-factor model, comprising a strong general mental health factor, alongside relatively weaker, but relevant, specific factors of Hedonic, Social and Psychological Well-Being. Such findings are consistent with Keyes' (2002; 2005) early theorising of a multi-dimensional mental health construct, with the added existence of an overarching mental health factor, as supported in studies amongst general and clinical populations (De Bruin & Du Plessis 2015; Hides et al., 2016; Jovanovic, 2015). Further, robust contributions of the general mental health factor were evident in the nomological validity assessments, and when statistically controlled for, relatively weak, but significant associations were revealed for the specific factors.

In the present study we specified a unidimensional and higher-order two, and three-factor solution as originally hypothesised by Keyes (2002). Athlete mental health studies tend to apply unidimensional (Vella et al., 2020; McGivern, Shannon & Breslin, 2021) or three-factor solutions (Salama-Younes, 2011; Foster & Chow, 2019), despite any converging, population-specific, evidence for either model. The limited fit statistics for the unitary or higher-order models presented in Table 1, particularly support several studies suggesting the need for improvement in the factor structure of the MHC-SF (i.e., De Bruin and Du Plessis 2015; Hides et al. 2016; Jovanovic´ 2015). Indeed, in one of the few MHC-SF factor analysis studies amongst athletes (albeit items were modified for sport-specific mental health) Foster & Chow (2019) outlined that an adequate fit for the three-factor solution would only be found if residual errors were correlated. A further study amongst athletes (Salama-Younes, 2011) removed five of the items to achieve adequate fit.

Correlating residual error terms is a controversial practice in factor analysis studies (Gerbing & Anderson, 1984). Some authors (Reise, 2012) suggest that unless clear semantic and/or theoretical overlap is evident, the correlation of error terms (both among and across subfactors) equates to an unanalysed association and essential omission of a theoretically meaningful variable(s). While Foster and Chow (2019) contended that all their correlated error terms loaded onto the specific social well-being factor, correlating item error terms within every specific MHC-SF factor (or across factors) would likely yield a much-improved model fit due to model saturation (Hermida, 2015). However, given little semantic, theoretical, or methodological grounds, we would advise against correlating error terms in further research.

Additionally, in reviewing our findings we examined Salama-Younes' (2011) decision to remove three items for the psychological well-being specific factor, and two items from the social well-being factor, a practice often referred to as "scale purification". Wieland et al. (2017) argued that scale purification should be made through a careful balance of judgmental and statistical criteria. While statistical criterion has been discussed earlier, judgmental criteria is based on a qualitative assessment of the appropriateness of survey items to reflect theoretical interpretation (Carpenter et al., 2017). Upon inspection of item wordings (see Table 3), we note that Salama-Younes' (2011) removed items reflective of personality, sense of purpose and meaning, and one's contribution in society, all deemed essential components of psychological well-being in philosophical traditions and contemporary theories (Diener & Emmons, 1984; Ryan & Deci, 2000; Keyes, 2002; 2005). As such, the removal of the aforesaid items in Salama-Younes' (2011) appeared to be based largely on statistical criteria (i.e., improvement of fit statistics), and lacking a qualitative justification.

We found that through testing a bi-factor model, that neither scale purification nor correlating error terms are required. Specifically, we found excellent CFA fit statistics for the retained model comprising a ‘general’ positive mental health factor, and three specific factors of ‘Hedonic well-being’, ‘Social well-being’ and ‘Psychological well-being’. As such, our findings amongst the athlete population support a number of factor analysis studies on the MHC-SF (De Bruin and Du Plessis 2015; Hides et al. 2016; Jovanovic´ 2015), including a recent multi-national study of 7,521 participants (Longo, Jovanović, Sampaio de Carvalho, & Karaś, 2020). Notably, further bi-factor specific calculations showed most of the I-ECV item values were <.80 rather than >.80 (see Table 3). Independent of the association between the external variables and composite MHC-SF score, significant associations remained with specific factors and external variables. Hence, we suggest a multi-dimensional structure provides researchers and practitioners to isolate specific mental health components alongside the unitary score.

