British Journal of Health Psychology (1997), 2, 361–376 *Printed in Great Britain*

The measurement of exercise motives: Factorial validity and invariance across gender of a revised Exercise Motivations Inventory

David Markland* and David K. Ingledew

School of Sport, Health and Physical Education Sciences, University of Wales, Bangor, Ffriddoedd Building, Victoria Drive, Bangor LL57 2EN, Wales

Objectives. The aim of this study was to further develop and refine the Exercise Motivations Inventory (EMI), a measure of individuals' reasons for exercising.

Design. Confirmatory factor analytic procedures using LISREL were employed to test the hypothesized 14-factor structure of the revised instrument (the EMI-2) and the invariance of the factor structure across gender.

Methods. Four hundred and twenty-five civil servants completed the revised instrument. Analyses were conducted in three phases. Phase 1 involved detailed examination of the fit of the 14-factors separately in order to detect and eliminate poor indicators. In Phase 2 each factor was paired with every other factor in order to detect and eliminate ambiguous items. In Phase 3 factors were grouped with conceptually related factors into five submodels and the fit and factorial invariance across gender of these submodels was tested.

Results. Item elimination at Phases 1 and 2 led to the development of a set of internally consistent factors with strong indicators and good discriminant validity. Phase 3 gave further evidence for the convergent and discriminant validity of the items and factors and strong support for the invariance of the factor structure across gender.

Conclusions. The EMI-2 is a factorially valid means of assessing a broad range of exercise participation motives in adult males and females, applicable to both exercisers and non-exercisers.

The study of participation motives has formed a cornerstone of exercise adherence research. A number of authors have pointed out that given the well-documented physical and psychological benefits of an active life-style, an understanding of why some individuals choose to exercise and others do not would be of great practical value (e.g. Duda, 1989; Willis & Campbell, 1992). Furthermore, several theoretical positions that have been applied to the physical activity domain consider individuals' exercise goals or objectives to be a central determinant of participation (Markland & Hardy, 1993; Weiss & Chaumeton, 1992).

*Requests for reprints.

For example, applications of Deci & Ryan's (1985, 1990) self-determination theory to exercise participation suggest that specific exercise motives can be either intrinsically or extrinsically oriented. Intrinsic motives are concerned with experiences of competence and interest–enjoyment whereas extrinsic motives are focused on the achievement of outcomes that are extrinsic to participation in the activity *per se* (Deci & Ryan, 1985; Pelletier, Fortier, Vallerand, Tuson, Brière & Blais, 1995; Ryan, Vallerand & Deci, 1984). It has been suggested that these different motivational orientations will have different cognitive, emotional and behavioural consequences. Extrinsic motives may lead to tension, pressure to perform and feelings of compulsion whereas intrinsic motives allow freedom from pressure and the experience of choice (Deci & Ryan, 1985). Participation motives such as exercising for enjoyment, challenge, skill improvement and affiliation have been characterized as intrinsic whilst exercising for reasons such as appearance improvement, weight control and social recognition have been considered to reflect extrinsic motivation (Duda & Tappe, 1989*a*; Frederick & Ryan, 1993, 1995; Markland, Ingledew, Hardy & Grant, 1992; Willis & Campbell, 1992).

A number of instruments have been developed to assess individuals' exercise motives. One of the most widely used is the Reasons for Exercise Inventory (REI; Silberstein, Streigel-Moore, Timko & Rodin, 1988), which originally comprised seven scales labelled weight control, attractiveness, tone, fitness, health, mood and enjoyment. In a recent confirmatory factor analytic study Davis, Fox, Brewer & Ratunsy (1995) found support for a less differentiated, six-factor version of the REI in which fitness and health items were collapsed into a single subscale. Frederick & Ryan's (1993) Motivation for Physical Activity Measure (MPAM) taps an even more restricted range of motives, comprising only three scales labelled enjoyment, competence and body-related motives. Whilst both the REI and MPAM have proven useful in testing theoretically driven research questions, it seems clear from the exercise participation literature that individuals have a broader and more differentiated conception of exercise motives (Duda, 1989; Duda & Tappe, 1988, 1989*a*; Ingledew, Hardy & de Sousa, 1995; Markland *et al.*, 1992; Markland & Hardy, 1993).

Duda & Tappe's (1989*b*) Personal Incentives for Exercise Questionnaire, comprising 10 scales, does assess a broader range of motives. However, Markland & Hardy (1993) criticized this instrument on the grounds that it fails to assess enjoyment as a motive for exercising and because a number of items could be read by respondents as reflecting general beliefs about exercise or about themselves, rather than specifically tapping their reasons for exercising. Markland & Hardy went on to describe the development and initial construct validation of the Exercise Motivations Inventory (EMI), an instrument that also assesses a broad range of reasons for exercising.

The EMI comprises 12 scales labelled stress management, weight management, recreation, social recognition, enjoyment, appearance, personal development, affiliation, ill-health avoidance, competition, fitness and health pressures. Further studies have supported the validity of the EMI. Markland *et al.* (1992) found that the EMI discriminated between women taking part in community aerobics classes and members of a Weight Watchers group taking part in aerobics as part of their weight-reduction programme. Drawing on Deci & Ryan's (1985) self-determination theory, the authors hypothesized that community aerobics participants' exercise motives would be more intrinsically oriented than those of the Weight Watchers. A discriminant analysis showed that the two groups were significantly differentiated in terms of a function defined by exercising for enjoyment, recreation, personal development, fitness, stress management and affiliation. Ingledew *et al.* (1995) found that EMI weight-management scores were differentially predicted amongst males and females by body mass index and body-shape dissatisfaction. Analyses showed that for men weight management was predicted by body mass index but not body-shape dissatisfaction, whilst for women weight management was predicted by body-shape dissatisfaction but not body mass index.

