Research Quarterly for Exercise and Sport
© 1997 by the American Alliance for Health,
Physical Education, Recreation and Dance
Vol. 68, No. 1, pp. 20–32

On the Factorial and Construct Validity of the Intrinsic Motivation Inventory: Conceptual and Operational Concerns

David Markland and Lew Hardy

The Intrinsic Motivation Inventory (IMI) has been gaining acceptance in the sport and exercise domain since the publication of research by McAuley, Duncan, and Tammen (1989) and McAuley, Wraith, and Duncan (1991), which reported confirmatory support for the factorial validity of a hierarchical model of intrinsic motivation. Authors of the present study argue that the results of these studies did not conclusively support the hierarchical model and that the model did not accurately reflect the tenets of cognitive evaluation theory (Deci & Ryan, 1985) from which the IMI is drawn. It is also argued that a measure of perceived locus of causality is required to model intrinsic motivation properly. The development of a perceived locus of causality for exercise scale is described, and alternative models, in which perceived competence and perceived locus of causality are held to have causal influences on intrinsic motivation, are compared with an oblique confirmatory factor analytic model in which the constructs are held at the same conceptual level. Structural equation modeling showed support for a causal model in which perceived locus of causality mediates the effects of perceived competence on pressure-tension, interest-enjoyment, and effort-importance. It is argued that conceptual and operational problems with the IMI, as currently used, should be addressed before it becomes established as the instrument of choice for assessing levels of intrinsic motivation.

Key words: intrinsic motivation, causal modeling

The Intrinsic Motivation Inventory (IMI) has been gaining widespread acceptance as a measure of intrinsic motivation in the context of sport and exercise. This development has been largely the result of McAuley, Duncan, and Tammen's (1989) and McAuley, Wraith, and Duncan's (1991) research describing the psychometric properties and use of IMI to determine levels of intrinsic motivation for a basketball shooting task and participation in aerobic dance, respectively. According to McAuley et al. (1989) and McAuley et al. (1991), the IMI determines an individual's level of intrinsic motivation as an addeditive function of four underlying dimensions: perceived competence, interest-enjoyment, pressure-tension, and effort-importance. In the aerobics version of the instrument, McAuley et al. (1991) included a fifth dimension: perceived choice. This was presumably held to reflect perceived locus of causality, a core construct of cognitive evaluation theory (CET; Deci & Ryan, 1985), which would seem to form the theoretical grounding for the instrument. McAuley et al. (1989) and McAuley et al. (1991) stated that the IMI possesses a number of unique psychometric properties. The full set of 27 items has been used rarely, and inclusion or exclusion of any one factor does not affect adversely the properties of the remaining factors. In addition, the subscales can be shortened by eliminating redundant items without compromising their reliability. Finally, the generic scale items can easily be modified to reflect intrinsic motivation for any particular activity.

Ryan (1982) is usually cited as having originated the IMI, with Ryan, Mims, and Koestner (1983) and Plant and Ryan (1985) also credited with some input into its development (e.g. Duda, 1992; McAuley et al., 1989, 1991; McAuley & Tammen, 1989; Whitehead & Corbin, 1991a, 1991b). In fact, Ryan (1982) merely briefly describes the use of a set of ad hoc scales to assess levels of interest and enjoyment, tension and pressure, effort, and importance, with respect to a puzzle-solving task. The number of items and the psychometric properties of the scales are not discussed. Ryan et al. (1983) also report the use of ad hoc scales comprising 26 items to measure interest and enjoyment, tension and pressure, effort, and importance. While these items are presumably the same as those used by Ryan (1982), no reference to their earlier use is given. Ryan et al. (1983) did, however, perform

Submitted: September 22, 1994
Accepted: August 23, 1996

David Markland and Lew Hardy are with the School of Sport, Health, and Physical Education Sciences at the University of Wales–Bangor.

a factor analysis of the items. Two main factors emerged from a principal components analysis with varimax rotation. The first comprised 11 items relating to interest, enjoyment, and attention (although the latter are not mentioned in the procedural section). The second factor comprised three items relating to pressure and tension. Plant and Ryan (1985) also discussed a factor analysis of a 17-item subset of the same set of items. In this study, three factors emerged: 11 items reflecting interest-enjoyment, 3 items relating to pressure-tension, and 2 items reflecting effort-importance. One item was eliminated. Neither Ryan (1982), Ryan et al. (1983), nor Plant and Ryan (1985) described these scales as the Intrinsic Motivation Inventory. Ryan (1982) mentioned a perceived competence scale, but only in the results section. Ryan et al. (1983) and Plant and Ryan (1985) did not mention a perceived competence scale at all.

The origins of the IMI, then, are somewhat shrouded in mystery. The real story starts with the two reports by McAuley and his colleagues: McAuley et al. (1989), and McAuley et al. (1991). In the 1989 paper, results were presented from a series of confirmatory factor analyses using LISREL, which examined the psychometric properties of a basketball-shooting version of the IMI. The primary purpose of their study was to test a hypothesized five-factor hierarchical model with four first order factors (perceived competence, interest-enjoyment, pressure-tension, and effort-importance) and a second-order general factor representing intrinsic motivation. This hierarchical model was attributed to Ryan (1982), which is rather odd given that he mentioned nothing about the factor structure of the scales he used. McAuley et al. (1989) explored both 18- and 16-item versions of the IMI, testing the hierarchical model against a four-factor (dimensions only) first-order model, a single-factor model, and a null or completely independent model. Results for the 16-item version were interpreted as supporting the hierarchical model, $\chi^2_{101} = 252.36$, $\chi^2/df = 2.499$, Goodness of Fit Index (GFI) = .788, Root Mean Square Residual (RMSR) = .136, Normed Fit Index (NFI) = .76. Similarly, McAuley et al. (1991) reported support for the hierarchical model using an aerobic dance version of the IMI, $\chi^2_{184} = 428.61$, $\chi^2/df = 2.329$, Adjusted Goodness of Fit Index (AGFI) = .824, GFI not reported, RMSR = .075, NFI = .87. In addition to the four first-order dimensions discussed above, the perceived choice dimension was also included in this model. However, McAuley et al. (1991) reported that this subscale was relatively low in reliability and that the relationship between perceived choice and overall intrinsic motivation was explained largely by a single item.

A major problem faced by researchers using confirmatory factor analysis is that there are no universally accepted criteria for what constitutes a good fit (Tanaka, 1993; Loehlin, 1992). Thus, there is room for argument in interpreting the results of a LISREL analysis. It is clear, however, that by conventional standards the results reported by McAuley et al. (1989) and McAuley et al. (1991) are far from optimal, and perhaps the best that can be said about them is that they do not contradict or support unequivocally the proposed hierarchical model.