However, we urge that specific factor scores should strictly be used to supplement the unitary score, as most of the MHC-SF data converged on an overarching strong general mental health factor. Specifically, and consistent with recent studies (Hides et al. 2016; Longo, Jovanović, Sampaio de Carvalho, & Karaś, 2020), relative to the specific factors, the general factor exhibited high reliability, and explained the majority of model and item variance. Moreover, the correlations between specific factors and external variables were weak when the unitary score was statistically controlled for. Such findings support structural equation modelling data (Hides et al. 2016) that the predictive validity of bi-factor version of MHC-SF’s is attributable to its general factor.

Lastly, to explain the somewhat contradictory finding that the higher-order unidimensional model exhibited poor model fit, whereas the bi-factor model's strength was attributable to the general, overarching positive mental health construct, Reise, Cook and Moore (2015) have suggested that global constructs such as mental health, intelligence and personality will inevitably exhibit multidimensionality. Hence, positive mental health can be considered a single construct pertaining to a global evaluation about one's subjective well-being, existing alongside multiple concepts, such as emotional, psychological, and social well-being (Longo, Jovanović, Sampaio de Carvalho, & Karaś, 2020).

Practical and methodological recommendations

Some practical recommendations from the study include the use of Keyes' (2002) two continua model of mental health when considering the design and evaluation of mental health literacy and awareness programmes. The two continua model provides a narrative around mental health that is less stigmatising, and less medicalised than what has previously been used (Hughes and Leavey, 2012). For example, Uphill, Sly & Swain (2016) outline that the use of the two continua model to mental health can offer athletes a narrative regarding how a successful, high functioning, athlete can simultaneously experience well-being and have a mental illness. Such examples are evident in recent studies among adolescent athletes (see Wynters et al., 2021). Moreover, willing athletes who self-characterise themselves as being well, yet experiencing or currently working through a mental illness could act as role models to help destigmatise mental illness. Additionally, the specific factors found within the MHC-SF in the present study could help practitioners explore the importance of social relationships, psychological meaning and purpose, and emotional health to one's overall mental health (Giles et al., 2020).

In a research capacity, the MHC-SF could be integrated into monitoring and evaluation of programme effectiveness, wherein the general score is calculated and supplemented by specific factors to determine any self-reported change. The MHC-SF is relatively quick to complete, easy to understand, and the proposed calculations are primary functions within statistical software packages such as SPSS (Longo et al., 2020). When examining more complex structural equation modelling, further epidemiological, cross-sectional studies could model the bi-factor version of the MHC-SF using the figure schematics presented in this study and specify predictions with relevant mental health variables (e.g., psychological needs satisfaction, drug misuse, trauma history). Doing so will help advance knowledge of athlete mental health such that athlete experiences and self-reports are grounded in Keyes' (2002) theory, helping ensure precision and an accurate representation of the correlates of interest (Giles et al., 2020). With the advances in sport psychiatry, the MHC-SF could also be used alongside ill-being measures in a more holistic assessment of athletes who present with psychological issues (Mistry, McCabe & Currie, 2020).

Limitations

There were several limitations, namely, the cross-sectional design meant that test-retest reliability remains unassessed. The mean age was 32, and any extrapolation to younger age groups involved in competitive sport is restricted. Although individual and coactive sports were represented in the sample, the types of sports was limited to equestrian, rugby, Gaelic-games and running, and the inclusion of more sports would have been more representative. The vast majority (86.4%) of participants classified themselves as non-or sub-elite (e.g., amateur, local leagues), and the 13.6% of participants who self-classified as elite is comparatively higher than the National Collegiate Athletics Association's (2019) estimate of 6%. While definitions of elite athlete level vary according to the sport, standard of participation, and global context (Swann et al., 2015), psychological pressure to succeed is higher at the elite level. Given a recommended participant-to-parameter ratio of at least 10:1 in structural equation modelling (Jackson, 2003) a larger sample of athletes could have warranted a splitting of the sample into subgroups (see Longo, Jovanović, Sampaio de Carvalho, & Karaś, 2020, for a multinational analysis), to determine if the bi-factor model of the MHC-SF holds true in the various competitive athlete levels and demographic factors. To this end, future research should aim to achieve a representative demographic sample when evaluating mental health measures, screening tools, and diagnostic practices. The present sample was mainly female, which may reflect a higher likelihood of females to complete mental health surveys or engage in the topic of mental health. Further, important data on race/ethnicity, socio-economic status and education/employment were absent. Finally, mental health literacy levels (i.e., knowledge of mental health, mental illness and self-management) differs across countries (McGivern et al., 2021), this may explain athlete participation rates and openness to engage in the survey.