Despite promising signs concerning the validity of the EMI, Markland & Hardy (1993) also pointed to a number of weaknesses with the instrument, particularly in relation to fitness-related motives. Specifically, the EMI fitness scale, which was found to be relatively low in internal consistency, comprises only three items. One of these refers to improving fitness in a general sense whilst the other two refer to improving flexibility. Other fitness-related motives, such as exercising to improve strength, speed and endurance, are not tapped by the EMI. The authors reported that other fitness-related items in the initial pool failed to load unambiguously on any factor. Given the range of physical benefits to be gained by exercising, this is clearly a shortcoming of the instrument.

Another weakness of the EMI concerns the measurement of health-related motives. The health-related scales, Health Pressures and Ill-health Avoidance, both have rather negative connotations. Although such extrinsic motives may be important to some individuals, particularly in the decision to adopt exercise in the first place (Dishman, 1987; McAuley, Wraith & Duncan, 1991), movement towards health could for others be a more positive, intrinsically oriented motivational force. In support of this notion, Kasser & Ryan (1996) found that health-related goals were closely related to intrinsically motivated aspirations such as affiliation and self-acceptance. Similarly, Duda & Tappe (1989*a*), using discriminant function analysis to examine motivational differences across age and gender, found that health benefits loaded positively on a function labelled wellness incentives which also comprised coping with stress and affiliation. Thus with respect to health it would seem appropriate to tap both approach- and avoidance-related motives.

A further problem with the EMI is that the phrasing of the original item stem makes the instrument only applicable to individuals who do exercise. It would be useful, however, to be able to tap the reasons that non-exercisers would have for exercising if they were to do so, in order to determine factors that might motivate initial involvement or reinvolvement. Such information would have practical benefits in terms of targeting interventions for sedentary individuals.

A more fundamental issue with the EMI, and indeed other measures of exercise motives, concerns the lack of a strong theoretical basis. Several authors have pointed out that the study of participation motives at the descriptive or surface level needs to be embedded within a more theoretical approach (e.g. Biddle, 1995; Gould & Petlichkoff, 1988). Whilst the EMI draws loosely on self-determination theory in the sense that some motives can be held to reflect intrinsic or extrinsic motivation, there remain problems with fitting some other motives into this framework. For example, fitness-related reasons for exercising could be held to be intrinsic to participation (if one exercises one does get fitter) or to reflect an extrinsic outcome to which an individual might aspire. Nevertheless, as Biddle (1995) has argued, knowledge of surface-level participation motives is important from a practical perspective in the promotion of exercise. This is because an understanding of individuals' participation motives can help in tailoring exercise interventions to meet personal needs (Willis & Campbell, 1992).

The aim of this study was to improve and refine the original EMI in order to produce a set of valid and reliable indicators of a broad range of participation motives. Based on an examination of the items and factor loadings in the original EMI development, several new items were generated, or resurrected from the original's initial item pool, to extend the range of fitness and health-related motives and to increase the number of indicators of some other scales. Modifications were made to the item stem to broaden the applicability of the instrument to non-exercisers as well as exercisers. In addition, analyses were conducted to test the invariance of the factor structure across gender. A confirmatory factor analytic approach was adopted. With the exception of Davis *et al.*'s (1995) analysis of the REI, previous examinations of the factorial validity of measures of exercise participation motives have relied on exploratory factor analyses. Given the existence of a clearly defined model of exercise motives with item-factor relationships specified *a priori*, it was deemed important that the factorial validity of the new EMI should be confirmed within a hypothesis-testing approach. In addition, confirmatory factor analysis allows one to test the fit of a model to multiple groups subject to the constraints that any one or any set of parameter estimates are invariant across groups, giving a powerful test of the generalizability of the factor structure across different samples (Marsh, 1993).

Method

Participants and procedure

The data were collected as part of a larger study examining exercise motivations and the stages of change for exercise, the results of which are not reported here. The participants were civil servants working in a large government establishment. One thousand questionnaires were distributed by a contact on the site, together with a letter explaining that the research concerned motivations for exercise and seeking informed consent. Questionnaires were returned by 425 participants of whom 282 were male (mean age = 38.66, SD = 9.95) and 143 were female (mean age = 36.14, SD = 9.62). When classified by stage of change (Prochaska & DiClemente, 1983), 82 participants (19.3 per cent) were in precontemplation (having no intention of changing behaviour), 57 (13.4 per cent) in contemplation (thinking of starting exercising in the next six months), 48 (11.3 per cent) in preparation (taking active steps to start exercising in the near future), 35 (8.2 per cent) in action (exercising regularly but only for the last six months) and 203 (47.8 per cent) in maintenance (having exercised regularly for more than six months).

