Article

EPER

European Physical Education Review
2020, Vol. 26(4) 782–796

DOI: 10.1177/1356336X19882060
journals.sagepub.com/home/epe
SAGE

What is health-related fitness? Investigating the underlying factor structure of fitness in youth

Úna Britton
School of Health & Human Performance, Dublin City University, Ireland

Johann Issartel
School of Health & Human Performance, Dublin City University, Ireland

Gerry Fahey
Dublin City University Business School, Ireland

Gerry Conyngham
Dublin City University Business School, Ireland

Sarahjane Belton
School of Health & Human Performance, Dublin City University, Ireland

Abstract

Health-related fitness (HRF) is theoretically defined as a multidimensional construct containing the components cardiorespiratory endurance, muscular strength, muscular endurance, flexibility and body composition. Given the range of field-based HRF tests available, health practitioners face a difficult task in selecting tests that best reflect the HRF construct as defined in the literature. This study aimed to investigate the underlying factor structure of the theoretical HRF construct with a view to identifying field-based tests representative of HRF in youth. Participants were 261 children (53% female, 47% male; mean age 12.22 ± 0.48 years). Indicators of four fitness components (20m shuttle run, curl-ups, push-ups, horizontal jump, vertical jump, handgrip, and modified back-saver sit-and-reach) were measured. Confirmatory factor analysis of the four-component model revealed a low contribution of flexibility (.10) to the overall model, leading to its removal. The subsequent three-component model showed better fit across all fit statistics (NFI, TLI, CFI, RMSEA, SRMR). Analysis of indicator loadings led to the removal of handgrip (.37), further improving model fit. The reduced three-component model was re-

Corresponding author:
Úna Britton, School of Health & Human Performance, Dublin City University, Dublin 9, Ireland.
Email: una.britton2@mail.dcu.ie

specified as a first-order model containing five indicators, and showed the best fit (NFI, CFI, SRMR). These findings suggest that a fitness construct in youth is adequately represented by three HRF components included in the theoretical definition (cardiorespiratory endurance, muscular strength, muscular endurance). These components load onto the same construct and can be tested using five tests (20m shuttle run, curl-ups, push-ups, horizontal jump, vertical jump) suitable for field-based research.

Keywords
Health-related fitness, youth, measurement, confirmatory factor analysis

Introduction

Physical fitness is a powerful indicator of health among youth (Ortega et al., 2008a; Ruiz et al., 2009). In children and adolescents higher fitness levels are associated with positive outcomes for bone health, mental health, obesity, and cardiovascular disease (Ortega et al., 2008; Ruiz et al., 2009). It is important to note, however, that physical fitness is a hypothetical construct (Marsh, 1993). In the past, studies assessing physical fitness have included a range of variables, from blood pressure, lung function, and body girth, to cardiorespiratory endurance (CRE), strength, coordination, and flexibility (Fleishman, 1964; Marsh, 1993). A definition of health-related fitness (HRF) was proposed by Caspersen and colleagues (1985), who suggested that HRF is a multidimensional construct comprising five components (Caspersen et al., 1985). Specifically, these five components are CRE, muscular strength (MS), muscular endurance (ME), flexibility, and body composition. CRE refers to the capacity of the respiratory and cardiovascular systems to carry out continuous strenuous exercise (Ortega et al, 2008). MS is the ability of the muscular system to produce force against a resistance in one maximal effort (Smith et al., 2014). Many researchers in the health field choose tests of explosive strength (e.g. horizontal and vertical jumps), or power, to represent MS as a component of fitness (Ortega et al., 2011; Tomkinson et al., 2018). ME is the ability of the muscular system to produce force over a prolonged period (Ortega et al., 2008; Smith et al., 2014). ME and MS together contribute to muscular fitness (MF), a term that refers to the ability to do work against a resistance either maximally, explosively, or repeatedly (Ortega et al., 2008; Smith et al., 2014). Flexibility is the range of motion at a joint (Pate et al., 2012). Body composition is the physical make-up of the body, often described as the percentage of muscle, fat, bone and water within the body (Caspersen et al., 1985). This definition of HRF, as a multidimensional construct consisting of five components, is predominantly accepted and utilised within health promotion research (ACSM, 2014; Payne and Isaacs, 2016). According to Pate (1988) these five components were included in the HRF construct because they are affected by physical training and are associated with important health outcomes.

Recent research provides evidence for associations between health and CRE, MS and ME (Ortega et al., 2008; Smith et al., 2014), but there is limited evidence to suggest an association between health and flexibility (Pate et al., 2012). The relationship between CRE and health is well established (Corbin et al., 2014; Pate et al., 2012; Ortega et al., 2008). In their review, Ortega and colleagues (2008) found strong evidence for associations between CRE and adiposity, cardiovascular disease risk, quality of life, and mental health. Associations between MF and health have been reported, but are less well established (Corbin et al., 2014; Ortega et al., 2008). Despite this there is growing evidence for the benefits of MF for health among youth. Specifically, superior MS and ME have been favourably associated with a reduced risk of insulin insensitivity (Benson et al., 2006), a reduced likelihood of excess adiposity (Grøntved et al., 2015; Janz et al., 2002; Smith et al., 2014), and a reduced risk of cardiovascular disease (Grøntved et al., 2015; Janz et al., 2002). Longitudinal studies have also identified the importance of developing MF at a young age for maintaining good health later in life (Grøntved et al., 2015; Janz et al., 2002; Ruiz et al., 2009). There is limited research on the relationship between flexibility and health in youth (Pate et al., 2012). Studies that have measured flexibility and health outcomes have not found significant associations (Casonatto et al., 2016; Stodden et al., 2015). It is worth questioning, then, the inclusion of flexibility as a component of HRF, given the lack of research evidence supporting its link to important health outcomes.