In addition to possible statistical interpretation problems, however, there is also a theoretical issue with respect to the hierarchical model of McAuley et al. In this model the IMI subscales are considered to be indicators of intrinsic motivation and held to be located at the same conceptual level. If one examines the formal propositions of CET, though (on which the IMI is presumably based), the theory appears to propose that changes in perceived locus of causality and perceived competence have a *causal* influence on intrinsic motivation. If this is the case, then a more accurate model of intrinsic motivation would place perceived locus of causality and perceived competence at a different conceptual level than the indicators of intrinsic motivation.

In fact, the position is not so clear, because Deci and Ryan (1985, p. 65) also argue that changes in perceived locus of causality and perceived competence are concomitant with, rather than causes of, motivational changes. This seems to be at odds with the causal implications of the CET propositions and the way in which the relationships between perceived locus of causality, perceived competence, and intrinsic motivation are generally characterized. For example, in a study of the effects of positive and negative feedback on intrinsic motivation for a stabilometer task, Vallerand and Reid (1984) presented a causal model using path analysis in which the type of feedback influenced intrinsic motivation through the mediating effect of perceived competence. Thus the perceived competence and intrinsic motivation constructs were located at different levels in the model. Moreover, Vallerand and Reid (1984) titled their paper, "On the Causal Effects of Perceived Competence on Intrinsic Motivation: A Test of Cognitive Evaluation Theory," which seems to be an unambiguous statement about the way in which these authors saw the constructs as being related.

In 1987, Deci warned that perceived locus of causality and perceived competence (which he described as needs) are quite different constructs from the emotional (and, one might add, the behavioral) responses that characterize intrinsic motivation. He stated that:

> It is interesting and important to study the emotions that accompany intrinsically motivated behavior, but it is especially important to keep clear the difference between needs and emotions *and to place them in appropriate theoretical relationship to each other* [Deci, 1987, p. 181, italics added].

These are clearly issues of fundamental theoretical importance. If perceived locus of causality, perceived competence, interest-enjoyment, pressure-tension, and effort-importance are indeed concomitants, then it makes sense to hold these constructs at the same level. The correct model, then, would be one in which the five constructs are allowed to covary (an oblique confirmatory factor analytic model). However, if perceived locus of causality and perceived competence do have causal influences on intrinsic motivation, then the constructs should not be situated at the same conceptual level.

These issues were explored within the context of exercise participation in a series of studies designed to assess a new measure of perceived locus of causality (Studies 1 and 2) and examine and compare a model in which the IMI constructs are held to be concomitants (an oblique confirmatory factor analytic model) with structural models in which perceived competence and perceived locus of causality have causal effects on interest-enjoyment, pressure-tension, and effort-importance (Study 3).

Study 1: Locus of Causality for Exercise Scale

Much of the research exploring CET in the context of sport and exercise has focused exclusively on the influence of perceived competence on intrinsic motivation. This is unfortunate, given that the perceived locus of causality construct plays a pivotal role in the theory. Because the present study involved an examination of the influences of both perceived locus of causality and perceived competence on intrinsic motivation, a measure of perceived locus of causality was an essential pre-requisite.

A number of authors have implied that perceived locus of causality is synonymous with Rotter's (1966) locus of control construct (e.g. Duda & Tappe, 1989a, p. 229; Duda & Tappe, 1989b, pp. 241, 243–244; Weiss & Chaumeton, 1992, p. 69). However, deCharms (1981), who can be credited as being the first to explore comprehensively the construct (although it was introduced to the literature by Heider in 1958), pointed out that perceived locus of causality and locus of control are quite different and have arisen from two different research traditions. The locus of control construct developed from a social-learning and expectancy-value context and refers to beliefs about control over reinforcement contingencies. Locus of causality, in contrast, grew from an explicitly motivational framework and an attempt to understand the energization of behavior. Thus, while locus of control is concerned with control over outcomes, locus of causality is concerned with the source of the initiation of behavior. Deci and Ryan (1985, pp. 166–168; 1990, pp. 248–249) also emphasize this point.

A central feature of self-determined behavior is the perception of choice (Deci & Ryan, 1985). As mentioned previously, McAuley et al. (1991) included a perceived choice subscale in their analyses of the aerobics version of the IMI but warned that its reliability was suspect. Furthermore, a close examination of the perceived choice items suggests that they may not fully capture the essence of the perceived locus of causality construct. According to Deci and Ryan (1985), an external locus of causality exists when a behavior is perceived as being initiated or regulated by a controlling event. Controlling events are defined as those experienced as pressure to think, feel, or behave in particular ways. In addition, it is argued that controlling events may be external to the individual, such as rewards, commands, or other situational influences, or they may be internal to the individual, such as needs, goals, and expectations. A distinction is drawn between internally controlling regulation of behavior (in which the perceived locus of causality is external) and internally informational regulation of behavior (in which the perceived locus of causality is internal). To the present authors, it seemed that at least two of the perceived choice items in the IMI might reflect rather extreme perceptions of external pressures to exercise (Item 9: I haven't really had a choice about participating in this aerobics class. Item 21: I participate in this aerobics class because I have no other choice). It is arguable, however, that in the exercise domain intrapersonal events such as the exercise goals that individuals adopt, which may be internally controlling or internally informational (Ryan, Vallerand, & Deci, 1984), will be a more central determinant of perceived locus of causality than external events (Thill & Brunel, 1995). For these reasons, the perceived choice subscale of the IMI was not considered appropriate as a measure of perceived locus of causality.

Several instruments have been developed which directly or indirectly assess perceived locus of causality. The Origin Measure (deCharms, 1981) is based on content analysis of short stories produced by participants in response to imaginary scenarios. However, scoring this instrument is complex, making its use with large samples difficult. A number of other instruments have also been developed based on a continuum conception of intrinsic and extrinsic motivation (e.g. Pelletier et al., 1995; Ryan & Connell, 1989; Vallerand & O'Connor, 1991; Vallerand et al., 1992). All these instruments are context-specific and could not be applied readily to the exercise domain without substantial revisions and revalidation. Consequently, the present authors decided to develop a new scale for measuring perceived locus of causality for exercise which could be administered easily and would sit comfortably with the IMI subscales.

Item Generation

Based on Deci and Ryan's (1985) description of the perceived locus of causality construct, and, in particular, on the concept of internally controlling events, nine items were generated which at face value reflected whether individuals exercise from choice or because they feel they *have to* exercise for some reason. A Likert-type scale, with scores ranging from 1 to 6, was used to register responses to items. Two items were worded positively, and the remainder were worded negatively. Items were recoded prior to the analyses such that high scores indicated a more internal, perceived locus of causality, and low scores indicated a more external locus.