Instruments

Participants completed the revised EMI (EMI-2, see below), a definition of regular exercise adapted from Loughlan & Mutrie (1995) and a measure of stage of change adapted from Marcus, Selby, Niaura & Rossi (1992). Regular exercise was defined as 'exercise (e.g. swimming, jogging, weight training, aerobics) 2–3 times per week, or sport (e.g. golf, hockey, football) 2–3 times per week'.

Item selection and hypothesized factor structure of the EMI-2

Twenty-five items were added to the 44 items of the original EMI. Fourteen factors were hypothesized to underlie the item set: stress management, revitalization (formerly labelled recreation in the original version of the EMI), enjoyment, challenge (labelled personal development in the original EMI), social recognition, affiliation, competition, health pressures, ill-health avoidance, positive health, weight management, appearance, strength and endurance, and nimbleness. Extra items were included for revitalization, enjoyment, appearance, challenge, affiliation, competition and ill-health avoidance. The new positive health scale comprised five items. The original fitness scale was replaced by two new scales, strength and endurance, and nimbleness. Based on results from the original EMI it was not considered worthwhile attempting to further differentiate fitness-related motives. Five items were included for strength and endurance and two new items were added for nimbleness. Two items loading on appearance in the original EMI were now considered to be more properly indicators of social recognition. Items were presented in random order.

The EMI instructions and item stem were reworded so that the items would be applicable to current non-exercisers. Participants were asked to indicate whether or not each statement was true for them personally, or would be true for them if they did exercise. The stem now read: 'Personally, I exercise (or might exercise) . . .'. This led to minor modifications to the wording of some of the original EMI items. Responses were made on a six-point Likert-type scale ranging from 'not at all true for me' to 'very true for me'.

Model-testing strategy and assessment of fit

Factorial validity was tested by analyses of covariance structures using LISREL 8.12 (Jöreskog & Sörbom, 1993*a*). A sequential model-testing approach was adopted (Jöreskog, 1993), involving three phases. In Phase 1 separate single latent variable models were tested for each of the 14 proposed factors. The aim here was to eliminate items that were poor indicators of their factors by examination of global goodness of fit and of the parameter estimates, along with detailed examination of the residuals. Using these sources of information, one or both of any problematic items were eliminated and the fit reassessed. In Phase 2 each factor was paired with every other factor in order to identify and eliminate ambiguously loading items. This was achieved again by examination of global goodness of fit, parameter estimates, and residuals, as well as examination of the modification indices for the regression matrices relating the factors to the items. Large modification indices indicate that the fit would be improved if items were allowed to cross-load on a non-intended factor and therefore that such items are factorially ambiguous. In addition, the discriminant validity of the scales was investigated by examining the 95 per cent confidence interval (± 1.96 standard errors) around each correlation between factors (Anderson & Gerbing, 1988). Ideally, the next phase would have tested a single model in which all the items and factors were included. However, the resultant model would have had far too many parameters to allow accurate estimation even given a much larger sample size than was available here. Consequently, in Phase 3 the complete model was approximated by grouping sets of conceptually related factors (and their items) into five separate submodels and these were tested separately. The submodels were labelled psychological motives (stress management, revitalization, enjoyment and challenge); interpersonal motives (social recognition, affiliation and competition); health motives (health pressures, ill-health avoidance and positive health); body-related motives (weight management and appearance); and fitness motives (strength and endurance, and nimbleness). The aim here was to determine that the factorial validity of the scales was good even when the factors were grouped within similar domains, where problems with convergent and discriminant validity were more likely to be evident.

Phase 3 also involved testing for the factorial invariance of the EMI-2 across gender by simultaneously fitting each submodel to the data from males and females. A hierarchical approach was adopted (Bollen, 1989; Byrne, 1989; Jöreskog & Sörbom, 1981, 1989). Each submodel was fitted to the combined data for males and females and then to the male and female data separately, to establish the adequacy of the five baseline models. Next, the invariance across gender of the pattern of fixed and non-fixed parameters was tested; this was followed by tests of the invariance of the factor loadings; followed by tests of the invariance of the factor loadings and measurement errors; followed by tests of the invariance of the factor loadings, measurement errors, and factor variances and correlations. Goodness of fit was assessed at each stage. If the fit of a model remains good when invariance constraints are imposed, then there is evidence supporting the addition of those constraints (Marsh, 1993).

Multiple criteria were employed in the assessment of fit, as outlined above. Global fit was assessed by the χ^2 likelihood ratio statistic, the root mean square error of approximation (RMSEA; Steiger, 1990) and its associated p value for RMSEA $< .05$, the Goodness-of-Fit Index (GFI; Jöreskog & Sörbom, 1981), the Non-normed Fit Index (NNFI; Tucker & Lewis, 1973) and the Comparative Fit Index (Bentler, 1990). In the tests of factorial invariance the Parsimony Normed Fit Index (PNFI; James, Mulaik & Brett, 1982) was also used. Due to the well-documented problems with employing χ^2 as a test statistic (e.g. Bentler & Bonett, 1980; Jöreskog, 1993; Marsh, Balla & McDonald, 1988; Mulaik, James, Van Alstine, Bennett, Lind & Stilwell, 1989), the χ^2 value was used as a badness-of-fit measure in the sense that small values were held to indicate a good fit and large values a poor fit, with the number of degrees of freedom being used as a standard by which to judge the size of χ^2 (Jöreskog, 1993; Jöreskog & Sörbom, 1989).