Another problem arising from a five-component HRF construct occurs when body composition is positioned as an outcome of relationships between HRF and other variables. This is common in health promotion research (Bailey et al., 2012; Casonatto et al., 2016; Grøntved et al., 2015; Janz et al., 2002; Lima et al., 2017; Rodrigues et al., 2016; Smith et al., 2014). A current conceptual model developed by Stodden et al. (2008) highlights physical activity, motor competence, HRF, and perceived competence as central to health in youth. This model proposes HRF as a mediator in the reciprocal relationship between motor competence and physical activity, while 'weight status' is presented as the outcome of these relationships. From a methodological perspective, indicators (e.g. body composition) cannot be both independent and dependent variables. Therefore, where measures of body composition are outcomes of the research question, HRF can only be treated as a four-component construct (should flexibility remain as a component of the construct). In addition, assessment of body composition involves passive tests such as measuring height, weight and body circumferences, compared with the four remaining components of HRF which are measured using active tests that require the participant to perform some action or skill (e.g. running to measure CRE; jumping to measure MS). This further highlights the difference in nature between body composition and the other components of HRF included in the traditional definition (Caspersen et al., 1985), lending support to the argument for removing body composition as a component, thus reducing HRF to a four-component construct.

Another issue of note is that within the field of health promotion research, all components of HRF are not always measured when associations between HRF and other variables are being examined. Often tests from popular fitness test batteries are chosen to measure a specific component of HRF. If HRF is a multidimensional five-component construct, then important information may be lost by measuring just one component and using this as a representation of the overall construct. For example, in recent studies testing Stodden et al.'s (2008) conceptual model, just one component of HRF (CRE) was measured (Lima et al., 2017). CRE is evidently an important HRF component (Ortega et al., 2008; Ruiz et al., 2009), but considering the reported contribution of both MS and ME to health in youth (Benson et al., 2006; Grøntved et al., 2015; Janz et al., 2002; Ortega et al., 2008; Smith et al., 2014) the exclusion of these components may be problematic.

Numerous valid and reliable HRF tests are available for use in field-based settings (Artero et al., 2011; Morrow et al., 2010; Ortega et al., 2008). In addition to a requirement for valid and reliable measures, time and equipment/facilities are extremely important considerations for researchers and health practitioners. As such, some of the available fitness tests are more applicable than others in certain settings. For example, tests of ME such as curl-ups and push-ups are easy to administer and require little equipment (Plowman and Mahar, 2013). In comparison, tests such as pull-ups, or flexed arm hang, require gym equipment that may not be available in some field settings (Council of Europe, 1983; Plowman and Mahar, 2013). Given the wide range of options available to field-based researchers, narrowing down this selection to tests that truly reflect the overall HRF construct can be problematic.

Despite numerous studies assessing components of HRF, to the authors' knowledge no study has assessed the factor structure of this theoretical construct among youth. Given that various individual components of HRF are often used as indicators of the overall construct, it is important to understand the underlying structure of HRF so that discussions of HRF are talking about a general and comparable concept. In addition, although sex differences in absolute scores for HRF are widely reported, with males consistently scoring higher than females across all HRF components except flexibility (Ortega et al., 2011; Santos et al., 2014), it is generally accepted that the overall HRF construct is the same for males and females. To the authors' knowledge, no study has tested this assumed sex invariance for HRF.

This research aims to test, using factor analysis, a four-component HRF construct based primarily on the theoretical definition of HRF (Caspersen et al., 1985) but excluding the component of body composition (due to the previously outlined problems posed by its inclusion as an independent variable), to determine the underlying structure of HRF within a youth population and to test its invariance across sex. In addition, this research aims to provide support for the selection of specific field-based tests which can adequately represent the HRF construct within a youth population.

Methods

Participants and settings

Data for this cross-sectional study were collected as part of the first phase of a longitudinal study which aims to track changes in health-related variables in children as they transition from primary to second-level school. Principals from six mixed-gender primary schools were contacted via email to invite their 6th class students to participate in this study. In Ireland, 6th class is the final year of primary education. These schools were specifically chosen as they were identified as the main feeder primary schools for two second-level schools which were invited to, and confirmed participation in, phase two of the longitudinal aspect of this study. Informed consent was granted by all six primary school principals, and individual consent was obtained from parents/guardians. Ethical approval was obtained from the authors' institutional ethics committee. Prior to testing, a physical activity readiness questionnaire was also completed for each participant by their parent/guardian. The consenting sample consisted of 261 children (53% female and 47% male; mean age 12.22 ± 0.48 years).

Measures

Seven tests were used to measure four components (MS, ME, CRE and flexibility) of HRF (Table 1). Tests were selected following a review of methodologies in recent large-scale studies of HRF in youth (Ortega et al., 2011; Ruiz et al., 2011; Santos et al., 2014), with specific consideration paid to the applicability of tests within a school-based setting, test reliability, and the ability of the tests to measure each of the four specified components of the proposed HRF construct. Protocols for each HRF test can be found in the FITNESSGRAM manual (Plowman and Mahar,

Table 1. Measurement of HRF.

HRF component	Test	Source	Reliability
CRE	20MST	FITNESSGRAM; EUROFIT	0.78–0.99 (Artero et al., 2011)
MS	Handgrip	EUROFIT; HELENA Study	0.96–0.98 (Artero et al., 2011)
	HJ	EUROFIT; HELENA Study	$p > 0.05$ test–retest differences (Ortega et al. 2008)
	VJ*	HELENA Study	$p > 0.05$ test–retest differences (Ortega et al., 2008)
ME	Push-ups	FITNESSGRAM	0.77 (Morrow et al., 2010)
	Curl-ups	FITNESSGRAM	0.87 (Morrow et al., 2010)
Flex	BS S&R	FITNESSGRAM	0.89 (Morrow et al., 2010)

20MST: 20 metre shuttle run test; CRE: cardiorespiratory endurance; MS: muscular strength; ME: muscular endurance; Flex: flexibility; HJ: horizontal jump; VJ: vertical jump; BS S&R: back-saver sit-and-reach; EUROFIT (Council of Europe 1983). FITNESSGRAM (Meredith and Welk, 2010). The HELENA Study (Ortega et al., 2008). *VJ was assessed using the Abalakov jump test protocol outlined in the HELENA study (Ortega et al., 2008) and using a jump mat and belt (Coulson and Archer, 2009) in place of an infrared jump platform.

2013), the EUROFIT manual (Council of Europe, 1983), and the HELENA study (Ortega et al., 2008; Ortega et al., 2011) (Table 1).

