Administering self-concept interventions in schools: No training necessary? A meta-analysis

Alison J. O'Mara
SELF Research Centre, University of Western Sydney *a.omara@uws.edu.au*

Jasmine Green
SELF Research Centre, University of Western Sydney

Herbert W. Marsh
SELF Research Centre, University of Western Sydney

A meta-analysis of 105 studies reporting 152 self-concept interventions in school settings was conducted. The aims of the study were twofold: to explore the construct validity approach to self-concept interventions, and to examine aspects of the administration of the interventions, namely treatment setting, administrator type, administrator training, and implementation standardisation procedures. In support of the multidimensional perspective of self-concept, results of the random effects model analyses suggest that targeting specific self-concept domains when measuring self-concept outcomes lead to higher effect sizes (p < 0.001). Interestingly, the treatment setting, intervention administrator type, administrator training, and the use of standardisation procedures were not significant moderators. The implications for self-concept intervention administration in school settings are discussed.

Self-concept, self-esteem, meta-analysis, intervention, training

INTRODUCTION

Self-concept has long been considered an important factor for both the development and adjustment of children (Harter, 1990) plus a host of educational outcomes (for example, achievement; Marsh and Yeung, 1997). As such, many interventions have been designed to enhance the self-concepts of children and adolescents in educational settings. Indeed, positive self-concept development is frequently cited as a central goal of education. For instance, the Australian Education Council (1989), in a statement of the ten goals of education, established that schools should enable students to develop a positive self-esteem – a call that is also evidenced in Australian State/Territory syllabus documents (for example, Board of Studies New South Wales, 1998). Thus, many schools now implement at least informal self-concept enhancement programs; this trend is not restricted to Australia.

With the widespread implementation of self-concept enhancement programs in the schooling system, researchers have been keen to establish the most effective methods for achieving benefits. However, the self-concept intervention literature is plagued by inconsistent approaches – both conceptual and methodological. Conceptually, self-concept research has been clouded by the contradictions between unidimensional and multidimensional perspectives of the construct, and the consequent mismeasurement that often occurs due to the use of inappropriate self-concept instruments. Methodologically, self-concept enhancement researchers have presented a medley of program designs, making comparisons across research programs difficult using conventional literature review methods. As a result, it is still not clear which intervention practices are the most

effective, with confounding results suggesting both positive and negative effects of intervention programs.

Some of the problems in the self-concept intervention literature in a variety of settings have been highlighted through the use of meta-analysis (O'Mara, Marsh and Craven, 2004; Haney and Durlak, 1998; Hattie, 1992). Meta-analysis is the systematic synthesis of various studies on a particular research question (Rosenthal, 1984). It involves the calculation of effect sizes, which are a standardised metric for comparing the scores of treatment groups with control groups, and pre-test scores with post-test scores. Effect sizes with similar features are grouped together and compared; this allows the identification of moderator variables. Previous meta-analyses of self-concept enhancement programs have identified a range of moderator variables across a host of treatment settings in an attempt to unravel the intervention efficacy quandary and practical dilemma. That is, are self-concept interventions really effective? The current study examines some previously unexplored issues in relation to the administration of self-concept interventions specifically in school settings.

The Construct Validity Approach to Self-Concept Intervention Research

For many decades, self-concept was seen as a unidimensional construct, using global measures that tapped into an overall, general self-concept. However, more recently, self-concept has been increasingly considered a multidimensional construct, which differs across situations and contexts (Marsh and Craven, 1997). Thus, for example, one might have a high mathematics self-concept but a low reading or verbal self-concept. In response to this theoretical thrust towards a multidimensional conception of self-concept, which is frequently attributed to Shavelson, Hubner and Stanton's (1976) seminal work, measures have been developed that tap into a host of self-concept domains. One such instrument is the Self Description Questionnaire instruments (SDQ; Marsh, 1990a, 1990b), the adolescent version of which measures 11 distinct self-concept domains. These instruments have been well-validated using factor analysis (Byrne, 1996) and have been used in a host of applied research, thereby highlighting the importance of accounting for multiple domains of self-concept.