Initial analyses indicated that the multivariate distribution of scores was significantly non-normal (Mardia coefficients for skewness and kurtosis were 208.857 and 36.972 respectively). In addition, item responses employed six ordered categories. Therefore the model was estimated using the weighted least squares method. Polychoric correlations and asymptotic covariance matrices were calculated using PRELIS 2.12 (Jöreskog & Sörbom, 1993*b*) and used as data input.[1] Separate input matrices were calculated for each analysis using listwise deletion. There was no evidence of systematic non-responding to items.

Results

Phase 1: Single-scale analyses

Table 1 shows the items and results for the single-scale analyses following item deletion. Effective sample sizes ranged from 418 to 423. On the basis of an examination of the residuals and modification indices, two items were eliminated from the stress management scale, one from weight management, one from revitalization, three from social recognition, two from appearance, one from affiliation, one from positive health and two from strength and endurance. Following this all the factor loadings were statistically significant for all scales, as evidenced by very high t values (not shown to save space). The fit of all the models except ill-health avoidance was excellent with χ^2 non-significant, RMSEA small and not significantly greater than .05 at the 5 per cent level, and GFI, NNFI and CFI approaching or reaching unity. For ill-health avoidance, RMSEA was not significantly greater than .05 and GFI, NNFI and CFI were high, but a large χ^2 relative to its degrees of freedom indicated some degree of misfit in the model. However, the pattern of the residuals and modification indices did not suggest that a significantly better fitting model could be achieved by item elimination so it was decided to proceed with the scale as it stood.

Phase 2: Scale-pairing analyses

The second phase involved pairing each scale with every other scale in order to detect ambiguous items. Effective sample sizes ranged from 415 to 423. Since there were 91 possible pairings the results are not given in detail here. To summarize, the initial fit of all pairings except some of those involving revitalization, challenge, ill-health avoidance, positive health and nimbleness were good to excellent. Examination of the residuals and modification indices suggested that the fit of the problem pairings could be improved by eliminating one item from each of these scales. This was done and all pairings involving these scales were re-examined. In each case the fit was now acceptable. For all 91 pairings following item deletion only 12 models failed to achieve either a non-significant χ^2 or RMSEA not significantly greater than .05, and in all cases GFI, NNFI and CFI were good (minimum values were .974, .910 and .950 respectively). Factor loadings were similar in all pairings to those estimated at Phase 1. Furthermore, in the least well fitting cases the modification indices suggested that there would be significant improvements in fit if one item was freed to cross-load on the other factor. However, examination of the estimated changes in parameter estimates associated with these modification indices suggested that if these relaxations were made, in each case the loading of the item on its non-intended

[1] Raw data and computed input matrices are available from the first author on request.

factor would be small, indicating that there was in fact little ambiguity associated with those items.

Table 2 shows the factor correlations obtained from the pairings, together with their standard errors. Confidence intervals failed to include 1.0 in all cases. However, the upper bound of the confidence interval (.997) for the correlation between revitalization and enjoyment approached unity. Thus the discriminant validity of these scales with respect to one another is in doubt. Finally, in order to assess the internal consistency of the scales following item deletion, Cronbach's alpha reliability coefficients were calculated. These are shown in Table 3, together with means and standard deviations for the scales. Alphas indicated that the reliability of all the scales was good apart from health pressures, which was only just acceptable. With minor exceptions, therefore, the results from Phases 1 and 2 suggested that we now had a strong set of individual factors with excellent indicators. Fifty-one items remained with nine factors comprising four items and five factors comprising three items.

Phase 3: Submodel groupings and factorial invariance across gender

This phase involved testing the fit of five oblique confirmatory factor-analytic submodels and testing the factorial invariance of the submodels across gender. Table 4 shows the fit statistics for male and female data combined and for the submodels fitted to male and female data separately. The fit was good for all five submodels. Factor loadings did not differ substantially from those obtained in the Phase 1 analyses. Factor correlations were similar to those found at Phase 2. The fit remained good for all submodels at each stage in the hierarchy of tests of invariance. In each case RMSEA remained not significantly greater than .05 and values for the absolute and incremental fit indices for the more constrained models remained close to those of the corresponding baseline models. Indeed, in some cases they were better in the more constrained models. PNFI increased systematically as sets of parameters were constrained to be equal for each submodel. Thus these analyses gave strong support for the invariance of the submodels across gender.

Discussion

The aims of this research were to improve the Exercise Motivations Inventory, especially with respect to the measurement of health- and fitness-related motives, to extend its applicability to non-exercisers and to test the invariance of the factor structure across gender. The Phase 1 analyses showed that at the single-scale level, following item elimination, the items were all very good indicators of their factors. Phase 2 supported this and showed that there was little ambiguity associated with the items. Thus at this level the ideal of a unidimensional measurement structure with items indicating one and only one latent variable was realized (Anderson & Gerbing, 1982; Gerbing & Anderson, 1984). Examination of the patterning of residuals and modification indices proved a useful tool in the identification and elimination of problematic items. Item elimination inevitably led to some loss of information, since the patterning of residuals in certain cases suggested that there were further, unexplored latent variables underlying the observations. However, exploration of these further latent variables would have had little practical utility since it could only have led to the formation of single-item or at best two-item factors. Furthermore, such an approach would have risked capitalizing on chance.