A minimum of four trained field staff conducted the testing sessions. Each class group was tested separately, with approximately 25 students in each class group. Testing was performed over the course of three separate physical education classes. 20MST and handgrip were tested on the first day, curl-ups on the second day, and horizontal jump (HJ), vertical jump (VJ), push-ups, and back-saver sit-and-reach (BS S&R) on the third day. On each day participants were put into groups of 5–6 students and each group rotated through the tests in a circuit fashion. One trained field staff member provided every 5–6 participants with a demonstration and verbal instructions of the test, the exception being the 20MST where groups of 10–12 participants were given a demonstration and instructions prior to beginning the test. Trained field staff followed the standardised testing protocols (Council of Europe, 1983; Ortega et al., 2011; Plowman and Mahar, 2013) when administering all tests. For the 20MST verbal encouragement from field staff was given to motivate participants to reach maximum effort. Speed reached on the last completed shuttle in the 20MST was used to calculate VO_{2max} (Léger et al., 1988). Maximum jump height (VJ) and distance (HJ) were calculated as the highest score from three trials on each test. Handgrip was calculated as the mean of the maximum score achieved on the right and left hand, with three trials performed on each hand. The maximum number of curl-ups and push-ups achieved were recorded. The furthest distance reached in the BS S&R was recorded. No performance feedback was given to participants during testing, and participants were not informed of their scores on any of the tests.

Data analysis

Data were analysed using SPSS version 23 and AMOS version 23 for Windows. Raw scores for each HRF test were entered into SPSS and descriptive statistics for each HRF test were calculated. A one-way between-groups multivariate analysis of variance (MANOVA) was conducted to

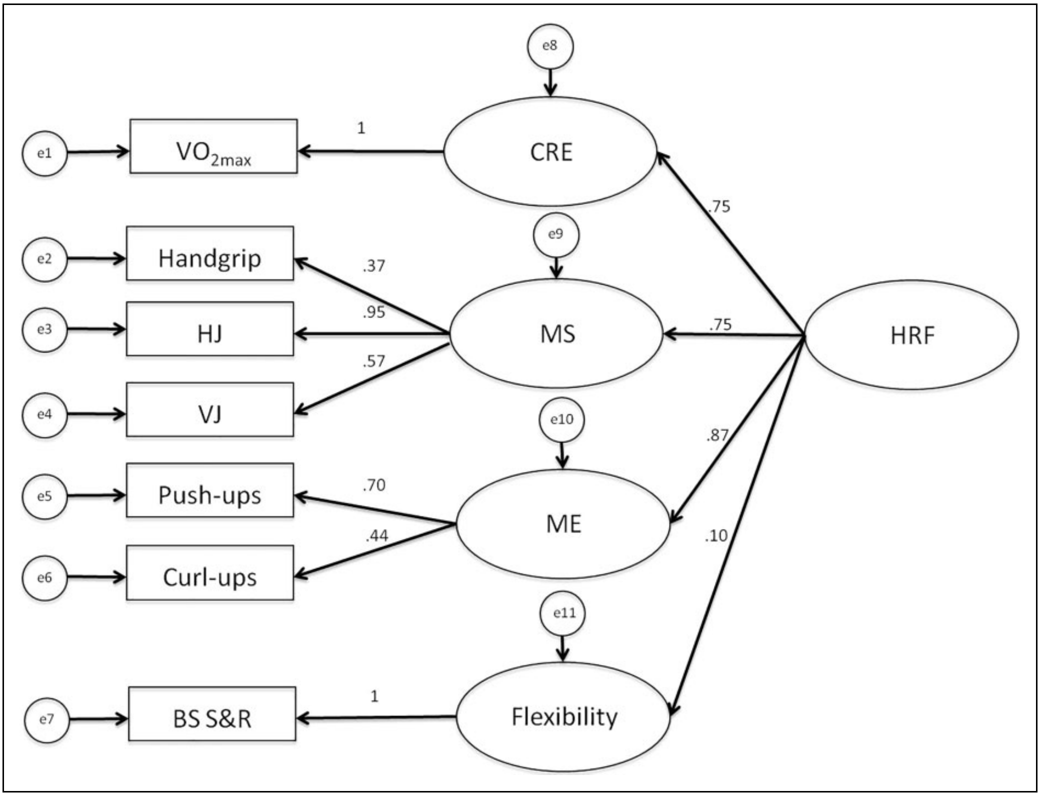

Figure 1. Four-component HRF model. HRF: health-related fitness; CRE: cardiorespiratory endurance; MS: muscular strength; ME: muscular endurance; HJ: horizontal jump; VJ: vertical jump; BS S&R: back-saver sit-and-reach.

identify sex differences for HRF tests. A Bonferroni adjusted alpha level of <0.017 was used to calculate significance. Effect sizes were calculated to identify the magnitude of any differences.

To analyse the underlying structure of HRF, a confirmatory factor analysis (CFA) using maximum likelihood estimation methods was conducted in AMOS version 23 (Arbuckle, 2014). CFA is widely used to test whether a hypothesised model fits the data (Fox, 2010). The initial model tested a four-component HRF construct (Model 1; Figure 1) based primarily on the theoretical definitions of HRF (ACSM, 2014; Caspersen et al., 1985; Pate, 1988; Payne and Isaacs, 2016) while taking into account the argument for excluding body composition as the fifth component. Within CFA, the presence of single-indicator latent variables (such as CRE in Figure 1) can cause identification issues (Kelloway, 1998; Kline, 2015). To address this identification issue in the current study, and enable the testing of the theoretically defined HRF construct, the error variance for CRE was set to zero and the loading to one, as recommended in Lämmle et al. (2010). To account for missing data, the nature of missing data was tested using Little's 'missing completely at random' (MCAR) test ($\chi^2 = 2016.83$; $p = 0.971$). A statistically non-significant result for Little's MCAR test allows for MCAR to be inferred (Tabachnick and Fidell, 2007). Where data are MCAR and there is a low percentage of data missing, most procedures to deal with this issue can be used successfully (Tabachnick

Table 2. Mean HRF scores according to sex.