Participants

Four hundred individuals from various backgrounds and age groups, all who were involved in some form of regular physical activity, were asked to complete the instrument. Of those, 241 (60.25%) agreed to participate. The participants comprised three fairly distinct subgroups: female aerobics and keep-fit class participants from the local community (n = 110, *M* age = 35.96 years, *SD* = 14.93); physical education students (n = 89, *M* age = 21.25 years, *SD* = 3.63, 34.83% women, 65.17% men); and members of local recreational sports clubs (n = 42, *M* age = 24.40 years, *SD* = 11.05, 19.10% women, 80.90% men).

Procedures

With the permission of class instructors, community aerobics and sports club participants were approached at the end of an exercise or sport session and asked to complete a questionnaire concerning their feelings about exercise. All participants gave informed consent to participate in the study. Participants were asked to take the questionnaires home and to return completed copies at their next exercise or sport session. The physical education students also gave their informed consent, and completed the questionnaire in a classroom environment. Participation in the study was not a course requirement, and no course credit was given for taking part. No measure of the amounts of exercise undertaken by the participants was obtained.

Analysis and Results

To reduce the items to an internally and factorially consistent set, the nine items were subjected to a principal components analysis with varimax rotation. Three factors emerged with eigenvalues greater than one, accounting for 61.4% of the variance. Table 1 shows the loadings, eigenvalues, and variance explained for the rotated factors. Each factor comprised three items. With the exception of two ambiguously loading items, the rotated factors appeared to be highly distinct statistically. In terms of item content, however, no consistent conceptual distinctions between the three factors were apparent. Cronbach's alpha reliability coefficients for the three factors were as follows: Factor 1, .738; Factor 2, .614; Factor 3, .426. Separate factor analyses were also performed for women (n = 149) and men (n = 92). In both these analyses, three factors emerged once again and were similar to those that emerged from the total sample, except Item 5 loaded on Factor 1 among women and Item 6 loaded on Factor 1 among men. Both these items also crossloaded strongly on other factors.

Because the aim of this study was to develop an easily administered instrument similar in format to the IMI subscales, it was not considered appropriate to further explore the multidimensional factor structure which had emerged. Instead, given that the first factor had the most acceptable reliability, its items appeared to have good face validity for the construct of interest, and its items consistently formed the core of Factor 1 in the three separate analyses, it was decided to proceed under the assumption that this might be a suitable measure of perceived locus of causality, and the second and third factors were dropped from further analyses. The remaining three items read as follows: Item 1: I exercise because I like to rather than because I feel I have to; Item 2: Exercising is not something I would necessarily choose to do, rather it is something that I feel I ought to do; Item 3: Having to exercise is a bit of a bind, but it has to be done. The scale was now labeled the Locus of Causality for Exercise (LCE) scale.

Table 1. Rotated factor loadings, eigenvalues, and percentage of variance explained for the exploratory factor analysis of the locus of causality items

Item	Factor 1	Factor 2	Factor 3
4	.843	.093	.156
3	.810	.168	-.002
2	.690	-.122	.286
7	.167	.783	-.049
1	-.092	.739	.062
5	.511	.550	.122
9	.123	.172	.769
8	.077	.172	.769
6	.263	.399	.572
Eigenvalue	2.267	1.732	1.636
Variance explained (%)	25.18	19.24	18.18

Study 2: Confirmatory Factor Analysis

The sample used for the exploratory factor analysis was clearly heterogeneous in terms of participants' age and gender and the types and, presumably, amounts of exercise they undertook. The LCE measure, thus, could be confounded by these variables. Therefore, the authors of this study sought confirmation of the factorial validity of the scale by administering the LCE to a second, more homogeneous sample and subjecting the data to confirmatory factor analysis.

Participants

The LCE was completed by a second sample of 169 female community aerobics and keep-fit class participants (*M* age = 34.41 years, *SD* = 12.76). All reported participating in aerobics or keep-fit classes at least once a week for at least 1 hour per week.

Procedure

After obtaining permission from class instructors, potential participants were approached at the end of their exercise classes and asked to take part in a study of how people feel about exercising. Informed consent was given by all those who agreed to participate. Before leaving the facilities, participants then completed the LCE along with Markland and Hardy's (1993) Exercise Motivations Inventory, an instrument that assesses individuals' reasons for exercising. Results pertaining to this instrument are not reported in this study.

Analysis

The data were analyzed using LISREL 8.03 (Jöreskog & Sörbom, 1993a). The variance-covariance matrix was computed and used as data input, and maximum likelihood estimation was employed[1]. Prior analyses revealed that the distribution of scores did not depart markedly from a normal distribution[2]. A problem with testing a single factor model with three observed indicators is that it is exactly identified. Consequently, to test whether the three items were loaded on a single factor and were equally good indicators of that factor, the factor loadings were constrained to be equal. Goodness of fit was assessed by examining the χ^2 goodness of fit test, the Comparative Fit Index (CFI; Bentler, 1990), the Non-Normed Fit Index (NNFI; Tucker & Lewis, 1973), the Incremental Fit Index (IFI; Bollen, 1989), and the Root Mean Square Error of Approximation (RMSEA; Steiger, 1990). The CFI, NNFI, and IFI are all members of the class of incremental fit indexes that compare the fit of a restricted model to a baseline model, usually the null model (Jöreskog and Sörbom, 1993b). For these indexes, minimum values of .90 generally are considered to represent an acceptable fit. RMSEA is a measure based on the extent to which a model holds approximately in the population, expressed as the discrepancy per degree of freedom (Browne & Cudeck, 1993). According to Browne and Cudeck (1993), a value of .05 generally indicates a close fit, and values of up to .08 represent a reasonable error of approximation. LISREL 8.03 (Jöreskog & Sörbom, 1993a) gives a test of RMSEA < .05 and, if the p value for χ^2 is not too small, computes a 90% confidence interval for RMSEA.

Results

Table 2 shows the fit indexes for the model, as well as Cronbach's alpha reliability coefficient for the scale. As can be seen, the results are encouraging, with a nonsignificant χ^2 and values for the incremental fit indexes and RMSEA suggesting an exact fit. The test of close fit indicates that RMSEA was not significantly greater than .05, and the lower limit of the confidence interval is zero, suggesting that the null hypothesis of an exact fit cannot be rejected, although the nonzero upper limit of the confidence interval indicates that the fit is not perfect. Finally, the factor loadings of the items were all statistically significant (completely standardized estimates: Item 1 = .786; Item 2 = .753; Item 3 = .787; $ps < .001$), and Cronbach's alpha for the scale was improved over the initial exploratory analysis. While caution is warranted in interpreting results for a model with so few parameters, these findings provide support for the factorial validity of the LCE in this sample.

Table 2. Fit indexes and Cronbach's alpha for the confirmatory factor analysis of the LCE

χ^2	df	p	CFI	NNFI	IFI	RMSEA	p	90% CI	Alpha
1.640	2	.440	1.00	1.00	1.00	.000	.575	.000–.144	.826

Note. CFI = Comparative Fit Index; NNFI = Non-Normed Fit Index; IFI = Incremental Fit Index; RMSEA = Root Mean Square Error of Approximation; CI = 90% confidence interval for RMSEA.