Table 1. Fit indices and item loadings following item elimination at Phase 1

Scale/items	Loading	χ^2	d.f.	$p(\chi^2)$	RMSEA	GFI	NNFI	CFI
Stress management ($N = 421$)		.512	2	.774	.000 n.s.	.999	1.000	1.000
To give me space to think	.752							
Because it helps to reduce tension	.908							
To help manage stress	.926							
To release tension	.954							
Revitalization ($N = 422$)		2.695	2	.259	.029 n.s.	.998	.998	.999
Because it makes me feel good	.862							
Because I find exercise invigorating	.886							
Because after exercising I feel refreshed	.936							
To recharge my batteries	.751							
Enjoyment ($N = 418$)		.842	2	.656	.000 n.s.	.999	1.000	1.000
Because I enjoy the feeling of exerting myself	.806							
Because I find exercising satisfying in and of itself	.907							
For enjoyment of the experience of exercising	.936							
Because I feel at my best when exercising	.804							
Challenge ($N = 421$)		4.690	5	.455	.000 n.s.	.998	1.000	1.000
To give me goals to work towards	.806							
To help me explore the limits of my body	.768							
To give me personal challenges to face	.929							
To develop personal skills	.704							
To measure myself against personal standards	.834							
Social recognition ($N = 419$)		1.120	2	.571	.000 n.s.	.999	1.000	1.000
To show my worth to others	.827							
To compare my abilities with other peoples'	.889							
To gain recognition for my accomplishments	.908							
To accomplish things that others are incapable of	.828							
Affiliation ($N = 421$)		1.358	2	.507	.000 n.s.	.999	1.000	1.000
To spend time with friends	.823							
To enjoy the social aspects of exercising	.948							
To have fun being active with other people	.935							
To make new friends	.822							

Table 1. Continued

Scale/items	Loading	χ^2	d.f.	$p(\chi^2)$	RMSEA	GFI	NNFI	CFI
Competition ($N = 422$)		2.254	2	.324	.017 n.s.	.999	.999	1.000
Because I like trying to win in physical activities	.917							
Because I enjoy competing	.970							
Because I enjoy physical competition	.968							
Because I find physical activities fun, especially when competition is involved	.937							
Health pressures ($N = 423$)		1.583	2	.453	.000 n.s.	.998	1.000	1.000
Because my doctor advised me to exercise	.736							
To help prevent an illness that runs in my family	.736							
To help recover from an illness/injury	.736							
Ill-health avoidance ($N = 423$)		9.326	2	.009	.093 n.s.	.997	.989	.996
To avoid heart disease	.827							
To prevent health problems	.956							
To avoid ill-health	.958							
Because I feel I have to exercise to stay healthy	.642							
Positive health ($N = 420$)		3.773	2	.152	.046 n.s.	.998	.995	.998
To help me live a longer, more healthy life	.750							
To have a healthy body	.873							
Because I want to maintain good health	.898							
To feel more healthy	.857							
Weight management ($N = 421$)		.377	2	.828	.000 n.s.	.999	1.000	1.000
To stay slim	.803							
To lose weight	.927							
To help control my weight	.953							
Because exercise helps me to burn calories	.863							
Appearance ($N = 422$)		1.158	2	.560	.000 n.s.	.999	1.000	1.000
To help me look younger	.635							
To have a good body	.825							
To improve my appearance	.944							
To look more attractive	.878							

Table 1. Continued

Scale/items	Loading	χ^2	d.f.	$p(\chi^2)$	RMSEA	GFI	NNFI	CFI
Strength and endurance ($N = 421$)		1.035	2	.596	.000 n.s.	.999	1.000	1.000
To build up my strength	.851							
To increase my endurance	.654							
To get stronger	.948							
To develop my muscles	.831							
Nimbleness ($N = 420$)		3.146	2	.207	.037 n.s.	.998	.997	.999
To get faster	.502							
To stay/become more agile	.854							
To maintain flexibility	.889							
To stay/become flexible	.941							

Note. RMSEA = Root Mean Square Error of Approximation; GFI = Goodness of Fit Index; NNFI = Non-normed Fit Index; CFI = Comparative Fit Index. n.s. = RMSEA not significantly greater than .05 at the 5 per cent level.

Table 2. Factor correlations (and standard errors) following item elimination at Phase 2