Despite the availability of sophisticated multidimensional instruments such as the SDQ, the self-concept intervention literature often neglects the multidimensionality of self-concept by measuring and reporting only global self-concept scores. This could potentially lead to an underestimation of the effectiveness of such interventions, since global self-concept instruments are unlikely to accurately assess the extent of change in self-concept after the implementation of an intervention targeting a specific self-concept facet (O'Mara, Craven, and Marsh, 2004). For instance, an intervention designed to enhance mathematics self-concept may not be adequately evaluated by a global scale, as global measures take into account facets of the self that are not likely to be affected by the mathematics-focused intervention.

The construct validity approach (Marsh and Craven, 1997; Marsh, 1993) suggests that domains of self-concept of highest relevance to those targeted in an intervention will be most impacted upon after administration of the intervention. O'Mara, Marsh, and Craven (2004) put this to the test in a comprehensive meta-analysis of self-concept interventions. They found frequent mismatch between the self-concept domain targeted in the intervention and that measured in the program evaluation phase. They also compared target self-concept domain outcomes with target-related and non-target self-concept domain outcomes. Target self-concept domains are those directly being enhanced in the intervention, target-related are those domains that are logically related to the intervention's goals, whilst non-target are self-concept domains completely unrelated to the goals of the intervention. For example, an intervention might be aimed at enhancing math self-concept and use a multidimensional instrument with several self-concept scales. The target self-concept domain outcome in this example would be the math self-concept score, while a target-

related domain might be general academic self-concept, and a non-target (unrelated) self-concept domain might be social self-concept. In this example, based upon the predictions of the construct validity approach, we would expect the largest change at post-test for math self-concept scores, a slight increase for the general academic self-concept scores and little or no change in social self-concept scores.

Consistent with the predictions of the construct validity approach, O'Mara, Marsh and Craven (2004) found that the self-concept domain/s directly targeted in an intervention showed greater improvement at post-test than those domains measured that were not directly targeted by the intervention. This supported the multidimensional perspective of self-concept, whilst also endorsing the use of domain-specific measures of self-concept when evaluating intervention effectiveness (Marsh, Craven and Martin, in press). The present paper will address this issue by utilising the construct validity approach employed by O'Mara, Marsh and Craven (2004), adopting the same terms: target, target-related, and non-target self-concept domain outcomes.

Differences in Intervention Administration

A further problem that makes comparison across self-concept interventions difficult is the differences in the way in which the interventions are designed and conducted. A number of issues have been examined elsewhere, such as the type of psychotherapy used (Hattie, 1992), the use of self-concept theory in developing the intervention (Haney and Durlak, 1998), and the effect of random group assignment on observed effect sizes (O'Mara, Marsh and Craven, 2004). However, the self-concept intervention literature has largely neglected considering the importance of treatment setting, administrator training, and implementation standardisation procedures on the effectiveness of self-concept interventions. Further, questions still remain as to the efficacy of various administrator types due to conflicting findings.

Previous meta-analyses of self-concept enhancement programs (O'Mara, Marsh and Craven, 2004; Haney and Durlak, 1998; Hattie, 1992) have not looked at whether the educational setting of an intervention is associated with higher or lower intervention success. This could have some important implications. For instance, very few self-concept interventions are conducted in early childhood settings (O'Mara, Marsh, Craven and Debus, in press), and yet pre-school age children have been found to exhibit self-concept differentiation (Marsh, Ellis and Craven, 2002; Marsh, Craven and Debus, 1991). However, the lack of interventions at such a young age may be justified, as Harter (1998, 1999) argues that the self-concepts of young children may be inaccurate, more polar than older children's, and more descriptive than representational. Thus, implementing interventions at this age may not be worthwhile given the lack of cognitive abilities to form realistic, representative self-concepts (Harter, 1999). The present study will therefore compare the relative effectiveness of interventions conducted in pre-school, primary and high school settings.

Another point of contention in the self-concept enhancement literature is the relative effectiveness of various change agents and the training they receive. Hattie (1992), in a meta-analysis of self-concept interventions, noted that teachers were less effective change agents in self-concept intervention programs. In contrast, O'Mara, Marsh and Craven (2004) found effect sizes for interventions conducted by mental health professionals were typically much lower than teachers, who were actually associated with the largest mean effect size. This may be partly attributable to the differences between the samples used in each meta-analysis. Hattie, for instance, included college students in the sample, while O'Mara, Marsh and Craven restricted the sample to high school or younger. Both meta-analyses included studies conducted in non-educational settings, which could also influence the comparison of administrator types. To extricate the differences in administrator effectiveness, the present study narrowed the inclusion criteria to only self-concept interventions conducted in school settings.

