Importance Models of the Physical Self:

Improved Methodology Supports a Normative-Cultural Importance Model but not the

Individual Importance Model

L. Francesca Scalas[1] Alexandre J. S. Morin[2] Herbert W. Marsh[2,3,4] Benjamin Nagengast[5]

[1]University of Cagliari, Italy; [2]University of Western Sydney, Australia; [3]University of Oxford, UK; [4]King Saud University, Saudi Arabia; [5]University of Tübingen, Germany

Corresponding author
L. Francesca Scalas
Dipartimento di Pedagogia, Psicologia, Filosofia
Università degli Studi di Cagliari
Via Is Mirrionis, 1
Cagliari (Loc. Sa Duchessa), 09123, Italy
E-mail: laura.scalas@unica.it

Abstract

We examine theoretical and methodological issues associated with the roles of individual and group-normative importance in self-esteem determination. Critical issues include multicollinearity among the physical self-subdomains, which may have affected previous results, and the need for a multidimensional perspective on importance models. Using Lindwall, Aşçi, Palmeira, Fox, & Hagger (2011)'s database, we apply state-of-the-art methodologies, including Exploratory Structural Equation Modeling and the product-of-indicators approach to latent interactions. Positive interactions would be required to support the Individually Importance-Weighted Average model, but none were observed in the multidimensional model, including all interaction effects; nonetheless, some effects were found in the country-based version of the model. Rather, we found support for the alternative Group Importance-Weighted Average model. We conclude that domain-specific self-concepts are weighted differently and thus differentially affect self-esteem, but these weights do not seem to depend on individual differences in importance. Although awaiting confirmation from further studies, our results suggest the idea that individuals use mainly normative importance processes based on cultural factors in weighting each domain specific component of self-concept.

Keywords: Physical self-concept, importance models, latent interactions, exploratory structural equation modelling

Importance Models of the Physical Self:

Improved Methodology Supports a Normative-Cultural Importance Model but not the

Individual Importance Model

Self-concept is formed by the self-perceptions of a person, which are developed through experience and in interpretations of one's environment (Shavelson, Hubner, & Stanton, 1976). It includes feelings of self-confidence, self-worth, self-acceptance, competence, and ability. It is influenced especially by the evaluations of significant others and by attributions to one's own behaviour (e.g., Harter, 1993, 1996; Shavelson, et al., 1976). Contemporary research highlights the multidimensional and hierarchical nature of self-concept (e.g., Marsh & Craven, 2006), which at the base of the hierarchy comprises perceptions of personal behaviour in specific situations (e.g., physical appearance, sport competence for the physical domain); inferences about self in broader domains (e.g., social, physical, and academic) in the middle of the hierarchy, and global self-concept (i.e. self-esteem) at the apex (Marsh & Shavelson, 1985; Shavelson, et al., 1976).

The role of self-domain importance in self-esteem determination

For decades researchers have been exploring the relations between specific and global components of self-concept (Hattie & Marsh, 1996). In particular, dynamic models of how global self-esteem is formed, maintained, and transformed have been developed to explain how the effects of specific multidimensional domains of self (e.g., physical, social, academic, appearance) are mediated/moderated by framing factors such as ideal standards, importance, schemas, and frames of reference (e.g., Hardy & Moriarty, 2006; Hoge & McCarthy, 1984; Marsh, 1993; Pelham & Swann, 1989). Here, we focus on the role of the importance of specific-domain self-concepts in self-esteem determination.

Individual importance of specific-domain self-concepts

William James (1890/1963) proposed that since a person cannot be all things, each individual must select carefully "the strongest, truest, deepest self . . . on which to stake his salvation" (p. 310) so that "I, who for the time have staked my all on being a psychologist, am mortified if others know much more psychology than I. But I am contented to wallow in the grossest ignorance of Greek." (p. 310). This Jamesian proposition is the precursor of what was later labelled the individual-importance

hypothesis (Hardy & Moriarty, 2006; Lindwall et al., 2011), or the Individually Importance-Weighted Average model (IIWA— Marsh, 1993; Scalas, Marsh, Nagengast, & Morin, 2013). This model proposes that the best representation of a person's overall self-evaluation is an appropriately weighted average of self-evaluations in specific domains. The IIWA model implies that domain-specific self-concepts differentially contribute to global self-esteem determination for different persons on the basis of both inter-individual and intra-individual comparison processes, wherein self-domains will be weighted differentially according to their individual importance.

Controversial empirical findings on the individual model of importance

The IIWA model is often cited as a well-established psychological principle (e.g., Guindon, 2010; Mruk, 2006; Schacter, Gilbert, Wegner, 2009), even though the existence of empirical support for it is contested. For example, Crocker and colleagues operationalized this model in terms of contingencies of self-worth (e.g., Crocker & Wolfe, 2001) and found support for it in relation to state self-esteem (i.e., fluctuations around more stable levels of trait self-esteem; e.g., Crocker, Sommers, & Luhtanen, 2002). However, these results were not replicated on state self-esteem (Leary et al., 2003) nor on trait self-esteem (Lemay & Ashmore, 2006) in later studies. Pelham and Swann (1989) found that a differential importance index based on the IIWA moderated the relationship between self-views in specific domains and self-esteem, and suggested that to test the IIWA appropriately, an intra-individual (within-person) approach is required (see also Hardy & Moriarty, 2006). However, a reanalysis of Pelham and Swann's data (1989) based on both within- and between-person approaches provided no evidence for the role of individual importance either way (Marsh, 1993). This result was replicated and extended in a longitudinal study by Shapka and Keating (2005), who found that weighting the domain-specific self-concepts of adolescents by their individual importance did not improve the ability of specific domains to predict general self-esteem. Some support for the moderating effect of importance was found in relation to narrowly defined domain-specific self-concepts considered unimportant for most people, but very important for some people (e.g., spiritual self-concept—Marsh, 1986; music self-concept—Vispoel, 2003); however, these effects have not always been replicated (e.g., for spiritual self-concept, Scalas et al., 2013).

Hardy and Moriarty (2006) found that out of twelve domain-specific self-concepts examined

in their study the three most important predicted a larger unique portion of the variance in self-esteem than did the three least important. Also, Harter (1990) found that the correlation with global self-esteem was higher for competence self-domains perceived to be important than for those deemed not important. However, both approaches seem to have confounded individual- and group-based relative importance ratings. For example, in a reanalysis of Hardy and Moriarty's data, Marsh (2008) found that the high-low differences in normatively important domain-specific self-concepts (i.e., important for all participants) were significantly related to self-esteem, whereas those based on purely individual ratings of importance were not; thus supporting a normative model of importance (see below). Marsh (2008) also suggested that the apparently demonstrated support for the IIWA was questionable, since most previous research on this topic had not distinguished adequately between individual and normative importance.

Normative or group importance of self-concepts

In line with Marsh (2008) we believe that the distinction between individual and group/normative importance of the various self-domains is crucial to better understanding the factors involved in self-esteem determination. Many factors, in addition to intra-individual factors, have been proposed to play a role in self-esteem determination. Among these, social factors and norms seem to occupy a central position. For instance, the symbolic interactionist perspective suggests that self-esteem is related to how one is regarded by others (Cooley, 1902; Mead, 1934). Similarly, the role of social groups in self-esteem determination and self-enhancement has been recognised by multiple classical social psychology theories (e.g., self-categorization theory—Turner, 1985; social identity theory—Tajfel, 1982; collective identity—Triandis, 1989). In essence, social and normative factors have long been recognised to play a role in self-esteem formation (e.g., Leary & Downs, 1995; Solomon, Greenberg, & Pyszczynski, 1991) and maintenance (e.g., Tesser, 1988).

For example, the Sociometer Theory (e.g., Leary, 2007; Leary & Downs, 1995) defines self-esteem as a sociometer that monitors the social environment. Specifically, when people find cues in the social environment that other people do not regard them positively or even reject them, they experience a loss in self-esteem, the direct consequence of not being valued in the eyes of others. This theory postulates that self-evaluations in domain-specific self-concepts should better predict global

self-esteem if an individual believes that a particular attribute is important for social approval

(positive effect) or disapproval (negative effect). Support for this theory has been found in relation to

both state (e.g., Leary et al., 2003) and trait self-esteem (e.g., Lemay & Ashmore, 2006; MacDonald,

Salztman, & Leary, 2003). Lemay and Ashmore (2006) suggest that seeking social approval is a sort

of universal and implicit motive influencing the level of self-esteem. According to the results from

their longitudinal study, "trait self-esteem appeared related to, and changed, as a function of perceived

regard from others even for people who claimed that their self-esteem was not dependent on others'

regard" (Lemay & Ashmore, 2006, p. 133).

Cultural factors and the normative model of importance

Many cultural factors, as well as ethnicity, have been found to affect self-esteem (e.g.,

Goodwin et al., 2012; Maïano et al., 2006; Morin, Maïano, Marsh, Janosz, & Nagengast, 2011;

Sedikides, Gaertner, & Toguchi, 2003). For example, Markus and Kitayama (1991; see also

Kitayama, Markus, & Lieberman, 1995) suggested different bases for self-esteem in Western and

Eastern societies, which could affect self-evaluations. They proposed self-enhancement as a central

motive in Western societies but not in Eastern societies. Similarly, Heine and colleagues (e.g., Heine

et al., 2001; Heine, Lehman, Markus, & Kitayama, 1999) suggested that unlike Western people,

Eastern societies might privilege self-improvement, and in turn self-criticism, over self-enhancement.

On the other hand, Sedikides and colleagues (e.g., Sedikides et al., 2003; Sedikides, Gaertner, &

Vevea, 2005) posited self-enhancement as a universal yet culturally shaped motive (see also Brown,

2003; Brown & Cai, 2010; Cai, Wu, & Brown, 2009), so that agentic-individualistic attributes might

influence self-esteem construction and enhancement in Western countries, whereas interdependent-

communal attributes could be more critical in Eastern countries. Recently, Gebauer et al. (Gebauer,

Wagner, Sedikides, & Neberich, 2013) found that the links between self-esteem and agency-

communion are moderated by culture, among other factors.

A common theme through these various perspectives has to do with the importance of cultural

factors, norms and standards in self-esteem determination. Going back to the proposal that domain-

specific self-concepts will contribute to a greater or lesser extent to self-esteem determination as a

function of the importance placed on those domains, it thus appears crucial to achieve a proper

differentiation between individual and group-normative ratings of importance. As shown by Marsh (2008), failing to make this distinction could cause bias in the interpretation of results, due to confounding effects between the individual and the group importance models. Indeed, when classic tests of the IIWA model examine whether importance ratings moderate the influence of self-domains on global self-esteem, they usually confound both. Thus, separation of the two components of importance is critical, both from theoretical and substantive perspectives. Interestingly, although the Jamesian perspective has usually been interpreted as being in direct line with the IIWA model, James' (1890) writings do not exclude the possibility of normative group-based importance ratings having an influence on self-esteem determination. Therefore, taking into account this additional perspective and differentiating it from the IIWA model, represents an important clarification of James' perspective and an important contribution of the present investigation.

Operationalization of the Importance Models

The role of self-domain importance in self-esteem determination has fascinated many researchers and has been operationalized in various ways (e.g., Crocker et al., 2002; Hardy & Moriarty, 2006; Harter, 2012; Hattie, 2003; Hattie & Fletcher, 2005; Hoge & McCarthy, 1984; Lindwall, Aşçi, Palmeira, Fox, & Hagger, 2011; Marsh, 1986, 1993, 1994, 1995, 2008; Pelham, 1993, 1995; Pelham & Swan, 1989; Rosenberg, 1982; Scalas et al., 2013). Here we discuss our operationalization of the individual and group-normative importance models (see Table 1 and Figure 1, based on Scalas et al., 2013). According both to individual and to group-normative importance models, depending on their importance, various self-concepts will have different effects on global self-esteem. Thus, individual and normative importance scores can be used to weight different domain-specific self-concepts, in order to provide an appropriate (weighted-average) representation of self-esteem.

Operationalization of the IIWA model versus the GIWA model

In the individual-importance approach (IIWA), weights vary both across domain-specific self-concepts and across individuals (e.g., math self-concept can be weighted more than spiritual self-concept for person A, who values math competence, but it could be the opposite for person B, who values spirituality more).

In the group-importance weighted approach (GIWA), weights differ according to the domain-

specific self-concepts but not according to the importance that an individual places on each domain. So the weights are different across domain-specific self-concepts but remain constant across individuals (e.g., people value math competence more than spirituality). Different strategies can be used to assign the weights for each self-concept facet. Here we consider two specific representations of this approach, the GIWA-free and the GIWA-normative (GIWA-norm) models, which need to be contrasted with one another in order to evaluate the general GIWA theoretical model. In the GIWA-free model, weights are determined empirically as freely estimated weights from regression models; these are optimal weights, since they provide the best representation of what happens in the examined sample—as long as the GIWA model proves to provide the best representation of the data when compared to alternative models, such as the IIWA. In the GIWA-norm, weights are established on the basis of group average importance ratings of each self-concept. Therefore, if the results of the two specific models (GIWA-free and GIWA-norm) are close in terms of fit, explained variance, and (in particular) estimates, this would confirm the premise of the GIWA theoretical model: that people use normative importance ratings to weight domain-specific self-concepts. In other words, the GIWA-norm model constrains the predictive coefficients to take on specific values, in order to systematically test the hypothesis that the relative predictive weights of different self-concept domains are really a function of the group-average level of importance attributed to each domain-specific self-concept. Alternatively, if the fit of the GIWA-free model proves to be much better than the fit of the GIWA-norm model, this means that the relative predictive weight of each domain-specific self-concept is a function of other factors, in addition to or instead of, the relative importance attributed to each of these domains. However, the empirically estimated weights of the GIWA-free model still remain constant across individuals, and thus are group-based weight indicators: that is, simply not a direct or unique function of group-average levels of importance.