Study Three: Structural Models

Having developed a measure of perceived locus of causality (although recognizing that only evidence for its factorial and not its construct validity has been presented), the research now moved on to examine the issues raised earlier concerning the causal effects of perceived competence and perceived locus of causality on intrinsic motivation. In this study an oblique confirmatory factor analytic model (M1: the concomitant model, see Figure 1) was compared with structural models in which perceived locus of causality and perceived competence have causal effects on the intrinsic motivation constructs. There are at least two ways in which the causal relationships between perceived locus of causality and perceived competence and intrinsic motivation could be usefully conceptualized. In the first, taking interest-enjoyment, pressure-tension, and effort-importance as indicators of intrinsic motivation, the two constructs would have direct and independent effects on the intrinsic motivation constructs (M2: the direct causal model, see Figure 2). However, according to CET (Deci & Ryan, 1985, p. 63; 1990, p. 269), increases in perceived competence will only enhance intrinsic motivation in the context of an internal perceived locus of causality. Fisher (1978) found support for this proposition in a study of the relationships between perceived competence and intrinsic motivation in which performance was perceived as either self-determined or constrained. Analyses revealed a strong positive correlation between perceived competence and intrinsic motivation when participants felt that their behavior was self-determined, but no correlation when the behavior was not experienced as self-determined. Thus, a second way to conceptualize the causal relationships between perceived locus of causality, perceived competence, and intrinsic motivation would be to posit interactive effects of perceived locus of causality and perceived competence on the intrinsic motivation constructs. Unfortunately, it is not straightforward to address interactions directly in structural equation models, at least in single-sample analyses. Consequently, in this study the interactional effects were approximated by testing a model in which perceived locus of causality mediated the effects of perceived competence on interest-enjoyment, pressure-tension, and effort-importance (M3: the mediational model, see Figure 3). This gave three models ranging from the least constrained (M1) to the most constrained (M3). Because one can always improve the fit of a model by allowing more free parameters, less constrained models are likely to evidence a better fit (Mulaik et al., 1989). However, given an acceptable fit for the more constrained models, one can argue that they are preferable on the grounds of parsimony or in terms of substantive theoretical concerns. Furthermore, it was hypothesized that if the mediational interpretation was an appropriate way of modeling the relationships between the constructs, the direct effects of perceived competence on the intrinsic

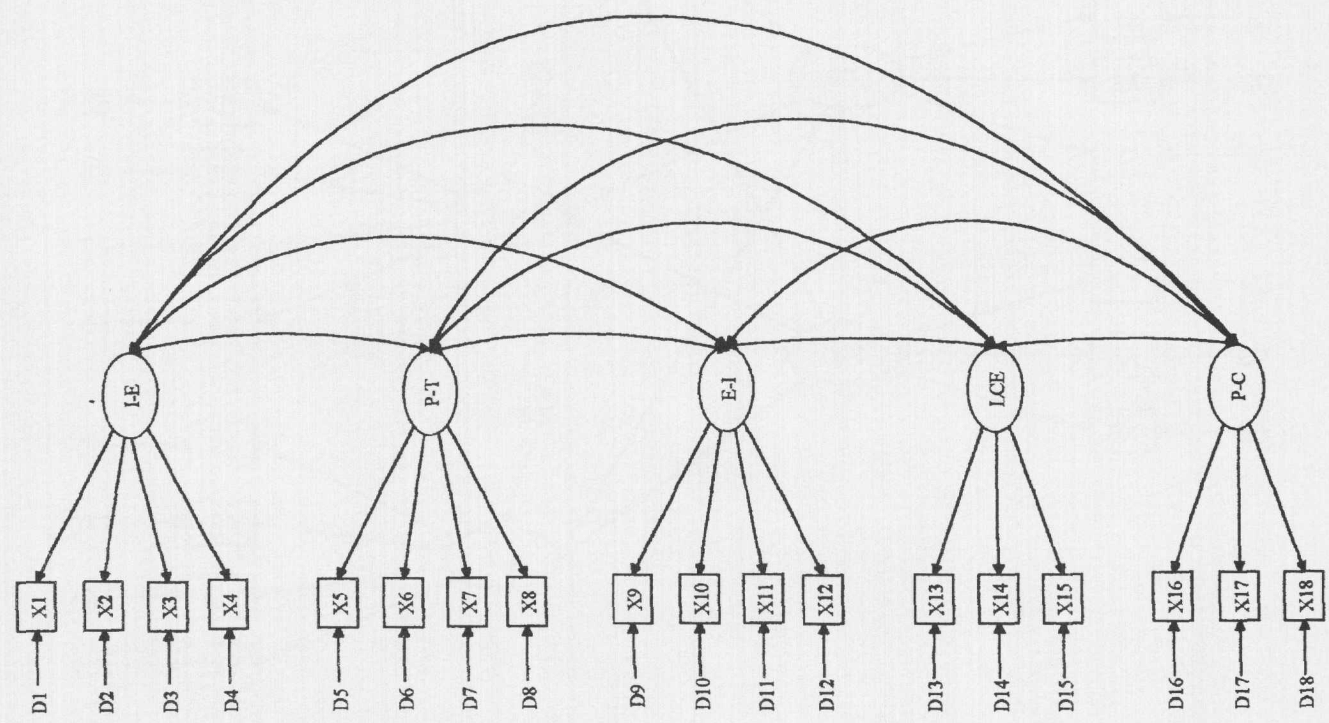

Figure 1. Model 1 (M1): concomitant model. I-E = interest-enjoyment; P-T = pressure-tension; E-I = effort-importance; LCE = locus of causality; P-C = perceived competence; D = measurement errors for observed variables.

motivation constructs in M2 would be weak compared to the indirect effects through perceived locus of causality in M3.

Participants

Data were collected from a new sample of 133 female aerobics participants from four different community aerobics classes (*M* age = 35.21 years, *SD* = 10.94). All participants reported that they participated in aerobics at least once per week.

Procedure

With the permission of class instructors, a request was made at the end of the aerobics sessions for participants to take part in a study examining how people feel about exercising. The exercisers who agreed to take part gave informed consent and then completed the data collection requirements.

Instruments

The participants completed the LCE along with a reduced, 15-item version of the aerobics form of the IMI developed by McAuley et al. (1991). The perceived choice items were eliminated entirely for the reasons outlined earlier. In addition, three items were removed to reduce the interest-enjoyment subscale from seven to four items (Items 14, 15, and 20). A previous, unreported factor analysis of the aerobics version of the IMI conducted by Markland and using a similar sample had found these items to have the lowest factor loadings on their latent variable. As noted earlier, McAuley et al. (1989) and McAuley et al. (1991) suggested that redundant items could be eliminated from the IMI subscales without adversely affecting their reliability. The scales administered comprised four items for each of the interest-enjoyment, pressure-tension, and effort-importance subscales and three items for the perceived competence subscale.