There is generally an implicit assumption that the training of intervention administrators is a desirable feature of any intervention. Indeed, in response to the Hattie's (1992) finding that teachers were the least effective administrators, Craven, Marsh and Burnett (2003) suggested that more comprehensive training of the intervention administrators might rectify this problem. This is a previously untested assumption, which will be explored as a potential moderator variable in the present meta-analysis.

As reiterated above, concerns have been raised as to the inconsistency with which self-concept intervention research has been conducted and reported. However, the focus was largely on between-study differences, such as different studies using different control group types or group assignment procedures (see O'Mara, Craven and Marsh, 2004). The moderators described above also tap into between-study differences. Within-study inconsistencies have received less attention. As such, the present study will attempt to explore the internal consistency of the interventions by including the use of standardisation procedures as a moderator variable. The use of standardisation procedures in an intervention refers to giving the same intervention to all participants, regardless of who the intervention leader is or at which point in time it is conducted. That is, we ask the question, "if I am receiving the treatment separately to you, are we still getting the same intervention?" Standardisation procedures could include a script that the intervention administrator must follow, or a set of detailed instructions. If some participants within an intervention are receiving more or less of the intervention than participants in other groups, this could affect the calculation of effect sizes.

METHOD

Sample of Studies and Coding Procedures

The sample of studies in the present meta-analysis was drawn from a larger database of self-concept interventions used by O'Mara, Marsh and Craven (2004). Full details of the sample of studies and coding procedures are described elsewhere (see O'Mara, Marsh and Craven, 2004), but are summarised briefly below.

The selection criteria were restricted to studies published in English between 1960 and 2000, involving children or adolescents in school settings, with a mean age of 18 or younger and a control group drawn from the same population as the intervention group. Studies also had to report self-concept or self-esteem scores. Studies were acquired primarily through searches of Psychological Abstracts, PsycINFO and ERIC databases, and through examination of the references of identified studies. The search yielded 105 studies reporting 152 self-concept interventions.

Coding categories included the relevance of the self-concept domain outcome to the intervention in order to establish construct validity; the setting of the intervention; treatment administrator type; training of the administrator; and the standardisation procedures of the intervention. Coding fidelity checks were in place, including comprehensive pilot coding, detailed coding manuals, and the employment of an independent second coder. This ensured that the coding between the two coders was satisfactorily consistent.

Computation of Effect Sizes

Of the 105 studies located, some had multiple treatment groups, and many measured more than one self-concept domain. Consequently, coding was conducted for 152 interventions comprising 366 unique effect sizes. Effect sizes were calculated for self-concept outcomes using the Comprehensive Meta-Analysis software program (Version 1.0.23; Borenstein and Rothstein, 1999). The program derives the effect size (d) using the standardised mean difference, which is

calculated by inputting either: means and standard deviations, *t*-values, or *p*-values. When post-test only scores are available, *d* is

$$d = (M_t - M_c) / SD_{\text{pooled}}$$

Where M_t = mean for the target intervention group,
   M_c = mean for the control group, and
   SD_{pooled} = pooled standard deviation based on both groups.

When pre-test and post-test scores are available for both groups, *d* is

$$d = [(M_{\text{post}} - M_{\text{pre}})_t - (M_{\text{post}} - M_{\text{pre}})_c] / SD_{\text{pre(pooled)}}.$$

Where $(M_{\text{post}} - M_{\text{pre}})_t$ = mean post-test score minus the mean pre-test score for the target intervention group,
   $(M_{\text{post}} - M_{\text{pre}})_c$ = mean post-test score minus the mean pre-test score for the control group, and
   $SD_{\text{pre(pooled)}}$ = pre-test pooled standard deviation based on both groups.

Alternative methods were used when means or SDs were not reported (cf. Lipsey and Wilson, 2001; Wolf, 1986). The sign of the difference was positive when the treatment group had a larger improvement than the control.