Operationalization of the simple unweighted model

Even though literature seems to suggest that self-concept importance, individual or normative, plays a role in self-esteem determination from a methodological point of view, as suggested by Marsh (e.g., 1986, 1993), it is important also to test a baseline model that does not consider the role of the self-domains' importance in self-esteem. This model has been called the simple unweighted model

(cf. Marsh 1993; see Table 1 and Figure 1); here, weights are constant across all domain-specific self-concepts and all individuals (i.e., all domain-specific self-concepts are valued equally by all individuals).

Alternative methods to test the importance models

In the literature, alternative methods have been developed to test the IIWA and GIWA models. For example, in relation to the IIWA, Pelham and Swann (1989) focused on the intra-individual component of importance and proposed a differential importance index: the correlation between domain-specific self-concepts and their importance ratings computed separately for each individual, which provides an index of the similarity between the individual profiles of domain-specific self-concepts and importance ratings. Along the same lines, Hardy and Moriarty (2006) proposed a discounting method in which the three (of twelve) most important domain-specific self-concepts and the three least important domain-specific self-concepts, determined separately for each participant, were identified and used to predict self-esteem. In relation to both IIWA and GIWA, Marsh (1993) proposed a generalized multiple regression approach, which evaluates the increment in self-esteem explained variance (R^2) associated with domain-specific self-concepts, importance ratings, and interaction between the two.

In our operationalization of the IIWA model, within a latent multiple regression approach, once the main effects of importance and domain-specific self-concept have been partialled out (evaluated as separate components in the context of a multivariate analysis), the importance-by-self-concept interaction effect represents whether those domains perceived as more important contribute more to the prediction of self-esteem. Therefore, support for the IIWA would imply that self-esteem is affected by positive interactions between domain-specific self-concept and their individual importance (e.g., Hattie, 2003; Hattie & Fletcher, 2005; Hoge & McCarthy, 1984; Lindwall et al., 2011; Marsh, 1986, 1993, 1994, 1995, 2008; Rosenberg, 1982; Scalas et al., 2013).

Of these methods, the generalized multiple regression is more versatile and has the advantage of allowing specific tests of both the differential importance index (e.g., Marsh, 1993), and the discounting method (Marsh, 2008). Unlike the other methods, this approach contrasts the different theoretical models of importance (IIWA, GIWAs, simple unweighted model), thus avoiding

confounding effects between individual and group-normative importance—the focus of the present investigation.[1]

Criticisms of the Generalized Multiple Regression Approach and its Evolution

Hardy (Hardy & Leone, 2008; Hardy & Moriarty, 2006) criticized Marsh's (1986, 1993) studies based on the generalized multiple regression approach described above, on the basis of potential ceiling effects, due to the high number of predictors included in these studies, and for multicollinearity problems. According to Hardy, if too many first order variables are included in a regression model, less variance remains to be explained by interaction effects. However, Marsh (2008) notes that, as with any other test of interaction, if the main effect of one variable varies as a function of the other variable, then there will be a significant interaction, no matter the number of predictors considered. In this case, the main effects of domain-specific self-concepts will not explain as much variance in self-esteem, when considered alone, as would be the case in the context of interactions. Clearly, there is also a difference when these relations are estimated in the context of fully latent models such as those considered here, versus the context of manifest multiple regression models, where the addition of measurement errors associated with multiple predictors may in itself limit the ability of the full model to detect any significant main or interaction effect. We also note that since only four subdomains are considered in the present study, ceiling effect threats seem not to apply to the present research.

According to Hardy (Hardy & Leone, 2008; Hardy & Moriarty, 2006), Marsh's approach suffers from multicollinearity because of the large number of predictor variables (between 36 and 45) generally entered in his moderation models (Cohen, Cohen, West, & Aiken, 2003; West, Aiken, Wu, & Taylor, 2007). Although multicollinearity is, in general, a potential threat when multiple independent variables are considered simultaneously, Marsh (2008) notes that in his models, not even the single construct analyses (performed separately for each of the 12 domain-specific self-concepts) provided clear support for the IIWA model. Another limitation of the original multiple regression approach is that tests of interaction effects tend to be less efficient in field studies (McClelland & Judd, 1993), mainly because of the use of scale scores that incorporate substantial measurement error (e.g., Bollen, 1989; Thorndike, Cunningham, Thorndike, & Hagen, 1991). In this regard, Marsh

(2008) suggested that interactions based on latent variables corrected for measurement errors could be methodologically more suitable to tests of the IIWA.

Lindwall et al.'s study

Building on work by Hardy & Moriarty (2006) and Marsh (2008), Lindwall et al. (2011) integrated different methods to test the IIWA model in relation to the links between self-esteem and physical self-concept domain and subdomains. Lindwall et al. (2011) used latent interactions to test the IIWA model (this was however a suboptimal version, based on item parcels, which are potentially biased and can camouflage misfit; see Bandalos, 2002; Marsh, Lüdtke et al., 2013; Nasser & Wisenbaker, 2003) based on single constructs, to solve the multicollinearity problem noted by Hardy and Moriarty (2006). For three out of four physical subdomains, Lindwall et al. (2011) found that low actual self-perceptions lead to lower self-esteem when importance is high than when importance is low, and thus the interaction effect was due to a negative effect of low self-perceptions in domains rated as important (see Figure 2 for a theoretical representation of the interaction effect expected in the IIWA model). However, the contribution of self-evaluations in domains rated as important never exceeded the contribution of self-evaluations in domains rated less important (with the only possible exception being the sport competence scale). Interestingly, no significant interaction was found for appearance self-concept, which is usually very strongly related to self-esteem and physical self-worth (Bowker, 2006; Morin & Maïano, 2011; Morin et al., 2011; Scalas & Marsh, 2008; Sonstroem, 1997).

Although Lindwall et al. (2011) partially supported the IIWA model, the single construct approach used in their study has been criticized (Scalas et al., 2013). Indeed, the IIWA model is inherently multidimensional, so that in order to understand whether self-domains valued differently by individuals have different effects on self-esteem, various domain-specific self-concepts should be included in the same model. Specifically, Scalas et al. (2013) transposed the generalized multiple regression approach into a Structural Equation Modeling (SEM) framework, using a full multiple item approach based on latent constructs to control for unreliability (Lindwall et al., 2011; Scalas & Marsh, 2008) and latent interactions to test the IIWA explicitly (Lindwall et al., 2011; Marsh, 2008). They did not use the single construct approach (e.g., Lindwall et al. 2011), but proposed to examine the joint effect of multiple self-domains on self-esteem. To do so they examined three, mostly unrelated,

domain-specific self-concepts (physical, academic and spiritual self-domains), thus producing a set of only 9 predictors in the more complex model—thus taking into account Hardy's criticism of the generalized multiple regression approach. As in Marsh's (1986, 1993, 2008) studies, they found no support for the IIWA model.

Empirical results from direct comparison of the IIWA and GIWA models

Interestingly, studies that have directly compared individual and group importance models, even though they used different operationalizations and methodologies, have found support for the idea that, unlike individual importance, normative importance is crucial to self-esteem (Hoge & McCarty, 1984; Lemay & Ashmore, 2006; Marsh, 2008; Scalas et al., 2013). For example, Marsh (1993; 2008) using the generalized multiple regression approach found that indices of the IIWA model (based on both intra-individual/ipsatized scores and between- subjects scores) were not particularly efficient in predicting self-esteem; they performed little better than the simple unweighted scores, and not nearly as well as other indices based on the GIWA. Scalas et al. (2013) found similar results with a latent version of the generalized multiple regression approach. They contrasted the IIWA and GIWA models using multiple domain-specific self-domains (physical, academic and spiritual self-concepts) in two independent samples. In relation to the IIWA, no positive interaction effect was found to predict self-esteem, whereas support was found for the GIWA. Indeed, two out of three self-domains significantly predicted self-esteem and were differentially weighted according to their average importance. Scalas et al. (2013) concluded that it is not the case that the individuals do not differentially weight the various components of self-concept. However, the weights tend to be determined by normative processes, with little influence gained from individual importance.

In the present study, we propose a direct comparison of the IIWA and GIWA models, using Lindwall et al.'s database (2011). Even though we used the same data, there are several methodological differences between our and Lindwall et al.'s study. We describe them below.

Methodological extensions of Lindwall et al.'s study

In this study, an important extension of Lindwall et al.'s (2011) study is in introducing a different statistical model, which allows us to include all interactions in a single model while reducing the potential multicollinearity issues pointed out by Hardy and Moriarty (2006). This approach is also

consistent with the theory: with current multidimensional perspectives of self-concept (Marsh, 2007a; Shavelson et al., 1976), with earlier research (e.g., Marsh, 1993), and with the analyses based on manifest indicators reported in the Lindwall et al. (2011) study.

We suggest that the failure to include multiple interactions in a single model in the main analyses of Lindwall et al.'s (2011) study was probably due to pragmatic rather than theoretical considerations, and was due mainly to two factors. First, the authors relied on the computationally intensive latent moderated structural equations (LMS) approach to test latent interactions (Klein & Moosbrugger, 2000), and this may have prevented them from testing more than one interaction at the time. This problem can be solved by adopting a different method to test latent interactions, such as the unconstrained product-of-indicators approach (e.g., Marsh, Hau, Wen, Nagengast & Morin, 2013; Marsh, Wen, & Hau, 2004; Marsh, Wen, Hau, Nagengast, 2013), which easily allows for the integration of multiple latent interactions in a single model. Second, the examined constructs are highly correlated to each other (e.g., correlations among physical self-concepts from .56 to .86, and among importance factors: from .59 to .84), and this could result in multicollinearity problems (Hardy & Leone, 2008; Hardy & Moriarty, 2006). We note here that these high correlations tend to be the norm in instruments based on Fox's (1990; Fox & Corbin, 1989) Physical Self-Perception Profile (PSPP; e.g. Hagger, Aşçi, & Lindwall 2004; Marsh & Cheng, 2012; Morin & Maïano, 2011). The lately developed Exploratory Structural Equation Modelling approach provides a potential solution to these problems.

Exploratory Structural Equation Modelling (ESEM). Marsh et al. (Marsh et al., 2009; also see Asparouhov & Muthén, 2009; Morin, Marsh, Nagengast, 2013) recently noted that the Independent Cluster Model (ICM) inherent in CFA—in which each item is allowed to load on a single factor—is overly restrictive for most multidimensional constructs. Furthermore, Morin and Maïano (2011) showed this to be the case for another PSPP-based instrument. Factor structures based on measures used in applied research typically include cross-loadings that can be justified by substantive theory, related to the formulation and wording of the items, or may simply represent another source of measurement error, wherein items are fallible indicators of the constructs and tend to present small residual associations with other constructs (Asparouhov & Muthén, 2009; Church & Burke, 1994).

From a statistical perspective, forcing real non-zero cross-loadings to be zero, as in the ICM-CFA model, typically positively biases factor correlations, potentially distorts factors and biases estimates in SEMs incorporating other constructs (Asparouhov & Muthén, 2009; Marsh et al., 2009; Marsh et al., 2010).

Although Exploratory Factor Analyses (EFAs) provide a way to circumvent this problem by freely estimating all cross-loadings, Marsh et al. (2009, p. 441) note that:

> Applied researchers have persisted with dubious approaches to CFA in the mistaken belief that EFA approaches were no longer acceptable. These misconceptions have been reinforced by the erroneous beliefs that many of the methodological advances associated with CFAs (e.g., goodness-of-fit assessment, complex error structures, growth modeling, latent mean structures, differential item functioning, tests of the full mean and measurement structures over multiple groups or time, introduction of method factors, bi-factor models) are not possible when latent constructs are inferred on the basis of EFAs rather than CFAs.

In addition, Morin and Maïano (2011, p. 3) note that part of this misconception is linked to the erroneous semantic dichotomy between "exploratory" and "confirmatory" methods when in fact EFAs are "perfectly well suited to theory-driven investigations, providing a stronger test that the items will relate to factors in the a priori hypothesized manner—imposing no ICM constraints on the model."

Although the argument can be made that cross-loadings could somehow modify the meaning of the constructs being measured, this argument becomes less relevant and even potentially misleading when cross-loadings remain small (simply allowing for control of items' fallibility as pure indicators of a construct). Indeed, the cross-loadings included in ESEM are not in themselves specified as having a meaning or as having a substantial effect on the meaning of the construct. Rather, they are mostly specified as potentially reflecting one specific type of measurement error linked to validity rather than reliability, wherein each item is presented as a potentially fallible indicator of the construct it is supposed to measure and thus is allowed to present some degree of irrelevant association with the other constructs included in the model. This argument is especially relevant for small (even non-significant) cross-loadings in the context of complex multidimensional models (as used in this study).