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 Stress management	–													
2 Revitalization	.828	–												
	(.024)													
3 Enjoyment	.627	.975	–											
	(.041)	(.011)												
4 Challenge	.569	.728	.809	–										
	(.046)	(.031)	(.027)											
5 Social recognition	.386	.517	.638	.871	–									
	(.049)	(.043)	(.035)	(.026)										
6. Affiliation	.256	.436	.514	.599	.718	–								
	(.054)	(.051)	(.048)	(.044)	(.034)									
7 Competition	.239	.506	.668	.734	.876	.739	–							
	(.057)	(.050)	(.036)	(.034)	(.021)	(.032)								
8 Health pressures	.332	−.075	−.265	−.054	.107	−.123	−.141	–						
	(.060)	(0.69)	(.061)	(.065)	(.064)	(.057)	(0.57)							
9 Ill-health avoidance	.394	.420	.251	.350	.148	.036	−.002	.473	–					
	(.052)	(.052)	(.053)	(.049)	(.054)	(.059)	(.053)	(.051)						
10 Positive health	.515	.724	.596	.516	.259	.136	.189	.123	.749	–				
	(.043)	(.038)	(.039)	(.042)	(.053)	(.058)	(.053)	(.064)	(.031)					
11 Weight management	.391	.381	.240	.269	.163	−.073	.021	.260	.417	.585	–			
	(.049)	(.052)	(.054)	(.056)	(.056)	(.056)	(.053)	(.062)	(.050)	(.040)				
12 Appearance	.409	.527	.465	.584	.459	.075	.115	.172	.464	.714	.760	–		
	(.051)	(.048)	(.048)	(.044)	(.049)	(.058)	(.056)	(.065)	(.047)	(.033)	(.030)			
13 Strength	.480	.624	.670	.805	.577	.189	.393	.152	.401	.717	.467	.707	–	
	(.048)	(.041)	(.035)	(.027)	(.041)	(.055)	(.052)	(.062)	(.048)	(.032)	(.050)	(.036)		
14 Nimbleness	.561	.678	.593	.588	.449	.275	.248	.262	.587	.763	.500	.694	.694	–
	(.044)	(.038)	(.044)	(.040)	(.045)	(.056)	(.052)	(.064)	(.044)	(.030)	(.051)	(.035)	(.033)	

Table 3. Means, standard deviations and alpha reliability coefficients for the scales following item elimination at Phase 2

Scale	Mean	SD	Alpha
Stress management	2.438	1.431	.916
Revitalization	2.644	1.340	.832
Enjoyment	2.350	1.448	.899
Challenge	1.745	1.315	.857
Social recognition	.905	1.059	.878
Affiliation	1.884	1.427	.910
Competition	1.581	1.632	.954
Health pressures	1.035	1.214	.686
Ill-health avoidance	2.924	1.374	.901
Positive health	3.470	1.188	.877
Weight management	2.829	1.591	.914
Appearance	1.976	1.352	.859
Strength	2.395	1.332	.864
Nimbleness	2.670	1.357	.899

The Phase 2 analyses also gave evidence for the discriminant validity of the scales, except in the case of enjoyment and revitalization. Nevertheless, there was no evidence that the items belonging to these scales were seriously ambiguous. Furthermore, the two sets of items were, as hypothesized *a priori*, conceptually distinct in that the enjoyment items reflected enjoyment of the act of exercising whilst the revitalization items reflected feeling invigorated by exercise. Consequently it was felt justifiable to maintain the distinction between these constructs.

The hypothesized factor structure of the EMI-2 was further supported at Phase 3. The fit of the baseline submodels was good even when the factors were grouped with conceptually related factors. In addition, this phase gave strong support for the invariance of the factor structure across males and females. This is an issue that has not been addressed in previous attempts to measure exercise motives and is important as it shows that the EMI-2 scale scores are not confounded by gender and that they can be used to make meaningful comparisons between levels of males' and females' exercise motives. However, factorial invariance was only demonstrated at the level of the five submodels. Clearly the inability to test a complete model comprising all 14 factors was a weakness of the present study. Nevertheless, as mentioned earlier it would be unrealistic to expect to obtain a good fit with such a large model, even given a very large sample size. When testing a large model it is difficult to pinpoint sources of poor fit simply because there are so many degrees of freedom and therefore so many ways in which the model could be wrong. Breaking the model down as was done here allows for a more precise diagnosis of possible misspecifications. A further potential problem is that only 56 per cent of the participants were currently exercising regularly, as judged by inclusion in the action and maintenance stages of change. It remains possible that the factor structure could differ for exercisers and non-exercisers and future research will need to test for invariance across activity level groupings.

As discussed earlier, recent applications of Deci & Ryan's (1985, 1990) self-determination theory to the exercise domain have characterized specific participation motives as

Table 4. Fit statistics for submodels at Phase 3, including results for male and female groups combined and separately

Model	N	χ^2	d.f.	$p(\chi^2)$	RMSEA	GFI	NNFI	CFI
Psychological motives								
Model fit to males and females combined	413	177.147	84	<.001	.052 n.s.	.985	.987	.989
Model fit to males only	273	137.615	84	<.001	.048 n.s.	.988	.993	.994
Model fit to females only	140	101.489	84	.094	.039 n.s.	.984	.996	.997
Interpersonal motives								
Model fit to males and females combined	418	133.118	51	<.001	.062 n.s.	.992	.993	.994
Model fit to males only	273	104.040	51	<.001	.062 n.s.	.992	.994	.995
Model fit to females only	140	92.799	51	<.001	.076 n.s.	.987	.991	.993
Health motives								
Model fit to males and females combined	417	62.660	24	<.001	.062 n.s.	.987	.981	.987
Model fit to males only	281	61.466	24	<.001	.075	.987	.984	.989
Model fit to females only	139	25.621	24	.373	.022 n.s.	.990	.999	.999
Body-related motives								
Model fit to males and females combined	419	52.870	19	<.001	.065 n.s.	.990	.987	.991
Model fit to males only	277	62.248	19	<.001	.091	.988	.984	.989
Model fit to females only	142	27.631	19	.091	.057 n.s.	.988	.992	.995
Fitness motives								
Model fit to males and females combined	420	14.699	13	.327	.018 n.s.	.996	.999	.999
Model fit to males only	278	15.365	13	.285	.026 n.s.	.996	.999	.999
Model fit to females only	142	25.643	13	.019	.083 n.s.	.988	.987	.992