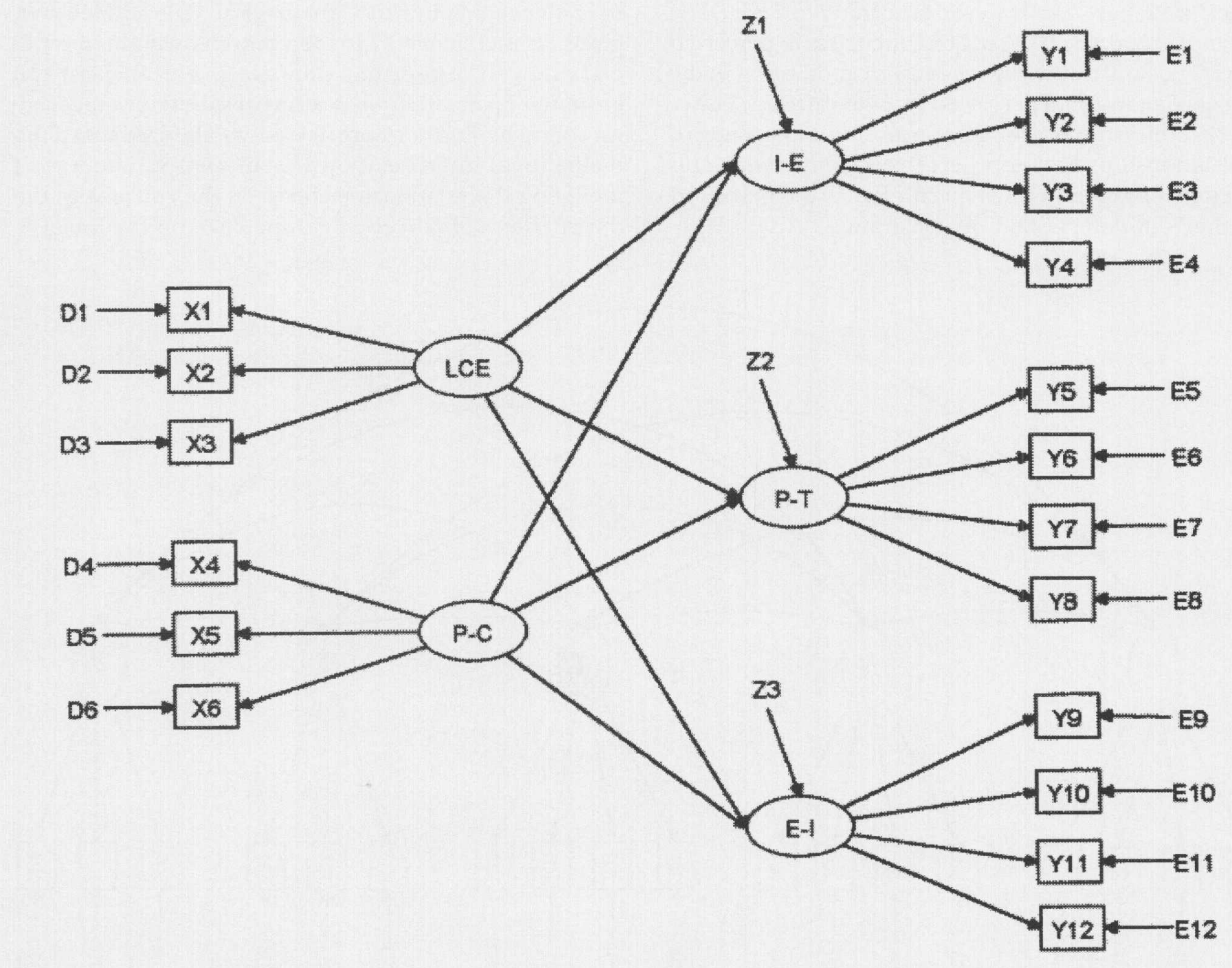

Figure 2. Model 2 (M2): direct causal model. I-E = interest-enjoyment; P-T = pressure-tension; E-I = effort-importance; LCE = locus of causality; P-C = perceived competence. D = measurement errors for exogenous observed variables. E = measurement errors for endogenous observed variables; Z = disturbance terms for endogenous latent variables.

Analyses

The data were again analyzed using LISREL 8.03 (Jöreskog & Sörbom, 1993a). The variance-covariance matrix was used as data input, and maximum likelihood estimation was employed[3]. Preliminary analyses revealed that the scores on the items were approximately univariately normally distributed, but there was evidence of multivariate skewness and kurtosis[4]. At this stage, in addition to the fit indexes reported for the previous analyses, the Parsimony Normed Fit Index (PNFI; James, Mulaik, & Brett, 1982) and Bozdogan's (1987) Consistent Akaike's Information Criterion (CAIC) were examined. These measures take into account both parsimony (in terms of the number of parameters in a model) and goodness of fit, thereby avoiding the problem of improving fit at the expense of reduced degrees of freedom. The model with the minimum value for CAIC or maximum value for PNFI is taken to be potentially the most useful (Bentler, 1992; James et al., 1982).

Results

Table 3 shows the means, standard deviations and Cronbach's alpha reliability coefficients for the LCE and IMI subscales. Table 4 shows the factor loadings, their *t* values, and factor correlations for the concomitant model. The factor loadings were all moderate to strong and statistically significant. Table 5 shows the fit indexes for the structural and confirmatory factor analytic models. In all cases, the fit was not ideal, with χ^2 being significant and relatively large compared to the degrees of freedom, RMSEA just outside the normally acceptable bounds (Browne & Cudeck, 1993), and the incremental fit indexes indicating some degree of misfit in the models. The indexes for the concomitant model (M1) were slightly better than for the structural models, but as mentioned previously, this is to be expected as it is less constrained. The fit for M2 was particularly poor, despite its being a less constrained model than M3. However, CAIC was lowest and PNFI highest for M3, with M2 having the highest value for CAIC and the lowest for PNFI. Thus, among these models the mediational model gives the most parsimonious explanation of the relationships between the constructs. Table 6 shows the completely standardized parameter estimates, together with their standard errors and *t* values, and disturbance terms for the endogenous constructs for M2 and M3. The paths are all moderate to strong and statistically significant for M3. For the direct causal model (M2), however, the paths from perceived competence to the intrinsic motivation constructs are weak, albeit significant. The