Analyses

The random effects method of moments model was used to conduct the meta-analysis. This includes a random error variance component that enables generalisation to the greater population of self-concept intervention studies (Raudenbush, 1994; for a beginner's summary of meta-analytic methods, see O'Mara, Marsh and Craven, 2005). An SPSS macro (Lipsey and Wilson, 2001) was used to calculate the mean effect size across all studies. Moderator analyses were conducted to test the ability of a single variable (e.g., administrator type) to explain the variability in the effect size distribution. Moderator analyses involve the calculation of the homogeneity of within-group and between-group statistics, Q_W and Q_B respectively, which is done using an analog to the ANOVA (Hedges and Olkin, 1985). A significant Q_B indicates that there is a significant difference between the groups. Another SPSS macro (Lipsey and Wilson, 2001) was used to run the moderator analyses.

Cooper's (1998) shifting unit of analysis technique was used to minimise violations of independence. For studies yielding multiple effect sizes, these effect sizes share many common features (for example., same intervention or administrator) and the effects are considered to be non-independent. In the shifting unit of analysis approach, the effect size for the multiple self-concept domains would initially be averaged to produce a single effect size for calculations involving the overall effect size for the sample. That is, the effect sizes are aggregated to produce one effect size for each intervention. For each moderator analysis, effect sizes are aggregated based upon the particular moderator variable, such that each study only includes one effect size per outcome on that particular variable. This approach helps to preserve as much of the data as possible (Cooper, 1998).

RESULTS

The overall mean effect size for the 152 self-concept interventions was 0.51, which suggests a moderate increase in self-concept for the treatment group at post-test compared to the control group.

Table 1 presents the findings from the moderator analyses testing the relevance of the self-concept outcome to the intervention, which taps into the construct validity approach to the

multidimensionality of self-concept. It was found that target self-concept domains yielded higher mean effect sizes ($d = 0.65$) than target-related outcomes ($d = 0.54$), and even higher still than non-target self-concept outcomes ($d = 0.22$). This moderator was statistically significant, as evidenced by the significant Q_B value ($Q_B (2) = 7.90$, $p < 0.02$).

Table 1. Results of the moderator analyses for the relevance of self-concept outcome

Homogeneity statistic	Q	df	p
Q_B	7.90	2	0.019
Q_W	260.41	182	< 0.001
SC Outcome Groups	**n**	**d**	**p**
Target SC outcome	92	0.65	< 0.001
Target-related SC outcome	51	0.54	< 0.001
Non-target SC outcome	42	0.22	0.091

Table 2 shows the results for the homogeneity analyses based upon the treatment setting. There was no statistically significant difference between the treatment settings, as evidenced by the non-significant Q_B value and the small differences between the mean d values for each group.

Table 2. Results of the moderator analyses for the different treatment settings

Homogeneity statistic	Q	df	p
Q_B	0.34	2	0.88
Q_W	176.89	149	0.06
Setting Groups	**n**	**d**	**p**
Early childhood school	4	0.54	0.21
Primary school	104	0.48	< 0.01
Secondary school	44	0.57	< 0.01

Table 3 illustrates the moderator analyses based upon the treatment administrator type. This moderator variable was not statistically significant ($Q_B (6) = 2.52$, $p = 0.87$), although both mental health professionals and professional trainees produced the lowest mean effect sizes ($d = 0.11$ and $d = 0.20$, respectively). Table 4 shows the results for the homogeneity analyses of the administrator training types. Again, the results are not statistically significant ($Q_B (6) = 2.38$, $p = 0.88$). Face-to-face training had a higher mean effect size ($d = 0.62$) than written training materials ($d = 0.34$), while programs that did not mention training at all had a moderate mean effect size similar to the overall mean effect size ($d = 0.57$ and 0.51, respectively).