Indeed, with multiple constructs measured by multiple items, many tiny cross-loadings are present. They are too small to impact the definition of the constructs clearly. However, together they combine and may inflate the factor correlations. Allowing for these small cross-loadings thus represents a way to control for this specific form of measurement error (we note here that this argument is not merely logical, but is also in line with previous simulation studies; e.g. Asparouhov & Muthén, 2009; Marsh, Lüdtke et al., 2013). In addition, latent measurement models specify the constructs as predicting the indicators, rather than the reverse. Thus, small cross-loadings likely reflect the small number of the indicators that are in fact associated with a secondary construct, without really changing the definition of the construct itself. This interpretation does not hold in the context of principal component analyses or formative measurement models where the items are specified as predicting the construct, or when the cross-loadings are substantial (e.g., Morin & Maïano, 2011). Recently developed Exploratory Structural Equation Modeling (ESEM; Asparouhov & Muthén, 2009; Marsh et al., 2009, 2010) provides an optimal combination of EFA and CFA and integrates many of the advantages of both techniques. ESEM allows items to load on different factors, providing a more accurate representation of the factor structure when non-zero cross-loadings are in fact present in the data, and typically results in substantially lower correlations among the latent factors (Asparouhov & Muthén, 2009; Marsh et al., 2009; Marsh et al., 2010; Morin & Maïano, 2011; Schmitt & Sass, 2011). Importantly, most of the methodological advantages of CFA also apply to ESEM, as demonstrated in the present investigation (see Appendix A of the online supplementary materials for a more technical description of the ESEM approach).

The Present Investigation

In the present investigation, an extension of Lindwall et al.'s (2011) investigation, we examine further theoretical and methodological issues associated with models of importance in self-esteem determination. Considering that literature has shown high, potentially inflated, intercorrelations among the scales of instruments derived from the PSPP, this study first investigates the usefulness of ESEM in providing an improved representation of answers to the revised PSPP, relative to classic CFA approaches (Lindwall et al., 2011). Second, it contrasts different theoretical

models of importance (see Figure 1). Regarding this point, ESEM can help to deal with

multicollinearity problems, thus providing more adequate tests of the IIWA model. Multicollinearity

was a major concern highlighted by Hardy and co-workers (Hardy & Leone, 2008; Hardy & Moriarty,

2006), so we believe that the models presented in this study could represent an interesting

compromise, allowing us to build bridges between the differences in perspective of Marsh and Hardy.

We note that this is the first time that ESEM has been applied to models of self-esteem determination,

and this also represents an original contribution of our research. Third, this study contrasts different

models (single constructs, all main effects and one interaction, all main effects and all interactions) to

provide further tests of the IIWA, showing that single construct models are not only inappropriate

from a theoretical point of view, but also empirically unconvincing. The overarching framework is

one of substantive-methodological synergy (Marsh & Hau, 2007), applying new and evolving

statistical methodology (e.g., ESEM) to better address substantive issues with important theoretical

implications. Since studies directly contrasting the individual and normative importance perspectives

found support for the normative model of importance (e.g., Lemay & Ashmore, 2006; Marsh, 2008;

Scalas et al., 2013), we expect to find similar results.

Method

Participants and Instruments

The study involved participants from four countries: Great Britain (283 females, *M* age =

21.38, *SD* = 2.62; 212 males, *M* age = 22.04, *SD* = 4.19), Turkey (344 females, *M* age = 20.55, *SD* =

1.85; 288 males, *M* age = 21.61, *SD* = 2.36), Portugal (237 females, *M* age = 16.49, *SD* = 1.04; 223

males, *M* age = 16.71, *SD* = 1.31), and Sweden (156 females, *M* age = 36.26, *SD* = 14.18; 88 males,

M age = 35.02, *SD* = 15.11). A full description of participants, instruments and procedures can be

found in Lindwall et al. (2011). Participants completed a revised version of the PSPP (PSPP-R:

Lindwall, Hagger & Aşçi, 2007) derived from Fox's (1990; Fox & Corbin, 1989) original instrument.

The original instrument has often been criticized (e.g., Marsh, Aşçi, Marco, 2002; Marsh, Richards,

Johnson, Roche, & Tremayne, 1994) for its idiosyncratic alternative response format, based on

Harter's instrument (1986). However, we note that this revised form has solved almost all the

problems associated with the previous version (e.g., Lindwall et al., 2007). The PSPP-R comprises 60

positively-worded items rated on a 4-point Likert scale, including 30 self-perception items forming

subscales of sport competence (Sport), physical conditioning (Conditioning), body attractiveness

(Body), physical strength (Strength), and general physical self-worth (PSW). Each competence item is

matched with a corresponding item to measure its perceived importance. For example: "I do very well

at all kinds of sports" is matched with "How important is it to you that you do well at all kinds of

sports?". Finally, six positively-worded items from the Rosenberg Self-Esteem Inventory (Rosenberg,

1989) provide a measure of global self-esteem (GSE). Also, we note that in tests recently conducted

by Lindwall, Aşçi, and Hagger (2011) the PSPP-R was found to be invariant over countries. Although

we acknowledge that the use of all positively worded items could have increased social desirability

and acquiescent response style, recent literature (Siemsen, Roth, & Oliveira, 2010) shows that

multivariate analyses including multiple predictors can help control this issue. Thus, we can conclude

that the instrument has shown good psychometric properties.

Analyses and Models

One of the main differences between the analyses conducted in this study and those in

Lindwall et al. (2011) is that we used a complete multiple-item approach and did not rely on

potentially worrisome parcels for the latent-variable models based on multiple indicators. Overall, we

examined single construct models involving one domain-specific self-concept, one importance factor

and one interaction (Model 1), but also models involving all domain-specific self-concepts, all

importance factors and one interaction (Model 2), and models with all domain-specific self-concepts,

all importance factors and all interactions (Model 3). Using ESEM-based models, we also contrasted

different theoretical models of importance: simple unweighted, GIWA-free, GIWA-norm, IIWA.

Specifically, for the ESEM predictive models, two sets of ESEM factors were used: one for domain-

specific self-concepts and one for importance factors. Only the predictors and moderators were

included in these sets of ESEM factors (i.e. the outcome variables GSE and/or PSW were specified

via regular CFA models). Moreover, correlated uniquenesses were specified between each domain-

specific self-concept item and its corresponding importance item. We used the Geomin oblique

rotation with an epsilon value of .5, as recommended by Marsh et al. (2009, 2010, also see Morin &

Maïano, 2011 for a comparison of rotation procedures in self-concept research). Finally, the ESEM

approach has been adapted to tests of latent interactions (see Appendix B of the online supplemental materials).

The different models of importance are described in Table 1 and Figure 1. In the simple-unweighted model, importance is not considered; all domains are weighted equally (i.e., constrained to be the same) in the prediction of self-esteem. In the GIWA-free model, domain-specific self-concept effects are freely estimated in order to obtain empirically optimal weights; the weights given to each domain differ from one another, but for any one domain the same weight applies to all individuals. In the GIWA-norm it is assumed that domain-specific self-concepts are weighted according to group norms. In previous studies based on non-latent frameworks, group norms typically have been derived on the basis of group-aggregate scores of importance (e.g., Hoge & McCarthy, 1984; Marsh, 1986, 1993; for a presentation of the rationale for this practice, also see Marsh, 1993). As noted by Hoge and McCarthy (1984), overall mean importance ratings are a measure of group saliency and thus can be used to weight the effects of different self-concepts' facets. Here, we have simply adapted this logic to the latent variable framework. In operationalizing the GIWA-norm approach, the unstandardized effect of each domain-specific self-concept on self-esteem was set to be equal to the latent mean of the corresponding importance factor; that is, an estimate of the average importance in the overall group. We used estimates of average importance ratings to reflect the unstandardized predictive coefficients in order to have a convenient way of ensuring that the standardized predictive coefficients are directly proportional to the relative importance of each domain-specific self-concept. Thus, similar results from the two specifications (GIWA-free and GIWA-norm), particularly in relation to the estimates of the actual physical subdomains (i.e. Sport, Body, Strength, Conditioning) on the two outcome latent variables (i.e. PSW, GSE), would support the theoretical normative model of importance. For the GIWA models two versions were specified, one not including latent importance factors and one including them (with paths to self-esteem freely estimated). This allowed us to evaluate the direct contribution of importance in self-esteem determination and to provide a basis of comparison for the IIWA model. Finally, the IIWA model assumes that the effects of the various domain-specific self-concepts on self-esteem are moderated by the importance that each individual places on each of them. Thus, supporting the IIWA model

requires that interaction effects are statistically significant (particularly considering the large sample involved in the present study) and positive (see Figure 2). In this model, all domain-specific self-concepts and importance factors are freely estimated, and all the latent interaction terms are included. We have also considered additional IIWA models, where the interaction regression paths (only for PSW, only for GSE, and for both PSW and GSE) have been fixed to zero; comparing the chi-square of these nested models with the basic IIWA comparison model (with all regression paths freely estimated) provides a direct test of the contribution of interaction effects in PSW and GSE determination.

If any of the above models are able to fit the data and predict global components as well as more complex models with latent interactions (IIWA), the results would support the more parsimonious models. In the case of difficulties in disentangling the best model, incremental R^2 for the outcome variables can be used to decide upon the best model to represent PSW and GSE determination, using the logic of the multiple regression approach (Marsh, 1993). The R^2 values also allow us to have a common metric to evaluate the results of the different models.

To test the moderation effects of importance, we used the product-of-indicators latent approach (e.g., Lin, Wen, Marsh, & Lin, 2010; Marsh et al., 2004; Marsh, Wen et al., 2013). This method can be easily implemented in all SEM software packages and does not require complex non-linear constraints. The interaction effects were standardized, based on the formulas provided by Wen, Marsh and Hau (2010).

Analyses were executed using Mplus 6.1 (Muthén & Muthén, 1998-2008). In order to deal with missing values, full information maximum likelihood estimation was used (Enders & Bandalos, 2001; Muthén & Muthén, 1998-2008). Several fit indices were used to evaluate model fit. Following Marsh, Balla, and Hau (1996; also see Marsh, 2007b; Marsh, Hau, & Wen, 2004), we considered the Tucker-Lewis index (TLI), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA) to evaluate goodness of fit, as well as the χ^2 statistic and an evaluation of parameter estimates. The TLI and CFI vary along a 0-to-1 continuum with values greater than .90 and .95 typically taken to reflect acceptable and excellent fits to the data respectively. RMSEA values of less than .06 are taken to reflect a reasonable fit, whereas RMSEA values greater than .10 are

unacceptable (Byrne, 1998; Hu & Bentler, 1999; Marsh, 2007b; Marsh, Hau et al., 2004). However, it

should be noted that no golden rule exists (Chen, Curran, Bollen, Kirby, & Paxton, 2008; Marsh, Hau

et al., 2004; see also Beauducel & Wittmann, 2005; Fan & Sivo, 2005; Yuan, 2005). This is true

particularly in consideration of the complexity of the models examined here, which were conducted

completely at the item level to avoid potential biases resulting from the use of item aggregates

(Marsh, Lüdtke et al., 2013). The CFI contains no penalty for a lack of parsimony, so that improved

fit due to the introduction of new parameters may reflect capitalization on chance, whereas the TLI

and RMSEA contain penalties for a lack of parsimony (Cheung & Rensvold, 2002; Hu & Bentler,

1999; Marsh, Hau et al., 2004). When contrasting nested models we used $\Delta\chi^2$, as well as ΔCFI and

ΔRMSEA, with decreases smaller than .01 and .015 respectively considered acceptable (Chen, 2007;

Cheung & Rensvold, 2002).

Results

First, we present results from the measurement models based on ESEM and CFA analyses to

show that, in line with recent literature (Asparouhov & Muthén, 2009; Marsh et al., 2009; Marsh et

al., 2010; Morin & Maïano, 2011; Schmitt & Sass, 2011), ESEM models typically provide a better fit

to the data than ICM-CFA models and reduce the sizes of the estimated factor correlations.

Subsequently, moving from issues that remained unsolved in Lindwall et al.'s study (e.g., multiple-

domain models), we present and compare among each other several predictive models applied to

different theoretical models of importance.

ESEM versus CFA Measurement Models

We begin by contrasting ESEM models with corresponding ICM-CFA solutions. First, we

performed separate ESEM analyses for domain-specific self-concepts and importance factors,

showing well-defined factors in both analyses and fit indices (e.g., CFI = .953; TLI = .942) that are

superior to the corresponding ICM-CFA models (e.g., CFI = .913; TLI = .907; see Table 2 for

details).[2]

Subsequently, the six self-concept factors (Body, Sport, Conditioning, Strength, PSW, GSE)

and the five importance factors (Body, Sport, Conditioning, Strength, PSW) were included in the

same model; in this case two sets of ESEM factors were required (one for domain-specific and general self-concepts, and one for importance)[3]. The a priori ESEM solution showed well-defined factors, better fit indices than the corresponding ICM-CFA solution (see Table 2) and, as expected, lower correlations among the latent factors (ICM-CFA mean r = .52, ESEM: mean r = .29), and the cross-loadings remained small (see Table 3).

In order to examine the predictive models, described in the next section, we also tested some hybrid CFA-ESEM measurement models, where PSW and GSE were defined in terms of CFA factors. Indeed, since these latent constructs (PSW and GSW) were specified as outcomes in the predictive models, it would have been inappropriate to specify them as having cross-loadings on factors predicting them and as being partly defined through items having their main loadings on these predictors. Doing so would make the models problematically non-recursive, in that the items used to define the outcomes (i.e., specified as predicted from the latent factors reflecting the outcomes) would be simultaneously used to define their own predictors (i.e., through cross-loadings, the items would also define the latent variables representing the predictors, which in turn would predict the latent outcomes that influence the same items). Furthermore, doing so would not have been possible within ESEM, as all factors forming a single set of ESEM factors need to be specified as similarly related to covariates (predictors and outcomes), so that a subsample from a set of ESEM factors cannot be used to predict another subsample from the same set of factors (Asparouhov & Muthén, 2009). Given this consideration, we specified the outcome variables of the predicting models (PSW and GSW) as CFA factors. In these hybrid models, there were two sets of ESEM factors: one including four domain-specific self-concepts for the physical subdomains, the other including the four corresponding importance factors. The outcome variables, PSW and GSE, were defined as CFA factors. The fit indices (see Table 2) associated with this hybrid model (e.g., CFI = .931; TLI = .920) remained better than for the complete ICM-CFA solution and the correlations remained substantially lower than for the ICM-CFA solution (mean = .36; see Table 4 for details).