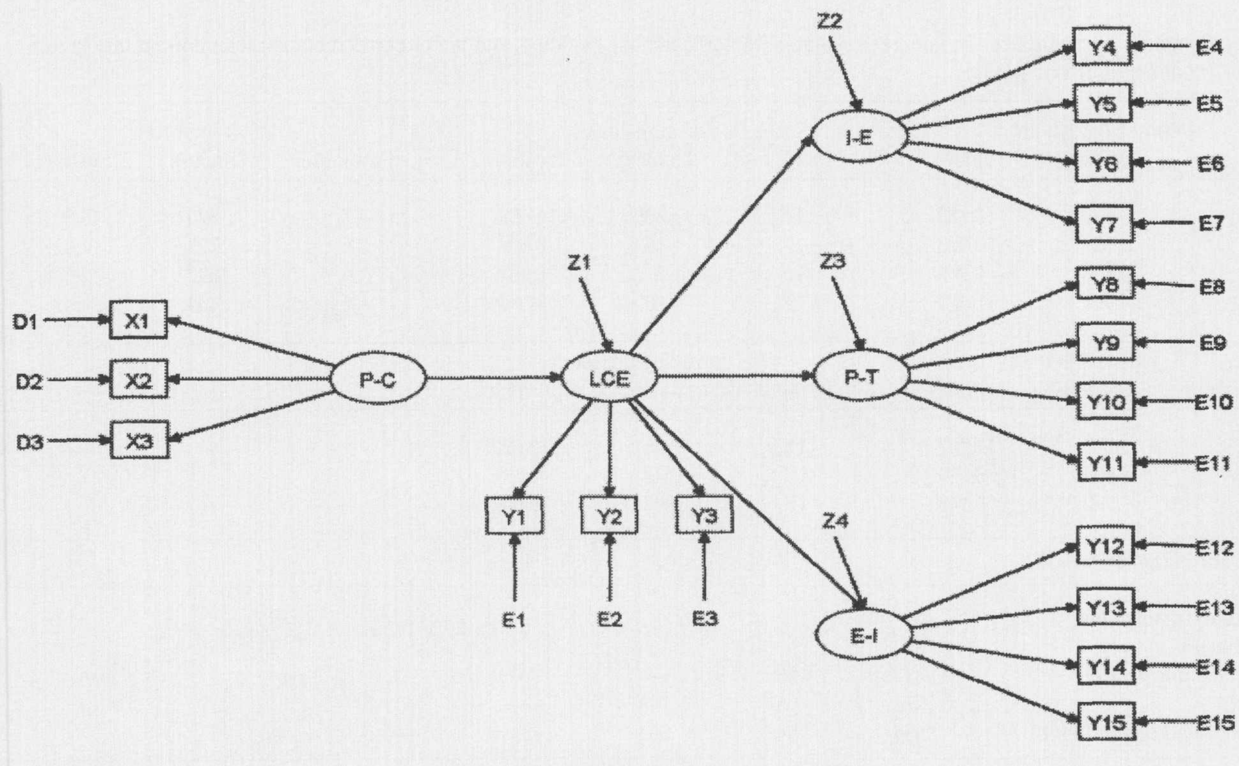

Figure 3. Model 3 (M3): mediational model. I-E = interest-enjoyment; P-T = pressure-tension; E-I = effort-importance; LCE = locus of causality; P-C = perceived competence. D = measurement errors for exogenous observed variables. E = measurement errors for endogenous observed variables; Z = disturbance terms for endogenous latent variables.

loadings of the observed variables on their factors in the structural models are not reported here because they did not differ substantially from those reported for the confirmatory factor analytic model.

Discussion

This series of studies was designed to examine a causal model of intrinsic motivation which placed the constructs embedded in the Intrinsic Motivation Inventory (IMI) in an appropriate theoretical relationship to one another as indicated by the tenets of cognitive evaluation theory. The first task was to develop an instrument to assess perceived locus of causality, because it is a central construct in the theory. Item reduction procedures led to a set of internally consistent, factorially valid items with high face validity, which was simple to administer and sat easily with the IMI subscales.

In the next stage, two structural models, in which perceived competence and perceived locus of causality had causal influences on interest-enjoyment, pressure-tension, and effort-importance, were compared with a model in which perceived competence and perceived locus of causality were concomitant with the other intrinsic motivation constructs. Although the overall fit of the models was not particularly good, the results were consistent with a causal explanation of the relationships between the constructs. The fit indexes for the concomitant model were marginally better than those of the mediational model, but the fit of the mediational model was clearly better than that of the direct causal model. Furthermore, when parsimony was taken into account, the mediational model was superior to both the concomitant model and the direct causal model. In addition, and as predicted, in the direct causal model the paths from perceived competence to the intrinsic moti-

Table 3. Means, standard deviations, and Cronbach's alphas for the LCE and IMI subscales

Scale	*M*	*SD*	Alpha
LCE	4.068	1.416	.815
Perceived competence	4.153	1.344	.805
Interest-enjoyment	5.694	1.398	.906
Pressure-tension	5.402	1.414	.817
Effort-importance	5.399	1.180	.721

Note. LCE = Locus of Causality for Exercise Scale; IMI = Intrinsic Motivation Inventory.

Table 4. Completely standardized parameter estimates (factor loadings), *t* values[a], and factor correlations for the confirmatory factor analytic (concomitant) model (M1)

Interest-enjoyment			Pressure-tension			Effort-importance		
Parameter	Loading	*t* value	Parameter	Loading	*t* value	Parameter	Loading	*t* value
$LX_{1,1}$	.922	13.783	$LX_{5,2}$	.839	11.438	$LX_{9,3}$	.820	10.791
$LX_{2,1}$	.929	13.973	$LX_{6,2}$	.562	6.697	$LX_{10,3}$	.628	7.622
$LX_{3,1}$	.650	8.242	$LX_{7,2}$	.817	10.991	$LX_{11,3}$	.620	7.508
$LX_{4,1}$	.867	12.421	$LX_{8,2}$	.685	8.577	$LX_{12,3}$	.408	4.631

LCE			Perceived competence		
Parameter	Loading	*t* value	Parameter	Loading	*t* value
$LX_{13,4}$	.867	11.702	$LX_{16,5}$	.925	12.620
$LX_{14,4}$	.694	8.626	$LX_{17,5}$	.548	6.481
$LX_{15,4}$	.762	9.148	$LX_{18,5}$	.809	10.517

Factor correlations	1	2	3	4	5
1 Interest-enjoyment	—				
2 Pressure-tension	.830	—			
3 Effort-importance	.877	.860	—		
4 LCE	.776	.781	.749	—	
5 Perceived competence	.565	.548	.681	.623	—

Note. LX = Lambda-X, the loadings of the observed variables on their latent variables; LCE = Locus of Causality for Exercise Scale. All *t* values and factor correlations significant ($p < .001$).
[a]LISREL 8.03 does not provide standard errors and *t* values for the completely standardized solution. These *t* values are derived from the unstandardized solution.

vation constructs were weak, whereas in the mediational model all the structural paths were moderate to strong. It should be noted that in the direct causal model the relationship between perceived locus of causality and perceived competence was constrained to zero. Ideally, this parameter would be free to be estimated. However, doing this led to an improper solution in that the completely standardized parameter estimate for the path from LCE to pressure-tension was greater than 1. As Marsh (1993) pointed out, a logical prerequisite to evaluating the goodness of fit of a model is that the empirical solution is proper. Constraining the relationship between the exogenous variables solved this problem.