Note. RMSEA = Root Mean Square Error of Approximation; GFI = Goodness of Fit Index; NNFI = Non-normed Fit Index; CFI = Comparative Fit Index. n.s. = RMSEA not significantly greater than .05 at the 5 per cent level.

reflecting intrinsic or extrinsic motivational orientations (Frederick & Ryan, 1993, 1995; Ryan *et al.*, 1984). However, it was also pointed out that it is difficult to categorize some motives as being exclusively intrinsic or extrinsic. Detailed examination of the pattern of factor correlations found in this study shows that specific exercise motives do not sit comfortably with a simple intrinsic–extrinsic dichotomy. For example, social recognition, which is conventionally considered to have an extrinsic focus (Duda & Tappe, 1989*a*), correlated moderately to strongly with revitalization, enjoyment, challenge and affiliation, factors which reflect constructs that are usually considered to represent intrinsic motivation (e.g. Frederick & Ryan, 1993, 1995; Ryan *et al.*, 1984; Wankel, 1993). Thus whilst some motives, such as enjoyment, challenge and appearance improvement conform reasonably well to conventional definitions of either intrinsic or extrinsic motivation, for others the position is not so obvious. This clearly presents problems from the perspective of trying to embed the study of surface-level participation motives within self-determination theory. However, Deci & Ryan (1985, 1990) have suggested that the simple intrinsic–extrinsic dichotomy may be misleading and that motivation may be better represented by a behavioural regulation continuum ranging from completely non-self-determined to completely self-determined forms of regulation. Deci & Ryan's (1985) organismic integration theory, one of three mini-theories that collectively define self-determination theory, describes the process by which individuals internalize regulation by external constraints and come to feel more self-determined in the regulation of their behaviour. This approach allows for an understanding of how an individual can be both extrinsically and intrinsically motivated at the same time. A way forward, then, may be to examine the role of reasons for exercising in this context in order to develop a stronger theoretical basis for the study of participation motives. Future research with the EMI-2 will seek to do this.

From a methodological perspective, the present series of analyses show that a considerable amount of detailed and useful information about the factorial validity of an instrument can be gained by adopting a rigorous and sequential approach to model testing. They also show that it is possible to use detailed examination of the fit of individual parameters to refine a measurement instrument without departing from a hypothesis-testing approach. Here problematic items were eliminated from the instrument, thus maintaining the basic integrity of the hypothesized factor structure without trying to make models fit by trawling for additional explanatory relationships or *post hoc* meddling with correlated error terms. Furthermore, the Phase 3 analyses showed that a large model can be meaningfully assessed even if sample size limitations preclude testing of the model as a whole.

In conclusion, this study gives strong support for the EMI-2 as a measure of a broad range of males' and females' exercise motives. It is anticipated that it will prove to be a valid and reliable means of assessing exercise motives across different populations and researchers are encouraged to use it to investigate both theoretical and applied questions in the area of exercise and health psychology.[2]

References

Anderson, J. C. & Gerbing, D. W. (1982). Some methods for respecifying measurement models to obtain undimensional construct measurement. *Journal of Marketing Research*, 19, 453–460.

[2]The final version of the EMI-2 is available from the first author on request.

Anderson, J. C. & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411–423.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238–246.

Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness-of-fit in covariance structures. *Psychological Bulletin*, 88, 588–606.

Biddle, S. J. H. (1995). Exercise motivation across the lifespan. In S. J. H. Biddle (Ed.), *European Perspectives on Exercise and Sport Psychology*, pp. 5–25. Champaign, IL: Human Kinetics.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.

Byrne, B. M. (1989). *A Primer of LISREL: Basic Applications and Programming for Confirmatory Factor Analytic Models*. New York: Springer-Verlag.

Davis, C., Fox, J., Brewer, H. & Ratunsy, D. (1995). Motivations to exercise as a function of personality characteristics, age, and gender. *Personality and Individual Differences*, 19, 165–174.

Deci, E. L. & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.

Deci, E. L. & Ryan, R. M. (1990). A motivational approach to self: Integration in personality. In R. A. Dienstbier (Ed.), *Nebraska Symposium on Motivation*, vol. 38, *Perspectives on Motivation*, pp. 237–288. Lincoln: University of Nebraska Press.

Dishman, R. K. (1987). Exercise adherence and habitual physical activity. In W. P. Morgan & S. E. Goldston (Eds), *Exercise and Mental Health*, pp. 57–83. Washington, DC: Hemisphere.

Duda, J. L. (1989). Goal perspectives and behavior in sport and exercise settings. In M. L. Maehr & C. Ames (Eds), *Advances in Motivation and Achievement: Motivation Enhancing Environments*, vol. 6, pp. 81–115. Greenwich, CT: JAI Press.

Duda, J. L. & Tappe, M. K. (1988). Predictors of personal investment in physical activity among middle-aged and older adults. *Perceptual and Motor Skills*, 66, 543–549.