Table 3. Results of the moderator analyses for the different treatment administrators

Homogeneity statistic	Q	df	p
Q_B	2.52	6	0.87
Q_W	178.77	145	0.03
Administrator Groups	**n**	**d**	**p**
Mental health professional	10	0.11	0.70
Professional trainee	13	0.20	0.03
Teacher	44	0.53	< 0.01
School counselor	24	0.54	< 0.01
Researcher	13	0.45	0.05
Mixed	29	0.58	< 0.01
Unknown	19	0.55	< 0.01

Finally, Table 5 demonstrates the results of the moderator analyses based upon the use of standardisation procedures during intervention implementation. The homogeneity results were just outside of reaching significance at the traditional 0.05 level ($Q_B (1) = 3.77$, $p = 0.052$), with

interventions employing standardisation procedures yielding higher mean effect sizes ($d = 0.66$) than those without ($d = 0.40$).

Table 4. Results of the moderator analyses for the various administrator-training procedures

Homogeneity statistic	*Q*	*df*	*p*
Q_B	2.38	6	0.88
Q_W	166.47	145	0.11
Training Type Groups	*n*	*d*	*p*
Written materials	11	0.34	0.21
Face-to-face	29	0.62	< 0.01
Other (incl. professional training)	19	0.47	0.02
Combination of above	37	0.38	< 0.01
Training mentioned, type not specified	5	0.77	0.05
Training stated as not necessary	1	0.42	0.62
Training not mentioned	50	0.57	< 0.01

Table 5. Results of the moderator analyses of the use of standardisation procedures during intervention implementation

Homogeneity statistic	*Q*	*df*	*p*
Q_B	3.77	1	0.052
Q_W	186.11	150	0.024
Standardisation Procedures Groups	*n*	*d*	*p*
Standardisation	65	0.66	< 0.001
No standardisation	87	0.40	< 0.001

DISCUSSION

The moderate positive overall mean effect size of 0.51 is promising indeed. This suggests that on average, most self-concept enhancement programs are producing benefits for the participants. A number of moderators were explored to see if they were associated with higher or lower mean effect sizes, including an examination of the construct validity approach to the multidimensionality of self-concept, and a host of moderators exploring the ways in which the interventions are administered.

The only statistically significant moderator of effect size was evidenced by the relevance of the self-concept outcome. That is, self-concept outcomes deemed a priori to be relevant to the intervention yield higher effect sizes than those related to or not at all targeted by the intervention. This supports the construct validity approach to intervention evaluation and the multidimensional structure of self-concept, and is consistent with the findings of O'Mara, Marsh and Craven (2004) and O'Mara et al. (in press), which used a larger database of studies not restricted to educational settings. It is also consistent with an expanse of research supporting the multidimensionality of self-concept in other research disciplines, such as personality (Marsh, Trautwein, Lüdtke, Köller and Baumert, 2003), adolescent mental health (Marsh, Parada and Ayotte, 2004), and developmental research (Crain, 1996). It is therefore recommended that self-concept intervention researchers adopt the well-established multidimensional perspective of self-concept by employing multidimensional measures such as the SDQ when evaluating intervention success. In this way, the true benefits of the intervention may be less likely to be concealed by global scales that fail to capture the impact upon specific self-concept domains.

Although not a significant moderator of effect size, interventions in secondary school settings are slightly more effective than in other settings. Interpretation of this variable is problematic in that

there could be a number of confounding variables. For instance, differences between settings could be due to the amount of time the participants are exposed to the intervention, as younger children are typically involved in less intensive programs. A host of other potential confounds could be considered, and the non-significant homogeneity results are likely to be a result of this. As meta-analytic procedures become more sophisticated, it may be possible one day to model covariates to tease out these effects. Additionally, more comprehensive reporting of the age breakdown in the interventions by researchers would assist meta-analysts in examining developmental moderators.

For the treatment administrator type, most types had relatively similar mean effect sizes, with the notable exception of mental health professionals and trainees, who both yielded lower mean effect sizes than other types. This is consistent with the findings of O'Mara, Marsh and Craven (2004). This could be due to the types of programs that such administrators are typically aligned with. For instance, it has been found elsewhere (O'Mara, Marsh, Craven and Debus, in press) that self-concept interventions using feedback are the most effective. Teachers are typically involved in programs implemented in the classroom such as attributional feedback or skills training programs, whereas mental health professionals were most commonly involved in programs involving counselling, and specialist programs such as those involving relaxation training programs. This might predispose teachers to be aligned with more effective treatments programs. In addition, it is feasible to assume that teachers might have more of a rapport with their intervention participants as they are exposed to the students for a greater amount of time, which could facilitate the participants' comfort with and involvement in the interventions. Thus, as with the treatment setting, future research could examine such covariates.