In summary, compared to traditional CFA solutions, the ESEM models provided a better fit to the data and substantially reduced factor correlations. In the next section, we compare different theoretical models of importance based on ESEM analyses. For the sake of completeness, the

corresponding CFA-based models are presented in the online supplemental materials (Appendix C).

Comparison between Different Weighted Models of Importance

In relation to our predictive models, we contrasted the previously described theoretical models of importance (simple unweighted, GIWA-free, GIWA-norm, and IIWA; see Table 1 and Figure 1) within the ESEM framework.[4]

Simple unweighted. The simple unweighted model represents a baseline model against which to evaluate the improvements associated with the other theoretical models of importance. In this model, the effects of the various domain-specific self-concepts on the outcomes (PSW, GSE) are constrained to be equal to each other. Consistently with the unrealistic nature of the highly restrictive assumptions underlying this model, the solution was improper. To further simplify the estimation of this model in order to get baseline comparison results, we completely fixed the measurement part of this model to the exact values found in the original ESEM analysis. As expected, this new, simpler model converged on a proper solution; it explained around 58% of the variance in relation to PSW and 29% of the variance for GSE (Table 5).

GIWA models. The GIWA-free model, in which the equality constraints on the weights of the domain-specific self-concepts were removed, resulted in a substantial improvement in the percentage of explained variance for both PSW and GSE (PSW: $R^2 = .808$; GSE: $R^2 = .504$). Therefore, differentially weighting the domains seems to make a big difference in the prediction of PSW and GSE. In line with Scalas et al. (2013), we also found that the fit indices of the GIWA-free and GIWA-norm models (in which the paths from domain-specific self-concepts to PSW and GSE are fixed to be equal to the latent means of the corresponding importance factors) were very similar (Table 5). Although the χ^2 difference test was significant [$\Delta\chi^2 (8) = 80.08$], decrements in other fit indices were trivial (ΔCFI $= .001$; ΔRMSEA $= 0$, see Chen, 2007; Cheung & Rensvold, 2002). More importantly, the R^2 values for the GIWA-norm (PSW: $R^2 = .803$; GSE: $R^2 = .471$) were nearly as high as for the empirically optimal GIWA-free model, thus supporting the hypothesis that domain-specific self-concepts are differentially weighted according to normative levels of importance.

In relation to the regression estimates in the total sample, for the ESEM solutions two effects

were significant in the GIWA-free model and remained significant in the GIWA-norm, in relation to

both PSW (effect of Body: GIWA-free .69, S.E. = .04, GIWA-norm .61, S.E. = .04; effect of

Conditioning: GIWA-free .37, S.E. = .04, GIWA-norm .44, S.E. = .05) and GSE (effect of Body:

GIWA-free .60, S.E. = .03, GIWA-norm .50, S.E. = .04; effect of Conditioning: GIWA-free .25, S.E.

= .05, GIWA-norm .36, S.E. = .05). The results appeared to be quite congruent, moving from the

GIWA-free to the GIWA-norm model; however, the correspondence between the GIWA-free and

GIWA-norm models was less marked than in Scalas et at.'s (2013) study, particularly for GSE. This

might be the result of the higher heterogeneity in the overall sample under consideration. Indeed,

although Scalas and colleagues (2013) found similar estimates in data from two different countries,

the samples were both formed by young adolescents—so that general developmental factors shared by

the two samples might have increased the similarities. On the other hand, in this study the four

subsamples that form the overall sample not only come from different countries, but are also different

in relation to age (e.g., adults in the Swedish sample and adolescents in the Portuguese sample).

Therefore, below we further examine the relation between domain-specific self-concepts and outcome

variables (PSW, GSE) in the four subsamples (also in Appendix D of supplemental materials we

report means and variances of the importance items for each country).

Specifically, we compared multigroup (across-countries) versions of the GIWA-free and

GIWA-norm models (see Table 6 and Table 7). In the multigroup GIWA-free model the regression

paths from the domain-specific self-concepts to PSW and GSE are freely estimated in all countries.

For the GIWA-norm, we tested a model where mean importance estimates, as well as regression paths

from domain-specific self-concepts to PSW and GSW, were specific to each country, taken separately

(GIWA-norm-countries; on the basis of country-specific average importance ratings), or allowed to

differ across countries. In the multigroup GIWA-free model, the regression paths from the domain-

specific self-concepts to the outcome variables showed a similar pattern of results across countries

(see Table 6). For example, the Body and Condition subscales significantly predicted both PSW and

GSE in all countries, while mixed results were found in relation to the Sport and Strength subscales.

Some differences emerged across countries in relation to the relative strength of the paths. For

example, in the Swedish sample, the effect of the Body subscale on PSW was .509, whereas in the

Turkish sample it was higher (.775). The multigroup version of the GIWA-free model was then

compared to the GIWA-norm, based on within-country importance means and to the GIWA-norm,

based on the overall sample importance means. The drop in χ^2 was significant for both models

(respectively: $\Delta\chi^2$ (32) = 262,88; $\Delta\chi^2$ (44) = 257,57), but trivial according to other fit indices (ΔCFI

=.003; ΔRMSEA = .001, for both models). The regression paths were similar in the multigroup

GIWA-free and GIWA-norm-countries models, particularly for the effects on PSW, with only some

exceptions. Overall, the pattern of results remained similar in relation to Body and Condition, with

only a few differences in the strength of the relations across countries. Again, the domains showing

higher differences across countries were Sport and Strength. For instance, the effect of Sport on PSW

in the Swedish sample was not significant in the GIWA-free model, but significant in the GIWA-

norm-country model. In relation to the percentage of explained variance (R^2), the results show very

few differences for PSW when moving from the GIWA-free model to the GIWA-norm model, based

on total sample importance ratings (British sample = .809 vs .810; Swedish sample = .892 vs .885;

Turkish sample = .821 vs .765; Portuguese sample = .830 vs .837). Conversely, the drop was more

marked for the GIWA-norm-country model (British sample = .781; Swedish sample = .878; Turkish

sample = .742; Portuguese sample = .801), suggesting the superiority of weights based on the total

sample over those on country-specific samples. A similar pattern of results was observed in the

prediction of GSW, where the R^2 remained similar for the GIWA-free and GIWA-norm, based on

total sample importance ratings (British sample = .458 vs .450; Swedish sample = .625 vs .543;

Turkish sample = .513 vs .537; Portuguese sample = .696 vs .660), but showed a substantial drop in

the GIWA-norm-countries model (British sample = .388; Swedish sample = .488; Turkish sample =

.479; Portuguese sample = .585).

Overall, these results support the idea that some country differences might be at work in the

normative standards used to evaluate the relative importance of self-concept subdomains in the

physical area. However, these differences do not seem to be substantial, and the prediction seems to

be more precise when based on average normative importance ratings based on the total sample, at

least in relation to the physical subdomains and the European countries examined here.

IIWA model. The IIWA model is an extension of the GIWA-free model, so that the two sets of models differ from each other only for the interaction effects, which are tested in the IIWA but not in the GIWA-free. The overall fit of the IIWA model was adequate (Table 5). In relation to R^2, the inclusion of the interaction effects did not add much to the explanation of both outcome variables (Table 5). However, caution should be used in the interpretation of this result, since interaction effects in applied research often explain a small proportion of variance (McClelland & Judd, 1993). Therefore, in the evaluation of the IIWA it is crucial to inspect the pattern of results associated with the interaction effects. Indeed, the critical test for the IIWA model is the presence of significant and positive interaction effects. Consequently, we performed additional tests. We have already noted, in line with previous literature that the IIWA model is inherently multidimensional; however, the single construct approach is sometimes considered a plausible alternative to the more comprehensive multidimensional model (e.g., Lindwall et al., 2011). Thus, here we juxtaposed three different models: 1) one self-concept, one importance factor, one interaction; 2) all self-concepts, all importance factors, one interaction; 3) all self-concepts, all importance factors, all interactions.

The estimates for the interaction effects (Table 8) changed from Model 1 to Model 3, thus confirming that models based on single constructs are suboptimal, in their failure to take into account the full multidimensionality of physical self-concepts. In Model 1, the one most similar to the models tested by Lindwall et al. (2011), most of the interaction effects (all but the Body Scale) positively predicted both outcome variables (see Table 8). This result is in line with Lindwall et al.'s findings based on item parcels and LMS latent interactions. Moving to Model 2, with all self-concepts as predictors in the model, all importance factors and a single interaction at a time (that is, a total of only 9 predictors in the model), the results changed. Indeed, some effects that were not significant became significant, whereas others that were, became non-significant. For example, most of the interaction effects on PSW became non-significant (see Table 8), whereas the effect of Body on PSW became significant. In the third model, where the effects of all interactions were added to the model, none of the interactions was significant. This result is different from what Lindwall et al. found in their multiple regression analyses, as well as in our CFA-based results (reported in the Appendix C), where the interaction between Sport self-concept and its importance significantly predicted global self-

esteem. Arguably, this third model provides the strongest test of the IIWA model and is more consistent with the inherent multidimensionality of this model. In addition, this model is based on less-correlated factors, due to the reliance on ESEM, is unaffected by measurement error due to the reliance on fully latent variables, and is estimated with sufficient statistical power, due to the large sample size (n = 1840) combined with relatively few predictors (n = 12).

In comparison to CFA models, ESEM models explained more variance for both PSW and GSE (see Appendix C). Not surprisingly, the R^2 values for both PSW and GSE increased in moving from Model 1 (one self-concept, one importance factor and one interaction) to Model 3 (all self-concepts, all importance factors and all interactions). In line with a multidimensional and hierarchical model of self-concept, the contribution of the domain-specific self-concepts (Body, Sport, Strength, Conditioning) to the explained variance of PSW was higher than the explained variance for GSE. Also, the specific contribution of importance factors was relevant in terms of R^2 for both outcome variables (6% for GSE and 3% for PSW). In relation to Model 3 we also tested three additional models: one where the interaction effects on PSW were fixed to zero, another where interaction effects on GSE were fixed to zero, and a third one where the interaction effects on both PSW and GSE were fixed to zero. We computed the $\Delta\chi^2$ between the basic IIWA and these nested models and found that none of them was significant (respectively, $\Delta\chi^2$: 7.48 with 4 d.f., p = .112; 5.39 with 4 d.f., p = .249; 13.70 with 8 d.f., p = .090). Finally, we performed a multigroup version of the IIWA model across countries. First of all, we tested a model with all regression paths constrained to be equal across countries, with the only exception being interaction effects. In this model, for the Turkish sample, the interaction effect from Body to PSW was significant and positive (beta = 077, S.E. = .039). Then we tested a model where interaction effects also were constrained to be equal across countries. The drop in χ^2 was non-significant ($\Delta\chi^2$ (24) = 12.61), and trivial according to other fit indices (ΔCFI =.002; ΔRMSEA = .000). In this model, an interaction effect from Sport to GSW was found in the four countries (beta =.086, S.E. = .042).

In summary, we contrasted different operationalizations of the IIWA model (i.e., Model 1, Model 2, Model 3). Among them, only Model 3 adequately reflected the multidimensional nature of

the IIWA model. In contrast to predictions based on the IIWA model, results showed no positive interactions in the multidimensional model. However, some interaction effects emerged in the country-specific models.

Discussion

In this paper we have proposed methodological extensions to Lindwall et al. (2011) that more appropriately address substantively important tests of the IIWA model. Many self-concept researchers take this model for granted, but mixed results are reported in the literature. Nevertheless, once the distinction between individual and group/normative importance is made clear (e.g., Marsh, 2008), the results seem to support the normative and not the intra-individual model (Hoge & McCarty, 1984; Lemay & Ashmore, 2006; Marsh, 2008; Scalas et al., 2013). Recent advances in statistical methodology available to applied researchers can help clarify the validity of the IIWA and GIWA models. Critical concerns include multicollinearity, due to high correlations among the physical self-subdomains measured by the PSPP (e.g., Hagger et al., 2004; Marsh & Cheng, 2012; Morin & Maïano, 2011) and the multidimensional nature of the IIWA model (e.g., Marsh, 1986, 1993; Scalas et al., 2013). Anchoring this extension within the framework proposed by Marsh (2008; Scalas et al., 2013), and incorporating the spirit of the substantive-methodological synergy discussed by Marsh and Hau (2007), we have also addressed a number of theoretical and methodological issues to clarify unresolved substantive problems, such as the importance models of self-esteem determination discussed here.

Methodological Contribution to the Study of Theoretical Models of Importance

The four physical subdomains in Lindwall et al.'s (2011) database are highly correlated, and we know from the literature that this situation could bias the results of structural equation models in various ways (Marsh et al., 2009; Marsh et al., 2010). Therefore, the investigation of these domain-specific self-concepts would benefit from ESEM, given the ability of this method to more accurately estimate factor structures that are not positively biased by abnormally high correlations (for a similar conclusion as applied to the physical self-concept, see Morin & Maïano, 2011). The ESEM measurement models not only fitted the data better than corresponding ICM-CFA models, but also resulted in substantially smaller correlations among both the self-concepts and the importance factors.