						Sex difference	
	All	Males	Females			Absolute	Fitnessgram
	(n=261)	(n=128)	(n=133)	F	$\eta 2$	Values	Norms
VO_{2max} (ml.kg.min)^F	49.66 (5.03)	51.11 (5.39)	48.21 (4.19)	20.2	0.08	m *	ns
Handgrip (kg)	17.73 (4.11)	18.10 (3.90)	17.34 (4.31)	1.7		ns	
HJ (cm)	37.21 (23.28)	144.93 (22.65)	129.85 (21.49)	23.2	0.10	m *	
VJ (cm)	35.49 (6.87)	36.74 (7.31)	34.30 (6.22)	6.2		ns	
Push-upsF	8.68 (8.01)	11.27 (8.74)	6.09 (6.25)	23.9	0.10	m *	ns
Curl-upsF	13.54 (10.77)	15.26 (11.88)	11.80 (9.23)	6.9		ns	m *
BS S&R (cm)^F	9.93 (7.64)	7.07 (7.13)	12.66 (7.11)	29.2	0.12	f *	ns

Data are shown as means with standard deviation in brackets. *Significant at Bonferroni adjusted p-value < 0.017.
η^2: partial eta squared effect size 'm' or 'f' in the sex difference column denotes the sex (m=male, f=female) that performed significantly better.
F: included in Fitnessgram; HJ: horizontal jump; VJ: vertical jump; BS S&R: back-saver sit-and-reach.

and Fidell, 2007). In this data set all measured variables had low levels of missing data (< 7%). Group mean substitution was therefore selected to estimate missing values (Tabachnick and Fidell, 2007).

To begin with, χ^2 was used to assess goodness of fit between the observed and fitted covariance matrices. A statistically insignificant χ^2 indicates a good fitting model, even though a poor χ^2 does not necessarily imply a poor model fit, since χ^2 can give erroneous results with large samples (Byrne, 2010). In addition, model fit was further assessed based on the Normed-Fit Index (NFI), the Comparative Fit Index (CFI), the Tucker–Lewis Index (TLI), the Root Mean Square Error of Approximation (RMSEA), and the Standardised Root Mean Square Residual (SRMR). NFI, TLI and CFI values of >0.9 signify acceptable model fit (Byrne, 2010; Schumacker and Lomax, 1996), with values of >0.95 for TLI and CFI considered to signify superior fit (Hu and Bentler, 1999). For RMSEA and SRMR, values of <0.06 and <0.08, respectively, indicate a good fitting model (Hu and Bentler, 1999). Based on analysis of the factor loadings for the four-component HRF construct (Model 1; Figure 1), additional models were subsequently tested, and a multiple-group analysis was used to test invariance in the HRF construct across sex.

Results

Performances on each HRF test are shown in Table 2. One-way MANOVA revealed significant sex differences for VO_{2max}, HJ, push-ups, and flexibility (Table 2). The importance of the differences can be seen by examining the effect sizes in Table 2. According to Cohen (1992) effect sizes of .20 to .49 are small, .50 to .79 are medium, and $\geq$ .80 are large. Males had significantly higher scores for VO_{2max}, HJ and push-ups. Females had significantly higher scores for flexibility. VO_{2max}, push-ups, curl-ups and BS S&R tests were all tested in accordance with FITNESSGRAM (Plowman and Mahar, 2013) protocols. When these HRF test-scores were assigned age and sex-specific values in accordance with FITNESSGRAM norms (Ortega et al., 2011; Plowman and Mahar, 2013), one-way MANOVA showed no significant sex differences for standardised scores.

Table 3. CFA fit statistics for HRF Models

Model	Description	χ^2	Df	Prob	NFI	TLI	CFI	RMSEA	SRMR
Model 1	Four-component model (CRE MS ME Flex)	31.91	12	0.001	0.91	0.90	0.94	0.080	0.057
Model 2	Three-component model (CRE MS ME)	17.74	7	0.013	0.95	0.93	0.97	0.077	0.045
Model 3	Reduced three-component model	10.80	3	0.013	0.97	0.91	0.97	0.100	0.037
Model 4	(CRE MS ME) First-order model	10.80	3	0.013	0.97	0.91	0.97	0.100	0.037

CFA: confirmatory factor analysis; $\chi 2$: Chi-square test; Df: degrees of freedom. Prob.: probability NFI: Normed Fit Index; TLI: Tucker–Lewis Index; CFI: Comparative Fit Index; RMSEA: Root mean square error of approximation; SRMR: Standardised Root Mean Square Residual; HRF: health-related fitness; CRE: cardiorespiratory endurance; MS: muscular strength; ME: muscular endurance; Flex: flexibility.

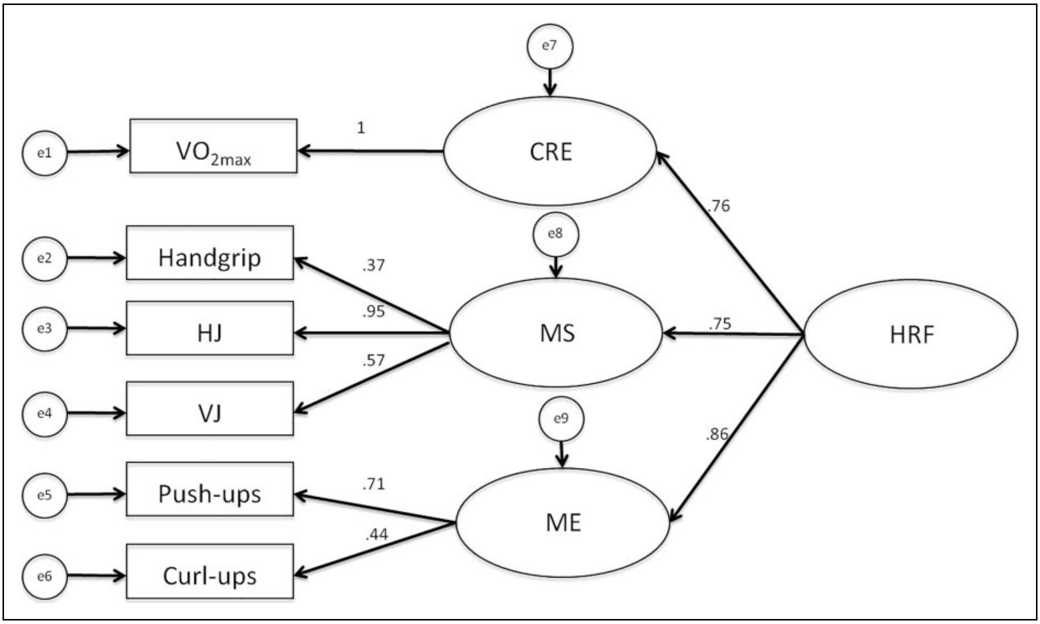

Figure 2. Three-component HRF model. HRF: health-related fitness; CRE: cardiorespiratory endurance; MS: muscular strength; ME: muscular endurance; HJ: horizontal jump; VJ: vertical jump.