Of the specifically mentioned training types, face-to-face training (for example, didactic instruction, role-playing.) yielded the highest mean effect size ($d = 0.62$), although this variable was not found to be a significant moderator of effect size. This is interesting in that it suggests that optimising intervention impact may be achieved by face-to-face training of the treatment administrators, since the mean effect size for this category was higher than the overall mean effect size of 0.51. The caveat here is that the moderator analyses were not significant, so the differences between the training types were not consistent. Thus, if time or finances restrict, a lack of specific training of administrators is not necessarily harmful to the overall effectiveness of the intervention. Future meta-analyses could explore whether training type and intervention type has an interaction effect.

The homogeneity analyses for the standardisation procedures variable nearly reached significance, with interventions employing standardisation procedures yielding higher effect sizes. This could be interpreted to suggest that evaluation of the programs is more consistent when the programs are uniformly presented to the participants. By not ensuring the fidelity of the intervention administration, participants in different groups within the same intervention evaluation program might receive less effective enhancement simply because of a lack of strict adherence to the program. This could undermine the entire evaluation procedure. As such, we recommend that future interventions take care to employ standardisation procedures. Some examples might include providing a standardised format such as a previously published intervention, lesson plans, directions for the administrator, outlines of program goals, checklists for each treatment session, and so on. Many of these procedures are relatively inexpensive and easy to implement, and should be a part of all self-concept intervention programs.

Thus, it seems that the use of measures of self-concept relevant to the aims of the intervention is more important to the measurement of self-concept intervention effectiveness in school settings than features associated with the administration of the intervention, although such factors should not be neglected. In particular, the standardisation of implementation procedures seems to be a

reasonable way to promote more consistent results a self-concept intervention. And, although training of treatment administrators was not a significant moderator, face-to-face instruction is likely to be more effective. However, most importantly, the construct validity approach to the multidimensionality of self-concept has been supported once again, this time for a collection of interventions set exclusively in educational settings.

REFERENCES[1]

Australian Education Council. (1989). *The Common and Agreed National Goals of Schooling.* Canberra: AGPS.

Board of Studies New South Wales (1998). *Human Society and its Environment K-6.* Sydney: Board of Studies NSW.

Borenstein, M. and Rothstein, H. (1999). Comprehensive meta-analysis: A computer program for research synthesis (Version 1.0.23) [Computer software]. New Jersey: Biostat.

Byrne, B.M. (1996). *Measuring Self-Concept Across the Life Span: Issues and Instrumentation.* Washington, DC: American Psychological Association.

Cooper, H. (1998). *Synthesising Research: A Guide for Literature Reviews* (3rd ed.). Beverly Hills, CA: Sage.

Crain, R.M. (1996). The influence of age, race, and gender on child and adolescent multidimensional self-concept. In B.A. Bracken (Ed.), *Handbook of Self-Concept: Developmental, Social, and Clinical Considerations,* (pp.395-420). USA: Wiley.

Craven, R.G., Marsh, H.W. and Burnett, P. (2003). Cracking the self-concept enhancement conundrum: A call and blueprint for the next generation of self-concept enhancement research. In H.W. Marsh, R.G. Craven and D. McInerney (Eds.) *International Advances in Self Research.* (Vol.1, pp.91-126).). Greenwich, CT.: Information Age Press.

Haney, P. and Durlak, J.A. (1998). Changing self-esteem in children and adolescents: A meta-analytic review. *Journal of Clinical Child Psychology, 27*, 423-433.

Harter, S. (1990). Processes underlying adolescent self-concept formation. In R. Montemayor, G. Adams and T. Gullotta (Eds.), *From Childhood to Adolescence*. CA: Sage.

Harter, S. (1998). The development of self-representations. In W. Daman (Ed.) and S. Eisenberg (Vol. Ed.), *Handbook of Child Psychology* (5th ed., pp. 553–617). New York, NY: Wiley.