These results support the causal interpretation of CET outlined earlier and suggest it is meaningful to hold perceived locus of causality, perceived competence, and the indicators of intrinsic motivation at different conceptual levels. Furthermore, although the mediational model only approximated the interactive effects of perceived locus of causality and perceived competence, the comparison between the structural models supports the contention drawn from CET that perceived competence will only influence intrinsic motivation within the context of some self-determination. After all, it is reasonable to assume that one can feel competent at an activity without necessarily feeling intrinsically motivated to engage in it.

It is important, however, to bear in mind the limitations of the study. First, the sample size for the final analyses was relatively small. According to Baldwin

Table 5. Fit indexes for the confirmatory factor analytic (concomitant) and structural models

	χ^2	df	*p*	CFI	NNFI	IFI	RMSEA	*p*	CAIC	PNFI
Concomitant model										
M1	254.671	125	< .001	.911	.891	.913	.089	< .001	525.627	.688
Structural models										
M2	332.907	129	< .001	.861	.835	.863	.109	< .001	580.302	.669
M3	287.881	131	< .001	.893	.875	.894	.095	< .001	523.495	.704

Note. CFI = Comparative Fit Index; NNFI = Non-Normed Fit Index; IFI = Incremental Fit Index; RMSEA = Root Mean Square Error of Approximation; CAIC = Consistent Akaike's Information Criterion; PNFI = Parsimony Normed Fit Index. (90% CI for RMSEA not included due to too small *p* value for χ^2.)

Table 6. Structural models: Parameter estimates (labeled paths: completely standardized solution) together with standard errors and *t* values[a], and disturbance terms for endogenous latent variables

	Perceived competence				LCE			Disturbance terms	
	Path	Error	*t* value		Path	Error	*t* value		
M2									
I-E	.172	.066	2.620	I-E	.865	.080	10.879	I-E	.221
P-T	.148	.073	2.017	P-T	.875	.094	9.353	P-T	.212
E-I	.353	.085	4.148	E-I	.836	.099	8.435	E-I	.177
M3									
LCE	.660	.097	6.826	I-E	.897	.085	10.607	I-E	.195
				P-T	.898	.096	9.332	P-T	.194
				E-I	.930	.103	9.028	E-I	.135
								LCE	.565

Note. I-E = Interest-enjoyment; P-T = Pressure-tension; E-I = Effort-importance. [a]LISREL 8.03 does not provide standard errors and *t* values for the completely standardized solution. These standard errors are rescaled by dividing the completely standardized parameter estimates by their *t* values derived from the unstandardized solution (Marsh, 1993).

(1989) and Bearden, Sharma, and Teal (1982), a minimum of 200 participants is required to produce stable results. Bentler (1992) recommended a 5:1 ratio of sample size to the number of parameters to be estimated, while Anderson and Gerbing (1988) recommended at least 150 participants. Clearly, the study would have benefited from a larger sample. In addition, there was evidence that the assumption of multivariate normality for maximum likelihood estimation was not upheld. This could bias downwardly the standard errors leading to an inflated number of statistically significant parameters (Byrne, 1994; Muthén & Kaplan, 1985). Due to the small sample size, it was not possible to use an asymptotic distribution free method of estimation.

A further limitation of the study is that it was cross-sectional. According to CET, it is *changes* in perceived locus of causality and perceived competence that influence intrinsic motivation, rather than their absolute levels at any given time. While the results are consistent with a causal interpretation, it would be necessary to collect data longitudinally to demonstrate these causal influences more convincingly. Finally, no evidence for the external validity of the LCE has been presented, although the strength of the relationships between it and the indicators of intrinsic motivation in this study suggests that it holds promise. However, neither has the external validity of the IMI been established firmly in previous work using the instrument. Indeed, McAuley et al. (1989), while expressing optimism about the ability of the IMI to assess intrinsic motivation, indicate that further work is required to validate the measure.

Deci (1987) and Deci and Ryan (1990, p. 263) have pointed out particular difficulties in the operationalization of intrinsic motivation. Deci (1987) warned of the dangers inherent in assuming that measures of persistence and self-reports of interest-enjoyment actually reflect intrinsic rather than extrinsic motivation. As Deci (p. 182) put it:

> There is simply no question that free-choice persistence and interest/enjoyment can reflect more than one source of internal motivation.

He went on to say (p. 183–184):

> Depending on the way questions are worded, participants may actually be answering different questions from those the investigator intended to ask. For example, imagine people who enjoyed doing a puzzle *because* they got paid for doing it. It was the payment they enjoyed, not the puzzle activity itself.

Thus, if self-reported enjoyment is used to infer intrinsic motivation, participants may appear to be motivated intrinsically, when, in fact, the motivation is extrinsic. In the current context, exercisers might report enjoying physical activity, not for the intrinsic rewards that stem from participation per se, but because they enjoy the extrinsic rewards associated with participation, such as weight loss, improvement of physical appearance, and social recognition (Ryan et al., 1984). Therefore, the IMI could fail to distinguish between the directional dimension of motivation (intrinsic or extrinsic) while still assessing the intensity of motivation. The important issue is what the items mean to the individual completing the questionnaire. From a conceptual perspective, the effort-importance subscale seems particularly troublesome, because it is obvious that people will very often put a great deal of effort into an activity to obtain extrinsic rewards. Therefore, effort, either subjectively or objectively operationalized, would not seem to be an unambiguous indicator of intrinsic motivation.

The present authors' concern is that the IMI in its present form is becoming accepted by researchers in sport and exercise psychology as a valid and reliable instrument, when, considering its theoretical grounding, the evidence for this is not strong. The authors contention is that a question mark remains over the psychometric integrity of the IMI as it is currently used. In particular, it would seem that the practice of summing scores on the IMI subscales to arrive at an overall indication of levels of intrinsic motivation fails to take account of the theoretical distinction between the antecedents of intrinsic motivation and its indicators.

In conclusion, this study has highlighted potential conceptual and operational weaknesses of the IMI as currently used. It also serves to point out the difficulties inherent in operationalizing intrinsic motivation and, in particular, the dangers of adopting operational definitions of constructs without ensuring that they adequately reflect the theory from which they are drawn. It is recommended that, before the IMI becomes firmly established as the instrument to use when assessing intrinsic motivation, further efforts should be paid to establishing its construct validity (and indeed that of the LCE). Moreover, if theory is to be advanced in this area, researchers should concentrate on building and testing conceptually sound models. In the current context, this means examining further the causal influences of perceived competence and perceived locus of causality on intrinsic motivation. As Deci (1987) argued, a preoccupation with simply predicting variance in operations rather than with the underlying theoretical issues is likely to lead eventually to confusion rather than clarity.