Indeed, as emphasized by Marsh et al. (2009, 2010), whenever non-zero cross-loadings are constrained to be zero, the fit will be poorer. Also, the factor correlations are likely to be positively biased and may introduce problems associated with multicollinearity, and unexpected biases in other parts of the model. Moreover, the superiority of ESEM over CFA models (see Appendix C) in relation to the analyses performed on this database was confirmed by better fit indices and the higher percentage of explained variance in relation to both outcome variables, and in the different models that were investigated.

In line with recent developments (Lindwall et al., 2011; Scalas & Marsh, 2008; Scalas et al., 2013), we adopted a latent approach based on multiple items to control for unreliability. Since the IIWA hypothesizes that domains valued differently by individuals would have different effects on self-esteem, a more adequate operationalization of this multidimensional model would include several domain-specific self-concepts. For this reason, we used the product-of-indicators approach, to enable the inclusion of several interactions at the same time, and we contrasted the results based on the single construct approach with those based on multiple subdomains.

To fully show that single construct approaches are suboptimal, we contrasted the results from different models: Model 1, based on single constructs—one domain-specific self-concept, one importance factor, one interaction; Model 2, including all domain-specific self-concepts, all importance factors and one interaction; and Model 3, based on all domain-specific self-concepts, all importance factors and all interactions. In line with Lindwall et al.'s (2011) results, using only one construct at a time produced several significant interaction effects for both outcomes. However, moving to Model 2 resulted in changes in the significance of interaction effects. Once we moved to the more appropriate specification of the model, including not only all first order factors but also all the interaction factors, all the ESEM interaction effects became non-significant, thus providing no support for the IIWA.

The Role of Importance in Determining Self-Esteem

Lindwall et al. (2011) focused on the individual model of importance, with particular reference to its conceptualization. We have extended their work, contrasting the individual model of importance with a normative model of importance, as well as a model that does not consider the role

of importance in self-concept determination (Marsh, 2008; Scalas et al., 2013). Specifically, we found that the simple unweighted model that does not take into account the importance of domain-specific self-concepts performs weakly in terms of explaining physical self-worth and global self-esteem. We found no support for the IIWA model. Positive and significant interactions would have been required to support the model, but we found none in the multidimensional version of the model estimated on the full sample (although some interactions effects emerged in the country-specific models), thus suggesting that individual importance plays a trivial role in the explanation of both outcome variables.

In relation to the other models of importance, we found support for the theoretical GIWA model. Indeed, the explained variances for physical self-worth and global self-worth in the GIWA-free model were very close to the values found in the GIWA-norm models. Also, all the significant effects in the GIWA-free model remained significant in the GIWA-norm. Overall, these results suggest that self-concept has an important role in self-esteem determination. Domain-specific self-concepts are weighted differently, but these weights do not depend on individual differences in importance. Interestingly, although some differences across countries emerged in relation to the strength of the regression paths from the physical subdomains to both outcome variables, the pattern of results remained highly similar across the four European countries considered here. These results suggest that further studies should attempt to contrast still more culturally different subgroups than those contrasted here. Indeed, domain-specific self-concept seem to be weighted on the basis of overall cultural-norms, which seem to prevail over peculiar country-based norms in self-esteem determination. In relation to the physical area examined in this study, the standards of beauty are quite homogeneous over Europe (and probably over the Western society). So it is not unexpected to have such similarities across countries. Realistically, moving to more culturally-different countries, such as Asian, African, or South American countries would amplify the differences. Therefore, additional empirical investigations should explore the generalizability of these results to even more diversified sets of countries, and in respect of even more diverse facets of the self that in previous research have been shown to present greater levels of cross-cultural variability (e.g. thinness, spirituality, etc.).

Our results supporting the GIWA model are in line with previous findings supporting a group-based importance model (Marsh, 2008; Scalas et al., 2013), as well as research on cultural (Kitayama

et al., 1995; Maïano et al., 2006; Markus & Kitayama, 1991; Morin et al., 2011) factors in self-esteem determination. Socio-cultural models conveyed by Western mass-media in particular may directly or indirectly affect physical self-worth and self-esteem by setting the stage on what is important through injunctive and descriptive norms (e.g., Cialdini, 2007). For example, the effects of media and social comparison are well documented in body-image literature, in relation to the internalization of the thin-ideal and its critical role in the development and maintenance of eating disorders (e.g., Ricciardelli & McCabe, 2001); however, they could also affect other domains and subdomains of self-concept. Therefore, although this is beyond the scope of the present study, potential insights could come from the integration of our GIWA model within the framework of social comparison (Gibbons, Buunk, 1999; Suls & Wheeler, 2000).

Limitations and Directions for Future Research

Notwithstanding the strengths inherent in the present study, limitations remain and should be taken into account in interpreting the results. In particular, the database used in this investigation was limited to the multiple dimensions of physical self-concept domain. Although more research is needed, we note that our results are consistent with those of Scalas et al. (2013), obtained with a wider set of self-domains, and thus support the generalizability of the GIWA model. To provide additional evidence and fully generalize our results, a wider age range of individuals should be investigated. For example, developmental factors, related to age differences between the samples, might have affected the correspondence of estimates between the GIWA-free and the GIWA-norm model results. The participants in the Scalas et al.'s (2013) study were in fact young adolescents aged 13- to 15-years old, whereas in the present database the participants were older, mostly young adults (mean age = 22.06) with an age range from 16- to 50-years old. The importance of peers during adolescence (e.g., Caldwell, Rudolph, Troop-Gordon, & Kim, 2004; McElhaney, Antonishak, & Allen, 2008) is well known. Young people tend to greatly value the opinion of peers, and this might affect their self-concepts and, in particular, the importance given to specific self-domains, especially in normatively-weighted models (GIWA-norm). Thus, normative importance ratings may well have more importance in younger populations undergoing pubertal development, where self-conceptions are still developing (e.g., Morin et al., 2011). Fully developed young adults thus are probably less influenced by group

opinion per se in self-esteem formation, and group-thinking may also be less prevalent than in younger peer groups, due to the cognitive development of critical thinking. However, normative-cultural processes clearly are still at work in young adults, as demonstrated by the similarities in explained variances across the GIWA models, as well as the pattern of regression paths over countries. Nevertheless, further research is needed to clarify this issue fully. Also, the fact that individuals apparently use normative importance processes in weighting each component of self-concept suggests that these effects may differ across cultural groups in which different norms regarding physical standards prevail (e.g. Morin et al., 2011; Siegel, Yancey, Aneshensel, & Schuler, 1999); a possibility that should be expanded upon in future studies including non European countries.

Another limitation of our approach is that it was not designed "explicitly" to test the intra-individual (within-person) perspective that some have attributed to James' writings (e.g., Hardy & Moriarty, 2006). As discussed in the introduction, in the past some approaches were developed to specifically address this point (e.g., the discounting method proposed by Hardy & Moriarty, 2006 and Hardy & Leone, 2008, and the similarity index proposed by Pelham & Swann, 1989). However, it has been noted that such approaches confound intra-individual sources of importance with normative source of importance (e.g., Marsh, 1993, 2008), and thus cannot help to disentangle the individual-importance (e.g., IIWA) and group importance (e.g., GIWAs) models—which were the focus of the present investigation.

Thus, even though the present ESEM-extension of Marsh's 1993 generalized multiple regression approach did not allow us to directly distinguish between the two components of individual importance that had been proposed to be part of the IIWA (intra- and inter- individual), our results clearly fail to support the IIWA model. Rather, they show that individual importance components had no meaningful effect. However, it should be noted that, implicit in the multiple regression framework (or the ESEM extension considered here) that we used is the fact that the effects of each predictor are estimated net of covariance with the other predictors (i.e., only the unique effects of each predictor over and above their shared variance with the other predictors are estimated).

Thus, by incorporating importance ratings for all domains within the same model, the model does not really estimate the effect of each importance factor, considered separately from the others,

but rather, the part of each importance rating that is not shared with the other importance factors (i.e., intra-individual differences). When these models are further extended to incorporate interaction effects between these unique components of importance and the self-concept factors, such models thus implicitly model the moderating role of intra-individual importance on the effects of self-concept factors. Although this control is implicit in the models tested here, it remains a fact that no explicit test of the differential effects of inter- versus intra-individual variations in importance ratings was conducted in this study. Such verification clearly should be the object of future studies. An interesting perspective on these verifications would be to rely on doubly latent multilevel models (e.g., Marsh et al., 2012; Morin, Marsh, Nagengast, & Scalas, 2013), with the importance ratings in each specific domain modeled as level 1 variables (within-person variations in importance ratings) and the global importance ratings as the level 2 variable (inter-individual differences in global importance). As yet this approach has not been implemented in the ESEM framework.

Conclusion

In conclusion, using advanced state-of-the-art methodologies, we have extended analyses of the IIWA model in Lindwall et al. (2011) using fully latent predictive ESEM models. Based on these extended analyses, we found no support for the individual model of importance. Even though the model is appealing from a theoretical point of view, rigorous empirical tests consistently fail to provide evidence for the moderation effect of individual importance in the relation between specific and general components of self-concept. Based on our results, as well as others from the literature (e.g., Farmer, Jarvis, Berent & Corbett, 2001; Marsh, 2008; Scalas et al., 2013; Shapka & Keating, 2005), we believe it is now time to revise this model: Individual importance does not play a major role in self-esteem determination. There is, however, some support for normative importance making a difference, even though this finding offers no support for the model of individual importance. Indeed, the confounding of these two sources of importance seems to be an ongoing source of confusion in the self-esteem literature and the basis, albeit unfounded, of claimed support for the IIWA model. Group importance does play a role in self-esteem determination, so that various self-concepts differentially affect global self-esteem based on normative-cultural standards; this issue clearly is deserving of further investigation.

Endnotes

1. However, it should be noted that, in relation to the IIWA, our approach does not directly disentangle inter-individual and intra-individual sources of variation in importance. We return to this issue in the discussion.

2. The fit indices are not great for the ICM-CFA measurement model. However, this is not surprising, considering that the analyses were conducted at the item level, thus increasing the complexity of the model. The ESEM model shows a much better fit, according to commonly used cut-off criteria.

3. One could argue that ESEM changes the nature of the instrument and the meaning of the factors. To dispel doubts about this, in Table 3 we report the factor loadings and cross-loadings for each factor estimated from the ESEM measurement model. As the reader will notice, the factor loadings clearly identify the original factors, the cross-loadings remain small, and the self-concept and importance factors show parallel patterns of loadings and cross-loadings, attesting to the fact that they reflect parallel constructs.

4. As discussed above, instruments derived from the PSPP often show high correlations between the subscales. This could negatively affect regression estimates by creating problems of multicollinearity. This is particularly true for models including interactions (Aiken & West, 1991). To minimize unnecessary multicollinearity in this study, all variables were standardized at the item level and the cross-product indicators for the latent interaction factors were all centred (Aiken & West, 1991). Nevertheless, since specific tests of multicollinearity are still not available within a latent framework, we also performed classical tests of multicollinearity for the manifest variables (i.e. scale scores) corresponding to the constructs used in our latent models with interaction. Interestingly, these analyses revealed no multicollinearity problems even though the correlations obtained with these scale scores tended to be generally higher (range = .20 - .75, M =.51) than those based on our fully latent ESEM models (range = .02-.72; M = =.37). More precisely, condition indices were all reasonably low (below 6.08 for PSW and below 6.07 for GSW); Tolerance (PSW: range = .25 - .72; mean =.44; GSW: range = .25 - .72; mean =.44) and VIF (PSW: range = 1.39 - 4.04; mean = 2.64; GSW: 1.39 - 4.04; mean = 2.65) values were also acceptable for all variables.

Acknowledgements

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: SAGE.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. Structural *Equation Modeling, 16*(3), 397–438.

Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modelling. *Structural Equation Modeling*, *9*(1), 78–102.

Beauducel, A. & Wittmann, W. W. (2005). Simulation study on fit indices in confirmatory factor analysis based on data with slightly distorted simple structure. *Structural Equation Modeling*, *12*, 41–75.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bowker, A. (2006). The relationship between sports participation and self-esteem during early adolescence. *Canadian Journal of Behavioural Science*, *38*(3), 214–229.

Brown, J. D. (2003). The self-enhancement motive in collectivistic cultures: The rumors of my death have been greatly exaggerated. *Journal of Cross-Cultural Psychology*, *34*, 603–605.

Brown, J. D. & Cai, H. (2010). Self-esteem and trait importance moderate cultural differences in self-evaluations. *Journal of Cross-Cultural Psychology*, *41*(*1*), 116–123.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic*

concepts, applications, and programming. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Cai, H., Wu, Q., & Brown, J. D. (2009). Is self-esteem a universal need? Evidence from the People's Republic of China. *Asian Journal of Social Psychology, 12,* 104–120.

Caldwell, M. S., Rudolph, K. D., Troop-Gordon, W., & Kim, D. (2004). Reciprocal influences among relational self-views, social disengagement and peer stress during early adolescence. *Child Development*, *75*, 1140–1154.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement. *Structural Equation Modeling*, *14*, 464–504.

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods and Research, 36*(4), 462–494.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255.

Church, A. T., & Burke, P. J. (1994). Exploratory and confirmatory tests of the Big 5 and Tellegens 3-dimensional and 4-dimensional models. *Journal of Personality and Social Psychology, 66*(1), 93–114.

Cialdini, R. (2007). Descriptive social norms as underappreciated sources of social control. *Psychometrika*, *72*(2), 263–268.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cooley, C. H. (1902). *Human nature and the social order*. New York: Scribner.