Conclusion

Overall, the bi-factor version of MHC-SF represents a theoretically grounded and valid measure of positive mental health in athletes. Sport organisations, researchers and practitioners may consider integration of the MHC-SF into monitoring and evaluation of programme effectiveness, and/or screening of positive mental health. We propose that future use of the MHC-SF should entail the calculation of a composite score in the knowledge that it is explaining the vast majority of MHC-SF model and item variance. However, given some contribution of the specific factors, supplementary analysis may involve the calculation of specific factors - albeit strictly to supplement the composite score. Keyes' (2002) two continua model and the specific factors found within the associated MHC-SF in the present study could serve as discussion points in future athlete mental health interventions, and continue to ground athletes' experiences in theory in future research studies. Limitations of this cross-sectional study relate to the higher distribution of female-to-male participants, higher age category of athletes, and the test-retest reliability of the MHC-SF remains unassessed in athlete populations. Finally, conducting further longitudinal factor analysis studies with a broader range of sports can provide a more comprehensive psychometric instrument for athletes.

References

Breslin, G., Shannon, S., Ferguson, K., Devlin, S., Haughey, T., & Prentice, G. (2019). Predicting athlete mental health stigma using the theory of reasoned action framework. *Journal of Clinical Sport Psychology*, *13*(1), 103-115.

Breslin, G., Shannon, S., Haughey, T., Donnelly, P., & Leavey, G. (2017). A systematic review of interventions to increase awareness of mental health and well-being in athletes, coaches and officials. *Systematic reviews*, *6*(1), 1-15.

Breslin, G., & Leavey, G. (Eds.). (2019). *Mental health and well-being interventions in sport: Research, theory and practice*. Routledge, London.

Byrne, B. M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. *International journal of testing*, *1*(1), 55-86.

Carpenter, N.C., Son, J., Harris, T.B., Alexander, A.L. and Horner, M.T. (2016), "Don't forget the items: item-level meta-analytic and substantive validity techniques for reexamining scale validation", *Organizational Research Methods*, Vol. 9 No. 4, pp. 616-650.

Donker, T., van Straten, A., Marks, I., & Cuijpers, P. (2011). Quick and easy self-rating of Generalized Anxiety Disorder: validity of the Dutch web-based GAD-7, GAD-2 and GAD-SI. *Psychiatry research*, *188*(1), 58-64.

de Holanda Coelho, G. L., Hanel, P. H., Medeiros Cavalcanti, T., Teixeira Rezende, A., & Veloso Gouveia, V. (2016). Brief Resilience Scale: Testing its factorial structure and invariance in Brazil. *Universitas Psychologica*, *15*(2), 397-408.

De Bruin, G. P., & Du Plessis, G. A. (2015). Bifactor analysis of the Mental Health Continuum-Short Form (MHC-SF). Psychological Reports, 116(2), 438–446.

Diener, E., & Emmons, R. A. (1984). The independence of positive and negative affect. *Journal of personality and social psychology*, *47*(5), 1105.

Dueber, D. M. (2017). Bifactor Indices Calculator: A Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. (accessed 04/03/2021: https://uknowledge.uky.edu/edp_tools/1/)

Ferentinos, P., Yotsidi, V., Porichi, E., Douzenis, A., Papageorgiou, C., & Stalikas, A. (2019). Well-being in patients with affective disorders compared to nonclinical participants: A multi-model evaluation of the Mental Health Continuum-Short Form. *Journal of clinical psychology*, *75*(9), 1585-1612.