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. 231–242). New York, Sport Science Associates.

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

Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen & J. S. Long (Eds.) *Testing structural equation models.* Newbury Park, CA: Sage.

Thill, E. E. & Brunel, P. (1995). Cognitive theories of motivation in sport. In S. J. H. Biddle (Ed.) *European perspectives on exercise and sport psychology* (pp. 195–217). Leeds, UK: Human Kinetics.

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

Vallerand, R. J. & O'Connor, B. P. (1991). Construction et validation de L'Échelle de Motivation pour les Personnes Agées. [Construction and validation of the Motivation Scale for the Elderly]. *International Journal of Psychology, 26*, 219–240.

Vallerand, R. J. & Reid, G. (1984). On the causal effects of perceived competence on intrinsic motivation: A test of cognitive evaluation theory. *Journal of Sport Psychology, 6*, 94–102.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senécal, C. & Vallières, E. F., (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement, 52*, 1003–1017.

Weiss, M. R. & Chaumeton, N. (1992). Motivational orientations in sport. In T. S. Horn (Ed.) *Advances in sport psychology* (pp. 61–99). Champaign, IL: Human Kinetics.

Whitehead, J. R. & Corbin, C. B (1991a). Youth fitness testing: The effects of percentile-based evaluative feedback on intrinsic motivation. *Research Quarterly for Exercise and Sport, 62*, 225–231.

Whitehead, J. R. & Corbin, C. B. (1991b). Effects of fitness test type, teacher, and gender on exercise intrinsic motivation and physical self-worth. *Journal of School Health, 61*, 11–16.

Notes

1. The variance-covariance matrix used as data input is available on request from David Markland (see address under Authors' Notes).
2. Univariate skewness values ranged from -1.247 to -.726 (M = -.926). Univariate kurtosis values ranged from -.518 to .667 (M = .116). Normalized Mardia coefficients computed using PRELIS 2.12a revealed that there was significant multivariate skewness ($z = 5.193$, $p < .001$, but not significant multivariate kurtosis ($z = .767$, $p > .10$).
3. The variance-covariance matrix used as data input is available on request from David Markland (see address under Authors' Notes).
4. Univariate skewness values ranged from -1.262 to 1.655 (M = -.001). Univariate kurtosis values ranged from -1.319 to 2.264 (M = -.238). Normalized Mardia coefficients computed using PRELIS 2.12a revealed that there was significant multivariate skewness ($z = 51.170$, $p < .001$, and significant multivariate kurtosis ($z = 15.698$, $p > .001$).

Authors' Notes

The authors would like to thank the reviewers of previous drafts of this manuscript and especially Fuzhong Li of the Oregon Research Institute for their extremely helpful comments and suggestions for improvement. Please address all correspondence concerning this article to David Markland, School of Sport, Health, and Physical Education Sciences, University of Wales–Bangor, Ffriddoedd Building, Victoria Drive, Bangor, Gwynedd, LL57 2EN, Wales, United Kingdom.

E-mail: pes004@bangor.AC.UK

References

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

Baldwin, B. (1989). A primer in the use and interpretation of structural equation models. *Measurement and Evaluation in Counseling and Development, 22*, 100–112.

Bearden, W. O., Sharma, S., & Teal, J. E. (1982). Sample size effects on chi-square and other statistics used in evaluating causal models. *Journal of Marketing Research, 19*, 425–430.

Bentler, P. M. (1990) Comparative fit indexes in structural models. *Psychological Bulletin, 88*, 588–606.

Bentler, P. M. (1992). *EQS: Structural Equations Program Manual.* Los Angeles, CA: BMDP Statistical Software.

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

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika, 52*, 345–370.

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 132–162). Newbury Park, CA: Sage.

Byrne, B. M. (1994). Testing for the factorial validity, replication, and invariance of a measuring instrument: A paradigmatic application based on the Maslach Burnout Inventory. *Multivariate Behavioral Research, 29*, 289–311.

deCharms, R. (1981). Personal causation and locus of control: Two different traditions and two uncorrelated measures. In H. M. Lefcourt (Ed.), *Research with the Locus of Control Construct. Vol. 1: Assessment Methods* (pp. 337–358). New York, Academic Press.

Deci, E. L. (1987). Theories and paradigms, constructs and operations: Intrinsic motivation research is already exciting. *Journal of Social Behavior and Personality, 2*, 177–185.

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. S. Dienstbier (Ed.) *Nebraska Symposium on Motivation Vol. 38: Perspectives on Motivation* (pp. 237–288). Lincoln, NE: University of Nebraska Press.

Duda, J. L. (1992). Motivation in sport settings: A goal perspective approach. In G. C. Roberts (Ed.) *Motivation in Sport and Exercise* (pp. 57–91). Champaign, IL: Human Kinetics.

Duda, J. L. & Tappe, M. K. (1989a) Personal investment in exercise among middle-aged and older adults. In A. C. Ostrow (Ed.) *Aging and Motor Behavior* (pp. 219–238). Indianapolis, IN: Benchmark Press.

Duda, J. L. & Tappe, M. K. (1989b) 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, IN: Benchmark Press.

Fisher, C. D. (1978). The effects of personal control, competence, and extrinsic reward systems on intrinsic motivation. *Organizational Behavior and Human Performance, 21*, 273–288.

Heider, F. (1958). *The psychology of interpersonal relations.* New York, Wiley.

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

Jöreskog, K. G. & Sörbom, D. (1993a). *LISREL 8.03.* Chicago, Scientific Software International.

Jöreskog, K. G. & Sörbom, D. (1993b). *LISREL 8: Structural equation modeling with the SIMPLIS command language.* Hillsdale, NJ: Scientific Software International.

Loehlin, J. C. (1992). *Latent variable models.* Hillsdale, NJ: Lawrence Erlbaum.

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.

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

McAuley, E., Duncan, T. E., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport, 60*, 48–58.

McAuley, E. & Tammen, V. V. (1989). The effects of subjective and objective competitive outcomes on intrinsic motivation. *Journal of Sport and Exercise Psychology, 11*, 84–93.

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.

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.

Muthén, B. & Kaplan, D. (1985). A comparison of methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38*, 171–189.

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.

Plant, R. & Ryan, R. M. (1985). Self-consciousness, self-awareness, ego-involvement, and intrinsic motivation: An investigation of internally controlling styles. *Journal of Personality, 53*, 435–449.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs, 80*, 1–28.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*, 450–461.

Ryan, R. M. & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57*, 749–761.

Ryan, R. M., Mims, V. & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. *Journal of Personality and Social Psychology, 45*, 736–750.