Crocker, J., Sommers, S. R., & Luhtanen, R. K. (2002). Hopes dashed and dreams fulfilled: Contingencies of self-worth and graduate school admissions. *Personality and Social Psychology Bulletin*, *28*, 1275–1286. doi:10.1177/01461672022812012

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, *108*, 593–623. doi:10.1037/0033-295X.108.3.593

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation*

Modeling, 8(3), 430–457.

Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indices to misspecified structural or measurement model components: rationale of two-index strategy revisited. *Structural Equation Modeling*, *12*(3), 343–367.

Farmer, R. F., Jarvis, L. L., Berent, M. K., & Corbett, A. (2001). Contributions to global self-esteem: The role of importance attached to self-concepts associated with the five-factor model. *Journal of Research in Personality*, *35*(4), 483–499.

Fox, K. R. (1990). *The physical self-perception profile manual*. DeKalb: Office for Health Promotion. Northern Illinois University.

Fox, K. R., & Corbin, C. B. (1989). The physical self-perception profile: Development and preliminary validation. *Journal of Sport and Exercise Psychology, 11*, 408–430.

Gebauer, J., Wagner, J., Sedikides, C., & Neberich, W. (2013). Agency-communion and self-esteem relations are moderated by culture, religiosity, age, and sex: Evidence for the "self-centrality breeds self-enhancement" principle. *Journal of Personality, 81*(3), 261–275.

Gibbons, F. X. & Buunk, B. P. (1999). Individual differences in social comparison: The development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, *76*, 129–142.

Goodwin, R., Marshall, T., Fülöp, M., Adonu, J., Spiewak, S., et al. (2012) Mate value and self-esteem: Evidence from eight cultural groups. PLoS ONE *7*(4): e36106. doi:10.1371/journal.pone.0036106

Guindon, M. H. (2010). *Self-esteem across the lifespan: Issues and interventions*. New York: Routledge.

Hagger, M. S., Aşçi, F. H., & Lindwall, M. (2004). A cross-cultural evaluation of a multidimensional and hierarchical model of physical self-perceptions in three national samples. *Journal of Applied Social Psychology*, *34*, 1075–1107.

Hardy, L., & Leone, C. (2008). Real evidence for the failure of the Jamesian perspective or more evidence in support of it? *Journal of Personality, 76*(5), 1123–1136.

Hardy, L., & Moriarty, T. (2006). Shaping self-concept: The elusive importance effect. *Journal of*

Personality, 74(2), 377–401.

Harter, S. (1986). Process underlying the construction, maintenance and enhancement of self-concept in children. In J. Suls & A. Greenwald (Eds.), *Psychological perspectives on the self*: (Vol. 3, pp. 136–182). Hillsdale, NJ: Erlbaum.

Harter, S. (1990). Self and identity development. In S. Feldman and G. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 352–387). Cambridge, MA: Harvard University Press.

Harter, S. (1993). Causes and consequences of low self-esteem in children and adolescents. In R. F. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 87–116). New York: Plenum.

Harter, S. (1996). Historical roots of contemporary issues involving self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept. Developmental, social, and clinical considerations* (pp. 1–37). New York: John Wiley.

Harter, S. (2012). *The construction of the self*. New York: Guilford Press.

Hattie, J. (2003). Getting self-concept back on the correct pathway for self-concept research in the new millennium: Revisiting misinterpretations and revitalizing the contributions of James' agenda for research on self. In H. W. Marsh, R. G. Craven & D. M. McInerney (Eds.), *International advances in self research* (Vol. 1, pp. 127–150). Greenwich, CT: Information Age.

Hattie, J., & Fletcher, R. (2005). Self-esteem = success/pretensions: Assessing pretensions/importance in self-esteem. In H. W. Marsh, R. G. Craven & D. M. McInerney (Eds.), *International advances in self research* (Vol. 2, pp. 123–152). Greenwich, CT: Information Age.

Hattie, J., & Marsh, H. W. (1996). Future directions in self-concept research. In B. Bracken (Ed.), *Handbook of self-concept* (pp. 421–462). New York: John Wiley & Sons.

Heine, S. J., Kitayama, S., Lehman, D. R., Takata, T., Ide, E., Leung, C., & Matsumoto, H. (2001). Divergent consequences of success and failure in Japan and North America: An investigation of self-improving motivations and malleable selves. *Journal of Personality and Social Psychology, 81,* 599–615.

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review, 106,* 766–794.

Hoge, D. R., & McCarthy, J. D. (1984). Influence of individual and group identity salience in the

global self-esteem of youth. *Journal of Personality and Social Psychology, 47*(2), 403–414.

Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.

James, W. (1890/1963). *The principles of psychology*. New York: Holt, Rinehart & Winston.

Kitayama, S., Markus, H. R., & Lieberman, C. (1995). The collective construction of self esteem. In J. Russell, J. Fernandez-Dols, A. Manstead, & J. Wellenkamp (Eds.), *Everyday conceptions of emotion: An introduction to the psychology, anthropology and linguistics of emotion* (pp. 523–550). Netherlands: Kluwer Academic.

Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*, 457–474.

Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, *58*, 317–344.

Leary, M. R., & Downs, D. L. (1995). Interpersonal functions of the self-esteem motive: The self-esteem system as a sociometer. In M. H. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 123–144). New York: Plenum.

Leary, M. R., Gallagher, B., Fors, E., Buttermore, N., Baldwin, E., Kennedy, K., et al. (2003). The invalidity of disclaimers about the effects of social feedback on self-esteem. *Personality and Social Psychology Bulletin*, *29*, 623–636.

Lemay, E. P., Jr. & Ashmore, R. D. (2006). The relation of social approval contingency to trait self-esteem: Cause, consequence, or moderator? *Journal of Research in Personality*, *40*, 121–139.

Lin, G. C., Wen, Z., Marsh, H. W., & Lin, H. S. (2010). Structural equation models of latent interactions: Clarification of orthogonalizing and double-mean-centering strategies. *Structural Equation Modeling*, *17*(3), 374–391.

Lindwall, M., Aşçi, & Hagger, M. S. (2011). Factorial validity and measurement invariance of the revised Physical Self-Perception Profile (PSPP-R) in three countries. *Psychology, Health and Medicine*, *16*, 115–128.

Lindwall, M., Aşçi, F. H., Palmeira, A., Fox, K. R., & Hagger, M. S. (2011). The importance of importance in the physical self: Support for the theoretically appealing but empirically elusive

model of James. *Journal of Personality*, *79*(2), 303–334.

Lindwall, M., Hagger, M., & Aşçi, H. (2007). Introducing the Physical Self-Perception Profile—Revised: Psychometric examination in three samples from Great Britain, Sweden and Turkey. Book of Abstract of the 12th European Congress of Sport Psychology. FEPSAC publication.

MacDonald, G., Saltzman, J. L., & Leary, M. R. (2003). Social approval and trait self-esteem. *Journal of Research in Personality*, *37*, 23–40.

Maïano, C., Ninot G., Stephan Y., Morin A. J. S., Florent, J. F., & Vallée, P. (2006). Geographic region effects on adolescent physical self: An exploratory study. *International Journal of Psychology*, *41*(2), 73–84.

Markus, H., & Kitayama, S. (1991). Cultural variation in the self-concept. In J. Strauss & G. R. Goethals (Eds.), *The self: Interdisciplinary approaches* (pp. 18–48). New York, NY: Springer-Verlag.

Marsh, H. W. (1986). Global self-esteem: Its relation to specific facets of self-concept and their importance. *Journal of Personality and Social Psychology, 51*(6), 1224–1236.

Marsh, H. W. (1993). Relations between global and specific domains of self: The importance of individual importance, certainty, and ideals. *Journal of Personality and Social Psychology, 65*(5), 975–992.

Marsh, H. W. (1994). The importance of being important: Theoretical models of relations between specific and global components of physical self-concept. *Journal of Sport and Exercise Psychology, 16*, 306–325.

Marsh, H. W. (1995). A Jamesian model of self-investment and self-esteem: Comment on Pelham (1995). *Journal of Personality and Social Psychology, 69*(6), 1151–1160.

Marsh, H. W. (2007a). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester, UK: British Psychological Society.

Marsh, H. W. (2007b). Application of confirmatory factor analysis and structural equation modeling in sport/exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of Sport Psychology* (pp. 774–798). New York: Wiley.

Marsh, H. W. (2008). The elusive importance effect: More failure for the Jamesian perspective on the

importance of importance in shaping self-esteem. *Journal of Personality, 76*(5), 1081–1122.

Marsh, H. W., Aşçi, F. H., & Marco, I. T. (2002). Multitrait-multimethod analyses of two physical self-concept instruments: A cross-cultural perspective. *Journal of Sport and Exercise Psychology*, *24*, 99–119.

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical properties. *Advanced structural equation modeling: Issues and techniques*, 315–353.

Marsh, H. W., & Cheng, J. H. S. (2012). Physical self-concept. In G. Tenenbaum, R. Eklund, & A. Kamata (Eds). *Measurement in sport and exercise psychology* (pp. 215–226). Champaign, IL: Human Kinetics.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, *1*(2), 133–163. doi:10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., & Hau, K-T. (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology*, *32*, 151–171.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320–341.

Marsh, H. W., Hau, K.-T., Wen, Z., Nagengast, B., & Morin, A. J. S. (2013). Moderation. In T.D. Little (Ed.), *Oxford Handbook of Quantitative Methods* (pp. xxx). New York, NY: Oxford University Press.

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U. & Nagengast, B. (2010). A new look at the big-five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22,* 471–491.

Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J. S., Trautwein, U., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—camouflaging misspecification with item-parcels in CFA models. *Psychological Method*s. doi:

10.1037/a0032773

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling, 16,* 439–476.

Marsh, H. W., Richards, G. E., Johnson, S., Roche, L., & Tremayne, P. (1994). Physical self-description questionnaire: Psychometric properties and a multitrait-multimethod analysis of relations to existing instruments. *Journal of Sport and Exercise Psychology*, *16*, 270–305.

Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107–125. doi:10.1207/s15326985ep2003_1

Marsh, H. W., Wen, Z., & Hau, K.-T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, *9*, 275–300.

Marsh, H. W., Wen, Z., Hau, K.-T., & Nagengast, B. (2013). Structural equation models of latent interaction and quadratic effects. In G.R. Hancock & R.O. Mueller (Eds.), *Structural Equation Modeling. A second course* (2nd Ed.) (pp. 436–458). Greenwich, CT: Information Age.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, *114*, 376–390.

McElhaney, K. B., Antonishak, J., & Allen, J. P. (2008). "They like me, they like me not": Popularity and adolescents' perceptions of acceptance predicting social functioning over time. *Child Development*, *79*, 720–731.

Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.

Morin A. J. S., & Maïano, C. (2011). Cross-validation of the short form of the physical self-inventory (PSI-18) using exploratory structural equation modeling (ESEM). *Psychology of Sport and Exercise*, *12*(5), 540–554.

Morin, A. J. S., Maïano, C., Marsh, H. W., Janosz, M., & Nagengast, B. (2011). The longitudinal interplay of adolescents' self-esteem and body image: A conditional autoregressive latent trajectory analysis. *Multivariate Behavioral Research*, *46*(2), 157–201.

Morin, A. J. S., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In

G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course, 2nd Edition* (pp. 395–436). Greenwich, Connecticut: IAP.

Morin, A. J. S. , Marsh, H. W., Nagengast, B., & Scalas, L. F. (2013). Doubly latent multilevel analyses of classroom climate: An illustration. *The Journal of Experimental Education.* doi: 10.1080/00220973.2013.769412.

Mruk, C. J. (2006). *Self-esteem: Research, theory, and practice: toward a positive psychology of self-esteem*. New York: Springer.

Muthén, L. K., & Muthén, B. O. (1998–2008). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement*, *63*, 729–757.

Pelham, B. W. (1993). The idiographic nature of human personality: Examples of the idiographic self-concept. *Journal of Personality and Social Psychology, 64*(4), 665–677.

Pelham, B. W. (1995). Self-investment and self-esteem: Evidence for a Jamesian model of self-worth. *Journal of Personality and Social Psychology, 69*(6), 1141–1150.

Pelham, B. W., & Swann Jr, W. B. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal of Personality and Social Psychology, 57*(4), 672–680.

Ricciardelli, L., & McCabe, M. (2001). Self-esteem and negative affect as moderators of sociocultural influences on body dissatisfaction, strategies to decrease weight, and strategies to increase muscles among adolescent boys and girls. *Sex Roles*, *44*, 189–207.

Rosenberg, M. (1982). Psychological selectivity in self-esteem formation. In M. Rosenberg & H. B. Kaplan (Eds.), *Social psychology of the self-concept* (pp. 535–546). Arlington Heights, IV. Harlan Davidson.

Rosenberg, M. (1989). *Society and the adolescent self-image*. Revised edition. Middletown, CT: Wesleyan University Press.

Scalas, L. F., & Marsh, H. W. (2008). A stronger latent-variable methodology to actual-ideal discrepancy. *European Journal of Personality, 22*(7), 629–654.

Scalas, L.F., Marsh, H.W., Nagengast, B., & Morin, A.J.S. (2013). Latent variable approaches to the Jamesian model of importance-weighted averages. *Personality and Social Psychology Bulletin*, *39*, 100-114

Schacter, D. L., Gilbert, D. T., & Wegner, D. M. (2009). *Psychology*. New York: Worth.