Duda, J. L. & Tappe, M. K. (1989*a*). Personal investment in exercise among adults: The examination of age and gender-related differences in motivational orientation. In A. C. Ostrow (Ed.), *Aging and Motor Behavior*, pp. 239–256. Indianapolis: Benchmark Press.

Duda, J. L. & Tappe, M. K. (1989*b*). The Personal Incentives for Exercise Questionnaire: Preliminary development. *Perceptual and Motor Skills*, 68, 1122.

Frederick, C. M. & Ryan, R. M. (1993). Differences in motivation for sport and exercise and their relations with participation and mental health. *Journal of Sport Behavior*, 16, 124–146.

Frederick, C. M. & Ryan, R. M. (1995). Self-determination in sport: A review using cognitive evaluation theory. *International Journal of Sport Psychology*, 26, 5–23.

Gerbing, D. W. & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*, 11, 572–580.

Gould, D. & Petlichkoff, L. (1988). Participation motivation and attrition in young athletes. In F. L. Smoll, R. A. Magill & M. J. Ash (Eds), *Children in Sport*, 3rd ed., pp. 161–178. Champaign, IL: Human Kinetics.

Ingledew, D. K., Hardy, L. & de Sousa, K. (1995). Body shape dissatisfaction and exercise motivations. *Journal of Sports Sciences*, 13, 60.

James, L. R., Mulaik, S. A. & Brett, J. M. (1982). *Causal Analysis: Assumptions, Models, and Data*. Beverly Hills, CA: Sage.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds), *Testing Structural Equation Models*, pp. 294–316. Newbury Park, CA: Sage.

Jöreskog, K. G. & Sörbom, D. (1981). *LISREL V: Analysis of Linear Structural Relations by the Method of Maximum Likelihood*. Chicago, IL: International Educational Services.

Jöreskog, K. G. & Sörbom, D. (1989). *LISREL 7: User's Reference Guide*. Mooresville, IN: Scientific Software.

Jöreskog, K. G. & Sörbom, D. (1993*a*). *LISREL 8.12*. Chicago, IL: Scientific Software International.

Jöreskog, K. G. & Sörbom, D. (1993*b*). *PRELIS 2: User's Reference Guide*. Chicago, IL: Scientific Software International.

Kasser, T. & Ryan, R. M. (1996). Further examining the American dream: Differential correlates of intrinsic and extrinsic goals. *Personality and Social Psychology Bulletin*, 22, 280–287.

Loughlan, C. & Mutrie, N. (1995). Recruitment of sedentary NHS staff for a workplace exercise programme using an adapted 'stages-of-change' exercise questionnaire. *Journal of Sports Sciences*, 13, 63–64.

McAuley, E., Wraith, S. & Duncan, T. E. (1991). Self-efficacy, perceptions of success, and intrinsic motivation for exercise. *Journal of Applied Social Psychology*, 21, 139–155.

Marcus, B. H., Selby, V. C., Niaura, R. S. & Rossi, J. S. (1992). Self-efficacy and the Stages of Exercise Behavior Change. *Research Quarterly for Exercise and Sport*, 63, 60–66.

Markland, D. & Hardy, L. (1993). The Exercise Motivations Inventory: Preliminary development and validity of a measure of individuals' reasons for participation in regular physical exercise. *Personality and Individual Differences*, 15, 289–296.

Markland, D. Ingledew, D. K., Hardy, L. & Grant, L. (1992). A comparison of the exercise motivations of participants in aerobics and weight watcher exercisers. *Journal of Sports Sciences*, 10, 609–610.

Marsh, H. W. (1993). The multidimensional structure of physical fitness: Invariance over gender and age. *Research Quarterly for Exercise and Sport*, 64, 256–273.

Marsh, H. W., Balla, J. R. & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391–410.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S. & Stillwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, 105, 430–445.

Pelletier, L. G., Fortier, M. S., Vallerand, R. J., Tuson, K. M., Brière, N. M. & Blais, M.R. (1995). Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: The Sport Motivation Scale (SMS). *Journal of Sport and Exercise Psychology*, 17, 35–53.

Prochaska, J. O. & DiClemente, C. C. (1983). Stages and processes of self-change in smoking: Towards an integrative model of change. *Journal of Consulting and Clinical Psychology*, 51, 390–395.

Ryan, R. M., Vallerand, R. J. & Deci, E. L. (1984). Intrinsic motivation in sport: A cognitive evaluation theory interpretation. In W. F. Straub & J. M. Williams (Eds), *Cognitive Sport Psychology*, pp. 230–242. New York: Sport Science Associates.

Silberstein, L. R., Striegel-Moore, R. H., Timko, C. & Rodin, J. (1988). Behavioral and psychological implications of body dissatisfaction: Do men and women differ? *Sex Roles*, 19, 219–232.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173–180.

Tucker, L. R. & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1–10.

Wankel, L. M. (1993). The importance of enjoyment to adherence and psychological benefits from physical activity. *International Journal of Sport Psychology*, 24, 151–169.

Weiss, M. R. & Chaumeton, N. (1992). Motivational orientations in sport. In T. S. Horn (Ed.), *Advances in Sport Psychology*. Champaign, IL: Human Kinetics.

Willis, J. D. & Campbell, L. F. (1992). *Exercise Psychology*. Champaign, IL: Human Kinetics.

Received 21 October 1996; Revised version received 12 June 1997