Factor analysis results showed that all of the models tested fit the data well (Table 3), with NFI, TLI, and CFI all above 0.9 (Byrne, 2010; Schumacker and Lomax, 1996). In Model 1, a four-component second-order model, HRF was described primarily by three components, CRE, MS and ME, with factor loadings ranging from 0.75 to 0.87 (Figure 1). Given the low loading for flexibility (0.10), it was removed as a component and a three-component model (Model 2; Figure 2) was subsequently tested.

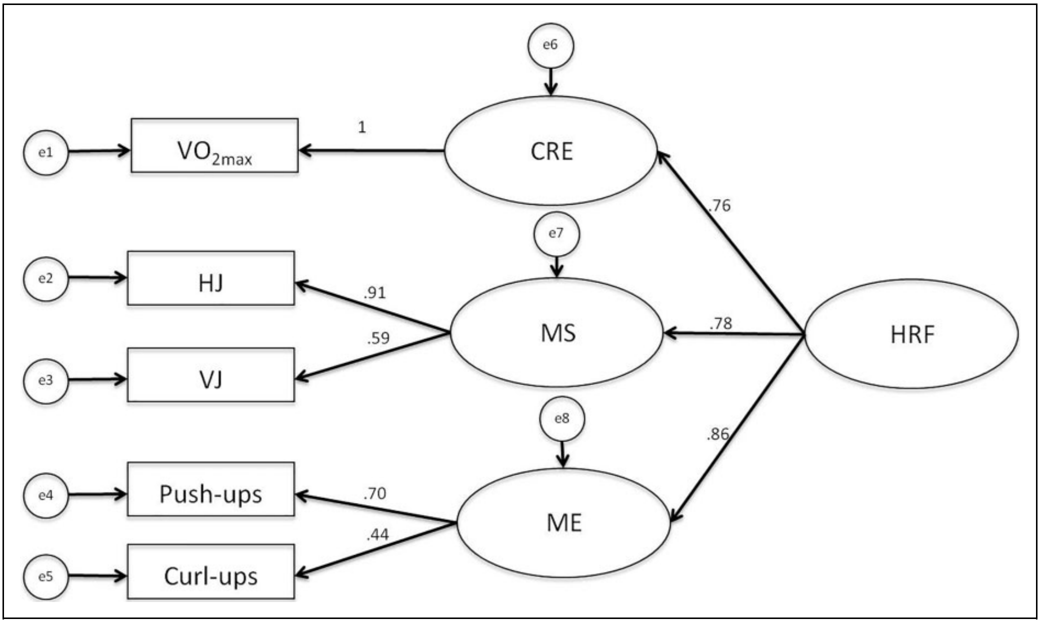

Figure 3. Reduced three-component HRF model. HRF: health-related fitness; CRE: cardiorespiratory endurance; MS: muscular strength; ME: muscular endurance; HJ: horizontal jump; VJ: vertical jump.

Factor loadings remained similar for CRE (0.76), MS (0.75) and ME (0.86) in the three-component (Figure 2) compared with the four-component model (Figure 1). On reviewing the loadings on the manifest variables for the three-component model (Figure 2), handgrip was removed from MS due to its low loading (0.37). Removing handgrip did not have a negative effect on fit. A number of the fit indices improved slightly (NFI, SRMR) and a number decreased slightly (TLI, RMSEA). CFI, TLI, NFI and SRMR all remained within acceptable limits for a good fitting model (Byrne, 2010; Hu and Bentler, 1999; Schumacker and Lomax, 1996). An examination of the confidence intervals for RMSEA showed the lower confidence limit met the rules of thumb for fit. The reduced three-component model (Figure 3) was therefore selected as it was a more parsimonious model.

The reduced three-component model was then re-specified as a first-order model given the identification issues posed by single indicator latent variables (Kelloway, 1998; Kline, 2015) and with a view to identifying the factor loadings for each individual test in the overall HRF construct (Figure 4). Error terms for HJ and VJ, and push-ups and curl-ups were correlated (Model 4; Figure 4) since HJ and VJ both measure MS, and push-ups and curl-ups both measure ME. The fit indices were identical for the second-order reduced three-component model (Figure 3), and the first-order model (Figure 4). This is expected, given that correlating the errors on the two MS and two ME indicators in the first-order model in effect connects these indicators in the same way as having them as indicators within a component, i.e. MS or ME.

Multiple-group analysis of the re-specified first-order HRF model (Figure 4) showed good fit (Table 3). This HRF model was invariant across sex (CMIN $= 5.547$; $p = 0.236$) for factor loadings. As expected, measurement intercepts were significantly different between males and

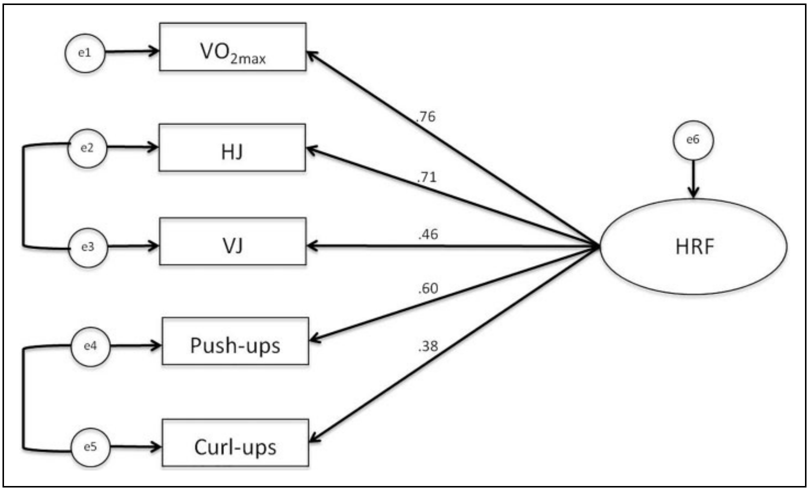

Figure 4. First-order HRF model. HRF: health-related fitness; HJ: horizontal jump; VJ: vertical jump.