Schmitt, T. A., & Sass, D. A. (2011). Rotation criteria and hypothesis testing for exploratory factor analysis: implications for factor pattern loadings and interfactor correlations. *Educational & Psychological Measurement*, *71*, 95–113.

Sedikides, C., Gaertner, L., & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of Personality and Social Psychology*, *84,* 60–79.

Sedikides, C., Gaertner, L., & Vevea, J. L. (2005). Pancultural self-enhancement reloaded: A meta-analytic reply to Heine. *Journal of Personality and Social Psychology*, *89*, 539–551.

Shapka, J. D., & Keating, D. P. (2005). Structure and change in self-concept during adolescence. *Canadian Journal of Behavioural Sciences*, *37*, 83–96.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*, 407–441.

Siegel, J. M., Yancey, A. K., Aneshensel, C. S., & Schuler, R. (1999). Body image, perceived pubertal timing, and adolescent mental health. *Journal of Adolescent Health*, *25*, 155–165.

Siemsen, E., Roth, A. & Oliveira, P. (2010). Common method bias in regression models with linear, quadratic, and interaction effects. *Organizational Research Methods*, *13*, 456–476.

Solomon, S., Greenberg, J., & Pyszczynski, T. (1991). A terror management theory of social behavior: The psychological functions of self-esteem and cultural worldviews. *Advances in Experimental Social Psychology*, *24*, 93–159. doi:10.1016/S0065-2601(08)60328-7

Sonstroem, R. J. (1997). The physical self-system: A mediator of exercise and self-esteem. In K. R. Fox (Ed.), *The physical self: From motivation to well-being* (pp. 3–26). Champaign: Human Kinetics.

Suls, J., & Wheeler, L.(2000). *Handbook of social comparison: Theory and research*. New York: Kluwer.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*, pp.

1–39.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181–227). New York: Academic Press.

Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). *Measurement and evaluation in education and psychology*. New York: Macmillan.

Triandis, H. C. (1989). The self and behavior in different cultural contexts. *Psychological Review*, *96*, 506–520.

Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behaviour. In E. J. Lawler (Ed.), *Advances in group processes: Theory and research* (Vol. 2, pp. 77–122). Greenwich, CT: JAI Press.

Vispoel, W. P. (2003). Measuring and understanding self-perceptions of musical ability. In H. W. Marsh, R. G. Craven & D. M. McInerney (Eds.), *International advances in self research* (Vol. 1, pp. 151–180). Greenwich, CT: Information Age.

Wen, Z., Marsh, H. W. & Hau, K-T. (2010). Structural equation models of latent interactions: An appropriate standardized solution and its scale-free properties. *Structural Equation Modeling*, *17*(1), 1–22.

West, S. G., Aiken, L. S., Wu, W., & Taylor, A. B. (2007). Multiple regression: Applications of the basics and beyond in personality research. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 573–601). New York: Guilford.

Yuan, K. H. (2005). Fit indices versus test statistics. *Multivariate Behavioral Research*, *40*(1), 115–148.

Table 1. *Theoretical Models of Importance*

	Simple unweighted approach	Group importance weighted approach (GIWA)		Individual importance weighted approach (IIWA)
Description of the model	All self-concepts contribute equally to self-esteem determination	Self-concepts differentially contribute to self-esteem determination, but this contribution is largely dependent on normative/group processes. Two variants of the model are considered, the GIWA-free (based on empirical optimal weighting) and the GIWA-norm (based on group importance factors)		Self-concepts differentially contribute to self-esteem determination, since importance of self-concepts depends on intra-individual processes. Self-domains are weighted differentially according to individual importance
Self-concept weights	The same/constant weights are attributed to all self-concepts	Different weights are attributed to all self-concepts		Different weights are attributed to all self-concepts
Individual weights	The same/constant weights are attributed to all subjects	The same weights are attributed to all subjects		Different weights are attributed to all subjects
Specification of the model	Self-concept weights are constrained to be equal	GIWA-free: All self-domains are freely estimated on the basis of empirical optimal group weights	GIWA-norm: All self-domain weights are constrained to be equal to the corresponding importance means of the overall group	Latent interactions between self-concept and importance factors are included in the model along with main effects
Evaluation of the model	Support for the SUA model requires fit indices and explained variance (R^2) to be as good as the corresponding GIWA-free model based on optimal weights	To evaluate the GIWA model a comparison of the GIWA-free and GIWA-norm is required. If the results in terms of fit, R^2, and critical estimates are similar, then the GIWA is supported		Support for the IIWA model requires significant increments in R^2 in comparison to the GIWA-free model, and significant and positive interactions between self-concepts and their corresponding importance factors
Role of importance for self-esteem	Support for this model would imply that importance scores are irrelevant for self-esteem determination	Support for this model would imply that normative importance scores affect self-esteem determination		Support for this model would imply that individual importance scores affect self-esteem determination

Table 2. *ESEM versus CFA-Restricted Measurement Models*

Model	χ^2 (df)	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR
Self concept only						
ICM-CFA a priori 6 factors	3260.533(579)*	.920	.913	.050	.048-.052	.045
ESEM a priori 6 factors	1635.297 (429)*	.964	.947	.039	.037-.041	.018
Hybrid ESEM: 4 a priori ESEM factors + GSE and PSW as CFA factors	2183.122(497)*	.950	.936	.043	.041-.045	.030
Importance only						
ICM-CFA a priori 5 factors	2779.470(395)*	.908	.899	.057	.055-.059	.047
ESEM a priori 5 factors	1411.909 (295)*	.957	.936	.045	.043-.048	.020
ESEM: 4 a priori ESEM factors, without a PSW importance factor	1213.547(186)*	.949	.924	.055	.052-.058	.025
Self Concept & Importance						
ICM-CFA with CUs	7916.559(1994)*	.913	.907	.040	.039-.041	.048
ESEM: 2 sets with CUs	4948.930(1744)*	.953	.942	.032	.031-.033	.029
Hybrid ESEM: 2 sets of ESEM factors with CUs, GSE & PSW as CFA factors, without a PSW importance factor	5739.465(1521)*	.931	.920	.039	.038-.040	.039

Note. * $p < .001$

Table 3. *Factor Loadings of the ESEM Factors*

Item	Body			Strength			Sport			Conditioning		
	Self-concept factor loading	Importance factor loading	Interaction factor loading	Self-concept factor loading	Importance factor loading	Interaction factor loading	Self-concept factor loading	Importance factor loading	Interaction factor loading	Self-concept factor loading	Importance factor loading	Interaction factor loading
SPORT1	-.006	-.027	.027	.099	.038	-.024	**.579**	**.681**	**.555**	.087	.085	-.025
SPORT2	.067	.067	.039	.155	.154	.084	**.628**	**.660**	**.684**	.108	.052	.023
SPORT3	.091	.105	.076	-.045	.014	-.052	**.539**	**.443**	**.283**	.295	.321	.373
SPORT4	.096	.075	.044	.179	.208	.185	**.647**	**.565**	**.455**	.098	.132	.212
SPORT5	.040	.118	-.005	.067	.233	.071	**.518**	**.492**	**.319**	.083	.086	.201
SPORT6	-.010	-.019	-.064	.083	.235	.118	**.608**	**.514**	**.387**	.233	.169	.255
COND1	.241	.179	.157	.076	.046	.021	.226	.351	.314	**.296**	**.182**	**.181**
COND2	-.086	-.032	-.030	.104	.093	.031	.043	.024	.099	**.605**	**.641**	**.406**
COND3	.100	.078	.029	.164	.082	.090	.192	.180	.212	**.491**	**.597**	**.465**
COND4	.131	.144	.084	.002	.027	.028	.209	.260	.147	**.549**	**.471**	**.480**
COND5	.083	.173	.104	.068	.039	.020	.078	.060	.123	**.676**	**.658**	**.618**
COND6	.144	.025	.067	.215	.175	.179	.169	.071	.092	**.496**	**.657**	**.512**
BODY1	**.611**	**.520**	**.541**	.055	.068	.042	.151	.130	.211	-.019	-.017	-.072
BODY2	**.562**	**.544**	**.460**	-.020	.051	-.044	.049	.115	.096	.176	.145	.095
BODY3	**.701**	**.699**	**.655**	.028	.060	.028	.019	.002	.022	.110	.048	-.006
BODY4	**.695**	**.662**	**.546**	.156	.165	.119	.094	.049	.079	-.090	-.039	.012
BODY5	**.636**	**.676**	**.655**	.149	.187	.158	.051	.027	.036	.141	.071	.076
BODY6	**.707**	**.656**	**.540**	-.056	-.069	-.003	-.025	.007	-.238	.188	.283	.403
STREN1	.022	.147	.085	**.598**	**.429**	**.479**	.095	.261	.170	.016	-.018	-.079
STREN2	.043	.073	.067	**.714**	**.634**	**.614**	.081	.180	.132	.056	.015	-.051
STREN3	-.014	.018	-.007	**.558**	**.523**	**.428**	.155	.138	.053	.137	.124	.090
STREN4	.046	.050	.022	**.565**	**.495**	**.516**	.141	.082	.004	.213	.321	.250
STREN5	.150	.151	.185	**.665**	**.685**	**.623**	.085	.015	.065	.098	.131	.048
STREN6	.041	.066	.016	**.678**	**.679**	**.679**	.087	.130	.062	.154	.068	.083

Note. Non-significant correlations are in bold.

Table 4. *Correlations for the CFA and ESEM Restricted Measurement Models*

	Sport	Cond	Body	Stren	Psw	GSE	IMPsport	IMPcond	IMPbody	IMPstren
Sport	1	.574	.288	.449	.472	.327	.719	.434	**.024**	.359
Cond	.861	1	.361	.421	.631	.473	.463	.681	**.089**	.292
Body	.556	.662	1	.246	.805	.636	.295	.161	.295	.200
Stren	.745	.754	.517	1	.361	.177	.371	.323	.101	.678
Psw	.593	.712	.854	.526	1	.867	.397	.308	.086	.203
GSE	.417	.499	.666	.327	.867	1	.192	.224	**.002**	**.013**
IMPsport	.726	.642	.426	.574	.348	.179	1	.482	.301	.533
IMPcond	.606	.681	.387	.522	.348	.209	.813	1	.352	.398
IMPbody	.222	.288	.351	.248	.164	**.047**	.591	.657	1	.355
IMPstren	.536	.526	.372	.677	.286	.084	.839	.742	.643	1

Note. Non-significant correlations are in bold. Below the diagonal are reported the CFA-based correlations; above the diagonal are reported the ESEM-based correlations.

Table 5. *Fit Indices and R² Incremental Change for Different Models of Importance in the ESEM Based Analyses*

Model	Effects on self-esteem	χ^2 (df)	CF	CFI	TLI	RMSEA	90% CI	PSW R^2	GSE R^2
Simple-unwgted	SCs	3313.221 (609)*	1.148	.919	.917	.049	.047-.051	.584	.293
GIWA-norm	SCs	2791.518 (527)*	1.141	.932	.919	.048	.047-.050	.803	.471
	SCs and IMPs	5819.570 (1529)*	1.136	.930	.919	.039	.038-.040	.805	.495
GIWA-free	SCs	3012.005 (519)*	1.043	.926	.910	.051	.049-.053	.808	.504
	SCs and IMPs	5739.487 (1521)*	1.136	.931	.920	.039	.038-.040	.814	.530
IIWA	SCs, IMPs and INTs	8992.609 (3107)*	1.156	.916	.906	.032	.031-.033	.816	.545

Note. $* p < .001$; simple-unwgted = group importance weighted approach—unweighted, with all self-concept paths to self-esteem constrained to be equal to each other (no importance factors); GIWA-norm = group importance weighted approach with regression paths of each self-domain to self-esteem set to be equal to the corresponding group importance latent mean; GIWA-free = group importance weighted approach with free estimates for the regression paths of each self-domain to self-esteem; IIWA = Individually Importance-Weighted approach, with free estimates for the regression paths of each self-domain and each importance factor to self-esteem, latent interactions included. SCs = self-concepts; IMPs = importance factors; INTs = interaction effects; CF= correction factor.

Table 6. *Standardized Estimates for the GIWA Models in the Four Countries*

	GIWA-free		GIWA-norm countries		GIWA-norm overall sample	
	Regression paths freely estimated		Regression paths freely estimated and proportional to country importance means		Regression paths proportional to the overall sample importance means	
	PSW	GSE	PSW	GSE	PSW	GSE
	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
British sample						
Body	.631 (.041)	.567 (.052)	.573 (.038)	.436 (.033)	.636 (.026)	.518 (.025)
Condition	.332 (.077)	**.164 (.106)**	.336 (.050)	.256 (.039)	.367 (.038)	.298 (.032)
Sport	.178 (.084)	**.056 (.117)**	.190 (.059)	.145 (.046)	.118 (.037)	.096 (.030)
Strength	.138 (.061)	**.158 (.083)**	.169 (.053)	.128 (.041)	.082 (.032)	.067 (.026)
Swedish sample						
Body	.509 (.048)	.564 (.069)	.504 (.042)	.382 (.036)	.648 (.026)	.531 (.026)
Condition	.487 (.082)	.422 (.139)	.428 (.070)	.324 (.055)	.373 (.041)	.306 (.034)
Sport	**.132 (.065)**	**.209 (.110)**	.194 (.056)	.147 (.043)	.120 (.037)	.099 (.031)
Strength	.294 (.071)	**.172 (.137)**	.286 (.062)	.217 (.048)	.084 (.032)	.068 (.026)
Turkish sample						
Body	.775 (.033)	.577 (.043)	.689 (.037)	.544 (.032)	.649 (.024)	.543 (.023)
Condition	.322 (.068)	.351 (.089)	.289 (.045)	.228 (.036)	.374 (.040)	.313 (.034)

Sport	**.015 (.062)**	**.002 (.085)**	**.095 (.049)**	**.075 (.039)**	.121 (.037)	.101 (.031)
Strength	**-.048 (.049)**	**-.094 (.068)**	**.038 (.041)**	**.030 (.033)**	.083 (.032)	.070 (.027)
Portuguese sample						
Body	.516 (.049)	.604 (.052)	.465 (.044)	.372 (.039)	.618 (.024)	.536 (.024)
Condition	.445 (.075)	.271 (.091)	.365 (.051)	.292 (.042)	.356 (.038)	.309 (.034)
Sport	.195 (.087)	.229 (.097)	.229 (.054)	.183 (.043)	.115 (.035)	.100 (.031)
Strength	**.106 (.069)**	**-.079 (.083)**	.183 (.054)	.147 (.043)	.079 (.031)	.069 (.026)

Note. Non-significant regression paths in bold.