Foster, B. J., & Chow, G. M. (2019). Development of the Sport Mental Health Continuum—Short Form (Sport MHC-SF). *Journal of Clinical Sport Psychology*, *13*(4), 593-608.

Giles, S., Fletcher, D., Arnold, R., Ashfield, A., & Harrison, J. (2020). Measuring well-being in sport performers: where are we now and how do we progress?. *Sports Medicine*, *50*(7), 1255-1270.

Gouttebarge, V., Bindra, A., Blauwet, C., Campriani, N., Currie, A., Engebretsen, L., & Budgett, R. (2021). International Olympic Committee (IOC) Sport Mental Health Assessment Tool 1 (SMHAT-1) and Sport Mental Health Recognition Tool 1 (SMHRT-1): towards better support of athletes' mental health. British journal of sports medicine, 55(1), 30-37.

Henriksen, K., Schinke, R., Moesch, K., McCann, S., Parham, W. D., Larsen, C. H., & Terry, P. (2020). Consensus statement on improving the mental health of high performance athletes. *International Journal of Sport and Exercise Psychology*, 18(5), 553-560.

Hides, L., Quinn, C., Stoyanov, S., Cockshaw, W., Mitchell, T., & Kavanagh, D. J. (2016). Is the mental wellbeing of young Australians best represented by a single, multidimensional or bifactor model? *Psychiatry Research*, 241, 1–7.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Hughes, L., and Leavey, G. (2012). Setting the bar: athletes and vulnerability to mental illness. *Br. J. Psychiatry* 200, 95–96. doi: 10.1192/bjp.bp.111.095976

Huppert, F. A. (2009). Psychological well-being: Evidence regarding its causes and consequences. *Applied Psychology: Health and Well-Being, 1*(2), 137-164.

Henriksen, K., Schinke, R., Moesch, K., McCann, S., Parham, W. D., Larsen, C. H., & Terry, P. (2020). Consensus statement on improving the mental health of high performance athletes. *International Journal of Sport and Exercise Psychology*, *18*(5), 553-560.

Keyes, C. L. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of health and social behavior*, 207-222.

Keyes, C. L. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. Journal of Consulting and Clinical Psychology, 73(3), 539.

Keyes, C. L., Wissing, M., Potgieter, J. P., Temane, M., Kruger, A., & Van Rooy, S. (2008). Evaluation of the mental health continuum–short form (MHC–SF) in setswana-speaking South Africans. *Clinical Psychology & Psychotherapy*, *15*(3), 181-192.

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of affective disorders*, *114*(1-3), 163-173.

Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N: q hypothesis. *Structural equation modeling*, *10*(1), 128-141.

Jackson, D. L., Gillaspy Jr, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods, 14*(1), 6.

Jovanović, V. (2015). Structural validity of the Mental Health Continuum-Short Form: The bifactor model of emotional, social and psychological well-being. *Personality and Individual Differences*, *75*, 154-159.

Jones, P. B. (2013). Adult mental health disorders and their age at onset. *The British Journal of Psychiatry*, *202*(s54), s5-s10.

Joshanloo, M., & Jovanović, V. (2017). The factor structure of the mental health continuum-short form (MHC-SF) in Serbia: An evaluation using exploratory structural equation modeling. *Journal of Mental Health*, *26*(6), 510-515.

Lamers, S. M., Westerhof, G. J., Bohlmeijer, E. T., ten Klooster, P. M., & Keyes, C. L. (2011). Evaluating the psychometric properties of the mental health continuum-short form (MHC-SF). Journal of clinical psychology, 67(1), 99-110.

Lonsdale, C., Hodge, K., & Rose, E. A. (2006). Pixels vs. Paper: Comparing Online and Traditional Survey Methods in Sport Psychology. *Journal of Sport & Exercise Psychology*, 28(1), 100-108

Longo, Y., Jovanović, V., Sampaio de Carvalho, J., & Karaś, D. (2020). The general factor of well-being: Multinational evidence using bifactor ESEM on the Mental Health Continuum–Short Form. *Assessment*, *27*(3), 596-606.

Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical care*, 266-274.

Lundqvist, C., & Sandin, F. (2014). Well-being in elite sport: Dimensions of hedonic and eudaimonic well-being among elite orienteers. *The sport psychologist*, *28*(3), 245-254.

McGivern, A., Shannon, S., & Breslin, G. (2021). Resilience, well-being, depression symptoms and concussion levels in equestrian athletes. *Journal of Public Mental Health* (in press).

Mistry, A. D., McCabe, T., & Currie, A. (Eds.). (2020). *Case studies in sports psychiatry*. Cambridge University Press.

Razykov I, Ziegelstein RC, Whooley MA, Thombs BD. The PHQ-9 versus the PHQ-8--is item 9 useful for assessing suicide risk in coronary artery disease patients? Data from the Heart and Soul Study. J Psychosom Res. 2012; 73(3):163-168.

Rejeski, W. J., Brawley, L. R., & Shumaker, S. A. (1996). Physical activity and health-related quality of life. Exercise and sport sciences reviews, 24(1), 71-108.

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. Journal of Personality Assessment, 95(2), 129-140.

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling a bifactor perspective. Educational and Psychological Measurement, 73(1), 5-26.

Rodriguez A, Reise SP, Haviland MG (2016). Applying bifactor statistical indices in the evaluation of psychological measures. Journal of Personality Assessment 98, 223–237. doi: 10.1080/00223891.2015.1089249

Salama-Younes, M. (2011). Validation of the mental health continuum short form and subjective vitality scale with Egyptian adolescent athletes. In *The human pursuit of well-being* (pp. 221-234). Springer, Dordrecht

Shannon, S., Breslin, G., Haughey, T., Sarju, N., Neill, D., Lawlor, M., & Leavey, G. (2019). Predicting student-athlete and non-athletes' intentions to self-manage mental health: Testing an integrated behaviour change model. *Mental Health & Prevention*, *13*, 92-99.

Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The brief resilience scale: assessing the ability to bounce back. *International journal of behavioral medicine*, *15*(3), 194-200.

Stucky, B. D. & Edelen, M. O. (2015). Using heierarchical IRT models to create unidimensional measures from multidimensional data. In S. P. Reise & D. A. Revicki (Eds.), Handbook of item response theory modeling: Applications to typical performance assessment, 183-206. New York: Routledge.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*, *166*(10), 1092-1097.

Swann, C., Moran, A., & Piggott, D. (2015). Defining elite athletes: Issues in the study of expert performance in sport psychology. *Psychology of Sport and Exercise*, *16*, 3-14.

Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development*, *58*(1), 134–146. https://doi.org/10.2307/1130296

Vella, S. A., Benson, A., Sutcliffe, J., McLaren, C., Swann, C., Schweickle, M. J., .& Bruner, M. (2020). Self-determined motivation, social identification and the mental health of adolescent male team sport participants. *Journal of Applied Sport Psychology*, 1-15.

United Nations Department of Economic and Social Affairs Population Division. (2015). World population, ageing. In World population ageing. New York: United Nations.

Uphill, M., Sly, D., & Swain, J. (2016). From mental health to mental wealth in athletes: Looking back and moving forward. *Frontiers in psychology*, 7, 935.

Wieland, A., Durach, C.F., Kembro, J. and Treiblmaier, H. (2017), "Statistical and judgmental criteria for scale purification", *Supply Chain Management*, Vol. 22 No. 4, pp. 321-328. https://doi.org/10.1108/SCM-07-2016-0230

Wynters, R., Liddle, S. K., Swann, C., Schweickle, M. J., & Vella, S. A. (2021). Qualitative evaluation of a sports-based mental health literacy program for adolescent males. *Psychology of Sport and Exercise*, *56*, 101989.

Wu, Y., Levis, B., Riehm, K. E., Saadat, N., Levis, A. W., Azar, M., & Ioannidis, J. P. (2019). Equivalency of the diagnostic accuracy of the PHQ-8 and PHQ-9: a systematic review and individual participant data meta-analysis. *Psychological medicine*, 1-13.