Harter, S. (1999). *The Construction of the Self: A Developmental Perspective.* New York, NY: Guilford Press.

Hattie, J.A. (1992). *Self-Concept.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Hedges, L.V. and Olkin, I. (1985). *Statistical Methods for the Meta-Analysis.* New York, NY: Academic.

Lipsey, M.W. and Wilson, D.B. (2001). *Practical Meta-Analysis.* Thousand Oaks, CA: Sage.

Marsh, H.W. (1990a). *Self Description Questionnaire (SDQ) I: A Theoretical and Empirical Basis for the Measurement of Multiple Dimensions of Preadolescent Self-Concept: A Test Manual and a Research Monograph.* Sydney: University of Western Sydney, Macarthur.

Marsh, H.W. (1990b). *Self Description Questionnaire (SDQ) II: A Theoretical and Empirical Basis for the Measurement of Multiple Dimensions of Adolescent Self-Concept: An Interim Test Manual and a Research Monograph.* San Antonio, TX: The Psychological Corporation (Republished in 1992, Publication Unit, Faculty of Education, University of Western Sydney, Macarthur).

Marsh, H.W. (1993). Academic self-concept: Theory measurement and research. In J. Suls (Ed.), *Psychological Perspectives on the Self.* (Vol. 4, pp.59-98). Hillsdale, NJ: Erlbaum.

---

[1] References used in the meta-analysis may be obtained from the first author by request.

Marsh, H.W. and Craven, R.G. (1997). Academic self-concept: Beyond the dustbowl. In G. Phye (Ed.). *Handbook of Classroom Assessment: Learning, Achievement and Adjustment.* USA: Academic Press.

Marsh, H.W., Craven, R.G. and Martin, A. (in press). What is the nature of self-esteem: Unidimensional and multidimensional perspectives. In M. Kernis (Ed). *Self-Esteem: Issues and Answers.* Psychology Press.

Marsh, H.W., Craven, R.G. and Debus, R.L. (1991). Self-concepts of young children aged 5 to 8: Their measurement and multidimensional structure. *Journal of Educational Psychology, 83,* 377–392.

Marsh, H.W., Ellis, L.A. and Craven, R.G. (2002). How do preschool children feel about themselves? Unravelling measurement and multidimensional self-concept structure. *Developmental Psychology, 38,* 376-393.

Marsh, H.W., Parada, R.H. and Ayotte, V. (2004). A multidimensional perspective of relations between self-concept (Self Description Questionnaire II) and adolescent mental health (Youth Self Report). *Psychological Assessment,* 16, 27-41.

Marsh, H.W., Trautwein, U., Lüdtke, O., Köller, O. and Baumert, J. (2003). Personality, and academic criteria: Multidimensional vs unidimensional perspectives of self-concept. SELF Research Centre, University of Western Sydney.

Marsh, H.W. and Yeung, A.S. (1997). The causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology, 89,* 41-54.

O'Mara, A.J., Craven, R.G. and Marsh, H.W. (2004, July). Unmasking the true effects of self-concept interventions and suggested guidelines for rectification. Paper presented at the Third International Biennial SELF Research Conference, Berlin, Germany.

O'Mara, A.J., Marsh H.W. and Craven, R.G. (2004, July). Self-concept enhancement: A meta-analysis integrating a multidimensional perspective. Paper presented at the Third International Biennial SELF Research Conference, Berlin, Germany.

O'Mara, A.J., Marsh, H.W. and Craven, R.G. (2005, December). Meta-analytic methods in educational research: Issues and their solutions. Paper presented at the AARE Conference, Parramatta, Australia.

O'Mara, A.J., Marsh, H.W., Craven, R.G. and Debus, R. (in press). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. Submitted for review.

Raudenbush, S.W. (1994). Random effects models. In H. Cooper and L.V. Hedges (Eds.), *The Handbook of Research Synthesis*, (pp.301-322). New York: Sage.

Rosenthal, R. (1984). *Meta-Analytic Procedures for Social Science Research (4ᵗʰ ed.)* (Vol. 6). CA: Sage.

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

Wolf, F.M. (1986). *Meta-Analysis: Quantitative Methods for Research Synthesis.* Beverly Hills, CA: Sage.