Table 7. *Latent Means for Multigroup GIWA Models*

	Sports	Conditioning	Body attractiveness	Physical Strength	Sports importance	Conditioning importance	Body attractiveness importance	Physical Strength importance
	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)	Estimate (S.E.)
Multi-group GIWA-free with regression paths freely estimated across countries								
British sample	-.021 (.051)	.065 (.052)	.028 (.050)	.028 (.050)	.007 (.052)	.030 (.050)	-.005 (.049)	.002 (.050)
Swedish sample	-.002 (.078)	.000 (.076)	.000 (.072)	.000 (.070)	-.001 (.074)	-. 002 (.075)	-.001 (.072)	-.001 (.070)
Turkish sample	-.001 (.045)	-.002 (.043)	.001 (.043)	-.001 (.045)	-.001 (.044)	-.001 (.040)	.000 (.044)	-.002 (.045)
Portuguese sample	.014 (.056)	-.026 (.077)	-.011 (.055)	-.009 (.054)	.004 (.053)	-.007 (.129)	.004 (.058)	.000 (.067)
Multi-group GIWA-norm country: model based on importance values of each separate country								
British sample	.009 (.039)	.132 (.041)	.122 (.050)	.065(.041)	.105 (.033)	.186 (.029)	.317 (.021)	.093 (.029)
Swedish sample	.053 (.064)	.377 (.072)	-.022 (.070)	.118 (.058)	-.122 (.037)	.566 (.049)	.196 (.028)	.113 (.039)
Turkish sample	.053 (.056)	.122 (.057)	.085 (.054)	.084 (.065)	.105 (.031)	.232 (.038)	.273 (.025)	.155 (.035)
Portuguese sample	-.014 (.037)	.044 (.038)	.098 (.041)	-.014 (.039)	.051 (.026)	.154 (.024)	.367 (.021)	. 020 (.022)
Multi-group GIWA-norm sample: model based on importance values of the overall sample								
British sample	-.007 (.034)	.121 (.037)	.117 (.049)	.052 (.039)	.083 (.000)	.177(.000)	.321 (.000)	.074 (.000)
Swedish sample	.017 (.052)	.074 (.049)	.089 (.064)	.025 (.049)	.083 (.000)	.177(.000)	.321 (.000)	.074 (.000)
Turkish sample	.021 (.035)	.072 (.035)	.094 (.041)	.027 (.035)	.083 (.000)	.177(.000)	.321 (.000)	.074 (.000)
Portuguese sample	.032 (.040)	.073 (.043)	.088 (.047)	.033 (.041)	.083 (.000)	.177(.000)	.321 (.000)	.074 (.000)

Table 8. *Standardized Estimates and Standard Errors of Main and Interaction Effects of Physical Subdomains on Physical Self-Worth and Global Self-Worth*

	Physical self-worth Estimate (S.E.)				Global self-worth Estimate (S.E.)			
	Body	*Sport*	*Condition*	*Strength*	*Body*	*Sport*	*Condition*	*Strength*
Condition 1*: one self-concept, one importance factor, one interaction—ESEM = CFA								
Self-concept	.900 (.018)	.650 (.034)	.884 (.028)	.594 (.035)	.725 (.023)	.602 (.039)	.652 (.034)	.472 (.039)
Importance	-.145 (.023)	**-.083 (.036)**	-.247 (.034)	-.129 (.036)	-.196 (.027)	-.264 (.041)	-.217 (.039)	-.253 (.040)
Interaction	.045 (.023)	.276 (.025)	.149 (.025)	.209 (.028)	.085 (.031)	.397 (.028)	.299 (.030)	.246 (.021)
Condition 2: all self-concepts, all importance factors, one interaction								
Self-concept	.673 (.044)	**.015 (.116)**	.365 (.123)	**.083 (.047)**	.579 (.049)	**.091 (.142)**	.243 (.157)	**.011 (.056)**
Importance	-.128 (.028)	**.089 (.047)**	**-.039 (.040)**	-.088 (.037)	-.127 (.034)	**-.082 (.060)**	**.083 (.058)**	-.163 (.052)
Interaction	.045 (.021)				.088 (.030)			
Self-concept	.686 (.043)	**.027 (.115)**	.355 (.125)	**.076 (.047)**	.597 (.049)	**.123 (.144)**	**.218 (.164)**	**-.009 (.055)**
Importance	-.141 (.028)	**.085 (.048)**	**-.042 (.041)**	**-.079 (.037)**	-.150 (.034)	**-.099 (.063)**	**.083 (.060)**	-.138 (.051)
Interaction		**.030 (.018)**				.123 (.026)		
Self-concept	.684 (.043)	**.018 (.118)**	.365 (.124)	**.079 (.047)**	.593 (.048)	**.099 (.144)**	.242 (.160)	**-.003 (.056)**
Importance	-.138 (.028)	**.091 (.047)**	**-.044 (.040)**	-.080 (.037)	-.142 (.034)	**-.075 (.060)**	**.080 (.059)**	-.144 (.051)
Interaction			**.023 (.019)**				.099 (.028)	
Self-concept	.685 (.043)	**.018 (.117)**	.363 (.123)	**.080 (.047)**	.594 (.047)	**.098 (.142)**	.233 (.159)	**.003 (.056)**
Importance	-.140 (.028)	**.092 (.048)**	**-.045 (.041)**	-.085 (.037)	-.153 (.035)	**-.073 (.062)**	**.081 (.060)**	-.163 (.052)
Interaction				**.012 (.021)**				.069 (.028)
Condition 3: all self-concepts, all importance factors, all interactions								
Self-concept	.677 (.039)	**.025 (.105)**	.357 (.117)	**.087 (.044)**	.588 (.048)	**.114 (.143)**	.227 (.161)	**-.002 (.058)**
Importance	-.124 (.029)	.086 (.039)	**-.038 (.041)**	-.087 (.039)	-.134 (.035)	**-.088 (.055)**	**.086 (.059)**	-.142 (.052)
Interaction	**.042 (.027)**	**.012 (.033)**	**.018 (.032)**	**-.016 (.026)**	**.042 (.039)**	**.063 (.065)**	**.067 (.055)**	**.000 (.039)**

Note. Non-significant effects in bold. *A similar pattern of estimates for the interaction effects was found with the LMS approach instead of the product indicators approach to latent interactions, although the estimates and standard errors were slightly different.
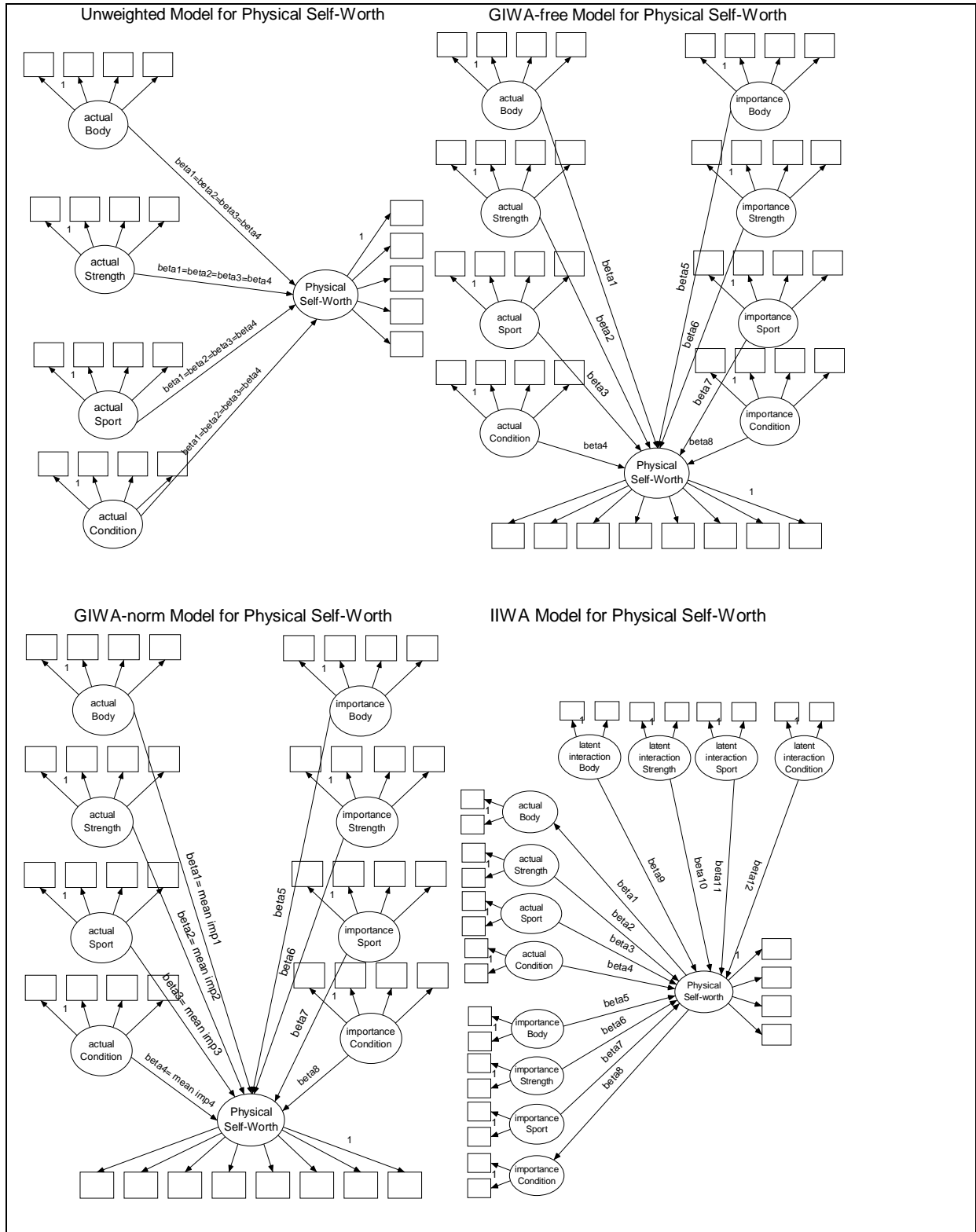
Figure 1. Theoretical models of importance. Here the Physical Self-Worth has been used as outcome

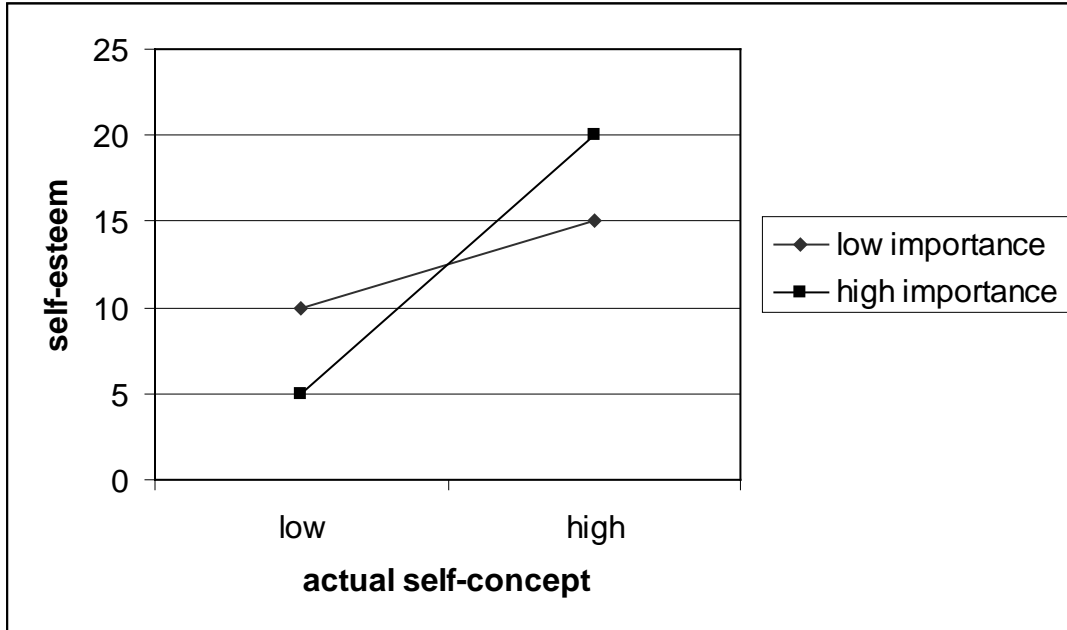variable. Corresponding models have been tested using Global Self-Esteem as outcome variable

Figure 2. Predictions based on the intra-individual model of importance