Table 4. Factor loadings (β) for indicators of HRF by sex.

	Male	Female
VO_{2max}	0.68	0.80
HJ	0.61	0.76
Push-ups	0.61	0.47
VJ	0.40	0.45
Curl-ups	0.28	0.43

HRF: health-related fitness, HJ: horizontal jump, VJ: vertical jump.

females (CMIN $= 51.618$; $p = 0.000$), but the overall pattern of factor loadings was invariant (Table 4). For example, in both males and females VO_{2max} loaded most strongly onto HRF while curl-ups had the lowest loading (Table 4).

Discussion

The primary purpose of this study was to test the factor structure of a four-component HRF construct, based largely on the dominant definition of HRF (Caspersen et al., 1985), but excluding body composition as a component due to the aforementioned problems associated with its inclusion. The underlying factor structure of the hypothetical HRF construct has not been reported in youth. Testing the structure of this construct was therefore important considering the widespread research in this area, and the need to present HRF as a construct that makes both theoretical and statistical sense.

Results from this study suggest a different definition of HRF to the five-component HRF construct most frequently referred to in the literature. Prior to model testing body composition was excluded as a component of HRF. It is reasonable to argue that due to the difference in nature between assessments of body composition and the remaining HRF components, and the frequent positioning of body composition as the dependent variable when HRF is an independent variable, the inclusion of body composition in an overall HRF construct makes neither theoretical nor statistical sense. Thus, the initial statistical testing of the HRF construct was of a four-component construct containing indicators of the components CRE, MS, ME, and flexibility.

The best fitting model (Hu and Bentler, 1999) proved to be a reduced three-component construct (Figure 3), which, given the known identification problems posed by single-indicator latent variables (Kelloway, 1998; Kline, 2015) (CRE in this case), was re-specified as a first-order model (Figure 4). Correlating the errors for the two MS indicators (HJ and VJ) and the two ME indicators (push-ups and curl-ups) in effect combined these into separate MS and ME components, highlighted by the identical fit statistics for the first-order and second-order models (Table 3). This model suggests then that HRF in youth is best represented by three rather than four components, where indicators with higher loadings (VO_{2max}, HJ, and push-ups) play a more important role. The three components retained in the HRF construct were CRE, MS, and ME. Flexibility was removed due to its low factor loading (0.10). The removal of flexibility as a component of HRF is supported by previous research which has highlighted the lack of association between flexibility and health in youth (Casonatto et al., 2016; Stodden et al., 2015). In their analysis of the factor structure of motor performance ability (MPA) in youth, a construct not dissimilar to HRF, Lämmle et al. (2010) also found a similarly low contribution of flexibility to the overall MPA construct. Results from Lämmle and colleagues (2010), and findings from the present study, suggest that flexibility might be better considered as an independent element rather than being included in an overall fitness construct.

Analysis of the factor loadings within the first-order model in the current study shows that VO_{2max}, as a measure of CRE, is the most important indicator, with a loading of 0.76. The fact that a measure of CRE is of such significance to the overall HRF construct is not surprising, given the widely reported association between CRE and health in youth (Corbin et al., 2014; Ortega et al., 2008; Pate et al., 2012), and the importance placed on this component in studies of youth health and fitness. In fact, CRE is the most extensively measured component of HRF, with researchers often choosing to test CRE exclusively as a measure of HRF when there is a desire to limit the number of tests in a research protocol. Results from this study, however, strongly suggest that HRF cannot be described by just one component. Recently, MF has begun to receive more attention in the health research field, with studies showing positive associations between MF and health in youth (Benson et al., 2006; Grøntved et al., 2015; Janz et al., 2002). Therefore, including measures of both MS and ME makes sense not only from a statistical, but from a theoretical perspective. If CRE alone is used as a measure of HRF, as has been the case in many studies (Barnett et al., 2008; Kriemler et al., 2010; Woods et al., 2010), then the important contribution of MS and ME would be ignored.

Results presented in this paper demonstrate that HJ as a measure of MS, and push-ups as a measure of ME, both have relatively high factor loadings ($r = .71$ and $r = .60$, respectively) within the HRF construct. Despite their lower factor loadings, VJ ($r = .45$) and curl-ups ($r = .38$), as measures of MS and ME, respectively, also contribute to overall model fit. HJ has been previously identified as a good overall measure of MS, with strong associations between it and other lower body ($R^2 = 0.829$–0.864) and upper body ($R^2 = 0.694$–0.851) strength tests (Castro-Pinero et al., 2010). What is more, HJ as a measure of MS has been found to negatively correlate with body mass index (BMI) (Brunet et al., 2007; Riddiford-Harland et al., 2006), a well-established marker of health (Baker et al., 2007), which further lends support to the relevance of HJ within the HRF construct. VJ has also been associated with body composition in youth, with performance on the VJ test significantly and inversely associated with body fatness (Minck et al., 2000) and BMI (Riddiford-Harland et al., 2006).

Push-ups and curl-ups as measures of ME are less well researched, with limited studies examining the association between push-up performance and health outcomes in youth. Despite this lack of research, it has been recommended by previous authors to include tests of ME when measuring HRF to give an overall picture of HRF (Plowman and Mahar, 2013). Handgrip was originally included as an indicator of MS in this study but was removed due to its low factor loading ($r = .37$), with a resultant improvement in model fit. Handgrip, while included in many HRF test batteries (Council of Europe, 1983; Ortega et al., 2011) and frequently tested as a marker of MS (Francis et al., 2016; Moliner-Urdiales et al., 2011; Ortega et al., 2011), is found to positively correlate with BMI, with children who have a higher BMI performing better on this test (Artero et al., 2010). In children, tests of MS have generally been found to correlate positively with BMI, while tests of ME have shown a negative correlation with BMI (Hassan et al., 2016). It is reasonable to suggest that MF tests which require the individual to support and move their own body weight (such as HJ, VJ, push-ups and curl-ups) may be more poorly performed by those with higher mass, whereas a test such as the handgrip test is less likely to be affected in this way as it requires minimal movement of the body. It is possible then that handgrip may have contributed to poorer model fit due to an opposing relationship with weight status compared with other tests used in this study.

Of further importance in testing the structure of the HRF construct in youth is assessing the degree to which the construct is the same for males and females. In the current study a multiple-group analysis of the first-order HRF model showed that while there were significant differences between males and females for absolute scores in all individual HRF tests, the structure of the proposed HRF construct was invariant across sex. This is to be expected, given that in studies on HRF in youth, males and females are typically measured on the same tests despite sex differences in absolute scores on HRF tests (Ortega et al., 2011; Santos et al., 2014). While the current study assessed the factor structure of HRF cross-sectionally to produce a HRF composite that proved to be invariant by sex, the sample is limited in terms of age-range. Future studies should test this composite across a broader age range, and longitudinally, to determine if this five-indicator HRF composite is invariant by age-group and by time.

Conclusion

HRF is a positive predictor of health in youth (Ortega et al., 2008). The association between CRE and health is well established (Corbin et al., 2014; Pate et al., 2012), and there is now a growing body of research identifying the positive effect of MS and ME on health outcomes in youth (Smith et al., 2014). As such, HRF is a primary construct of interest in youth health promotion research. Previously, many studies examining associations between HRF and other health-related variables have chosen to measure just one component, with CRE frequently the component of choice (Kriemler et al., 2010; Woods et al., 2010). Considering the strong positive association of MS and ME with health, however, excluding these components is not ideal. The problem that arises with a multi-component construct like HRF is how to combine the relevant components into a single score which reflects the construct and can be used in research examining associations between HRF, as a composite, and other variables. This study addresses these issues and provides researchers with a concise selection of tests that represent the HRF construct in youth, and that are suitable for use in a field-based setting. Findings from this study provide researchers with an effective method to measure HRF. Empirical evidence from this study suggests that HRF in youth is best represented by five indicators representing three components of fitness (CRE, MS and ME). These five indicators involve HRF tests that are widely used (Ortega et al., 2011; Ruiz et al., 2011; Santos et al., 2014), valid and reliable (Artero et al., 2011; Morrow et al., 2010; Ortega et al., 2008), and easily administered in a field-based setting.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Úna Britton https://orcid.org/0000-0002-4454-2085

References

ACSM (2014) *ACSM's Health-Related Physical Fitness Assessment Manual*. 4th edn. Baltimore: Wolters Kluwer, Lippincott Williams & Wilkins.

Arbuckle JL (2014) *IBM® SPSS® Amos™ 23 User's Guide*.

Artero EG, España-Romero V, Ortega FB, et al. (2010) Health-related fitness in adolescents: Underweight, and not only overweight, as an influencing factor. The AVENA study. *Scandinavian Journal of Medicine and Science in Sports* 20(3): 418–427.

Artero EG, Espana-Romero V, Castro-Piñero J, et al. (2011) Reliability of field-based fitness tests in youth. *International Journal of Sports Medicine* 32: 159–169.

Bailey DP, Boddy LM, Savory LA, et al. (2012) Associations between cardiorespiratory fitness, physical activity and clustered cardiometabolic risk in children and adolescents: The HAPPY study. *European Journal of Pediatrics* 171(9): 1317–1323.

Baker JL, Olsen LW, Sørensen TI, et al. (2007) Childhood body-mass index and the risk of coronary heart disease in adulthood. *New England Journal of Medicine* 357(23): 2329–2337.

Barnett L, van BeurdenN E, Morgan PJ, et al. (2008) Does childhood motor skill proficiency predict adolescent fitness? *Medicine & Science in Sports & Exercise* 40(12): 2137–2144.

Benson AC, Torode ME and Singh MAF (2006) Muscular strength and cardiorespiratory fitness is associated with higher insulin sensitivity in children and adolescents. *International Journal of Pediatric Obesity: IJPO* 1(4): 222–231.

Brunet M, Chaput JP and Tremblay A (2007) The association between low physical fitness and high body mass index or waist circumference is increasing with age in children: The 'Qu?? bec en Forme' Project. *International Journal of Obesity* 31(4): 637–643.

Byrne B (2010) *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. 3rd edn. New York: Routledge.

Casonatto J, Fernandes RA, Batista MB, et al. (2016) Association between health-related physical fitness and body mass index status in children. *Journal of Child Health Care* 20(3): 294–303.

Caspersen CJ, Powell KE and Christenson GM (1985) Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Reports (Washington, D.C.: 1974)* 100(2): 126–131.

Castro-Pinero J, Ortega FB, Artero EG, et al. (2010) Assessing muscular strength in youth: Usefulness of standing long jump as a general index of muscular fitness. *Journal of Strength & Conditioning Research* 24(7).

Cohen J (1992) A Power Primer. *Psychological Bulletin [PsycARTICLES 112(1)*. Available at: http://www2.psych.ubc.ca/~schaller/528Readings/Cohen1992.pdf (accessed 18 April 2018).

Corbin C, Welk G, Richardson C, et al. (2014) Youth physical fitness: Ten key concepts. *Journal of Physical Education, Recreation & Dance (JOPERD)* 85(2): 24–31.
Coulson M and Archer D (2009) *Practical fitness testing*. London: A & C Black Publishers Ltd.
Council of Europe (1983) *Testing Physical Fitness: Eurofit*. 1–18.
Fleishman EA (1964) *What Do Physical Fitness Test Measure? A Review of Previous Research*. Englewood Cliffs, NJ: Prentice-Hall.
Fox R (2010) Confirmatory Factor Analysis. *Encyclopedia of Research Design* 181–186.
Francis CE, Longmuir PE, Boyer C, et al. (2016) The Canadian Assessment of Physical Literacy: Development of a model of children's capacity for a healthy, active lifestyle through a Delphi process. *Journal of Physical Activity and Health* 13(2): 214–222.
Grøntved A, Ried-Larsen M, Møller NC, et al. (2015) Muscle strength in youth and cardiovascular risk in young adulthood (the European Youth Heart Study). *British Journal of Sports Medicine* 49: 90–94.
Hassan N, Kamal H and Hussein Z (2016) Relation between body mass index percentile and muscle strength and endurance. *Egyptian Journal of Medical Human Genetics* 17(4): 367–372.
Hu L and Bentler P (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.
Janz KF, Dawson JD and Mahoney LT (2002) Increases in physical fitness during childhood improve cardiovascular health during adolescence: The Muscatine Study. *International Journal of Sports Medicine* 23(Suppl 1): S15–S21.
Kelloway K (1998) *Using LISREL for Structural Equation Modeling; A Researcher's Guide*. London: Sage.
Kline R (2015) *Principles and Practice of Structural Equation Modeling*. 4th edn. London: Sage.
Kriemler S, Zahner L, Schindler C, et al. (2010) Effect of school based physical activity programme (KISS) on fitness and adiposity in primary schoolchildren: Cluster randomised controlled trial. *BMJ (Clinical Research Ed.)* 340: c785.
Lämmle L, Tittlbach S, Oberger J, et al. (2010) A two-level model of motor performance ability. *Journal of Exercise Science and Fitness* 8(1): 41–49.
Léger LA, Mercier D, Gadoury C, et al. (1988) The multistage 20 metre shuttle run test for aerobic fitness. *Journal of Sports Sciences* 6(2): 93–101.
Lima RA, Pfeiffer KA, Bugge A, et al (2017) Motor competence and cardiorespiratory fitness have greater influence on body fatness than physical activity across time. *Scandinavian Journal of Medicine and Science in Sports* (January): 1–10.
Marsh HW (1993) The multidimensional structure of physical fitness: Invariance over gender and age. *American Educational Research Journal* 30(4): 841–860.
Meredith MD and Welk GJ (2010) *Fitnessgram/Activitygram Test Administration Manual*. 4th edn. Champaign: Human Kinetics.
Minck MR, Ruiter LM, Van Mechelen W, et al. (2000) Physical fitness, body fatness, and physical activity: The Amsterdam Growth and Health Study. *American Journal of Human Biology* 12(March 1999): 593–599.
Moliner-Urdiales D, Ruiz JR, Vicente-Rodriguez G, et al. (2011) Associations of muscular and cardiorespiratory fitness with total and central body fat in adolescents: The HELENA study. *British Journal of Sports Medicine* 45(2): 101–108.
Morrow JR, Martin SB and Jackson AW (2010) Reliability and validity of the FITNESSGRAM®: Quality of teacher-collected health-related fitness surveillance data. *Research Quarterly for Exercise and Sport* 81. S24–S30.
Ortega F, Ruiz JR, Castillo MJ, et al. (2008a) Physical fitness in childhood and adolescence: A powerful marker of health. *International Journal of Obesity* 32(1): 1–11.
Ortega F, Artero E, Ruiz J, et al. (2008b) Reliability of health-related physical fitness tests in European adolescents. The HELENA Study. *International Journal of Obesity* 32(10): 549–557.
Ortega FB, Artero EG, Ruiz JR, et al. (2011) Physical fitness levels among European adolescents: The HELENA study. *British Journal of Sports Medicine* 45: 20–29.
Pate RR (1988) The evolving definition of physical fitness. *Quest* 40(3): 174–179.

Pate R, Oria M and Pillsbury L (2012) *Fitness Measures and Health Outcomes in Youth*. National Academies Press 11.
Payne G and Isaacs L (2016) *Human Motor Development: A Lifespan Approach*. 9th edn. New York: Routledge.
Plowman SA and Mahar MT (2013) *FITNESSGRAM /ACTIVITYGRAM Reference Guide* (4th Edition). Dallas, TX: The Cooper Institute.
Riddiford-Harland DL, Steele JR and Baur LA (2006) Upper and lower limb functionality: Are these compromised in obese children? *International Journal of Pediatric Obesity* 1(1): 42–49.
Rodrigues LP, Stodden DF and Lopes VP (2016) Developmental pathways of change in fitness and motor competence are related to overweight and obesity status at the end of primary school. *Journal of Science and Medicine in Sport / Sports Medicine Australia* 19(1): 87–92.
Ruiz JR, Castro-Piñero J, Artero EG, et al. (2009) Predictive validity of health-related fitness in youth: A systematic review. *British Journal of Sports Medicine* 43(12): 909–923.
Ruiz JR, Castro-Piñero J, España-Romero V, et al. (2011) Field-based fitness assessment in young people: The ALPHA health-related fitness test battery for children and adolescents. *British Journal of Sports Medicine* 45(6): 518–524.
Santos R, Mota J, Santos DA, et al. (2014) Physical fitness percentiles for Portuguese children and adolescents aged 10-18 years. *Journal of Sports Sciences* (May 2014): 1–9.
Schumacker ER and Lomax GR (1996) *A Beginner's Guide to Structural Equation Modeling*. New Jersey: Erlbaum.
Smith JJ, Eather N, Morgan PJ, et al. (2014) The health benefits of muscular fitness for children and adolescents: A systematic review and meta-analysis. *Sports Medicine* 44(9): 1209–1223.
Stodden D, Sacko R and Nesbitt D (2015) A review of the promotion of fitness measures and health outcomes in youth. *American Journal of Lifestyle Medicine* 11: 232–242.
Stodden DF, et al. (2008) A Developmental Perspective on the Role of Motor Skill Competence in Physical Activity: An Emergent Relationship. *Quest* 60(2): 290–306. DOI: 10.1080/00336297.2008.10483582.
Tabachnick BG and Fidell LS (2007) *Using Multivariate Statistics*. 5th edn. New York: Allyn and Bacon.
Tomkinson GR, Carver KD, Atkinson F, et al. (2018) European normative values for physical fitness in children and adolescents aged 9-17 years: Results from 2 779 165 Eurofit performances representing 30 countries. *British Journal of Sports Medicine* 52(22): 1445–1456.
Woods CB, Tannehill D, Quinlan A, et al (2010) The Children's Sport Participation and Physical Activity Study (CSPPA). Research Report No 1. *School of Health and Human Performance, Dublin City University and The Irish Sports Council, Dublin, Ireland*.

Author biographies

Úna Britton is a Researcher in Child and Adolescent Physical Activity and Health at Dublin City University.

Johann Issartel is an Associate Professor of Motor Development and Physical Education at Dublin City University.

Gerry Fahey is a Researcher in Work and Organisational Psychology, Structure and Measurement of Personality.

Gerry Conyngham is an Assistant Professor of Statistics at Dublin City University.

Sarahjane Belton is an Associate Professor of Physical Education.