EDUCATIONAL PSYCHOLOGIST, *41*(3), 181–206

Do Self-Concept Interventions Make a Difference? A Synergistic Blend of Construct Validation and Meta-Analysis

Alison J. O'Mara and Herbert W. Marsh
SELF Research Centre
University of Western Sydney, Australia
Department of Educational Studies
University of Oxford, England

Rhonda G. Craven
SELF Research Centre
University of Western Sydney, Australia

Raymond L. Debus
Faculty of Education and Social Work
University of Sydney, Australia

Traditional reviews and previous meta-analyses of self-concept interventions have underestimated effect sizes by using an implicitly unidimensional perspective that emphasizes global self-concept. In contrast, this research employed a synergistic blend of meta-analysis and multidimensional construct validation to evaluate the impact of self-concept interventions for children in 145 primary studies (200 interventions). Overall, interventions were significantly effective (d = .51, 460 effect sizes). However, in support of the multidimensional perspective, interventions targeting a specific self-concept domain and subsequently measuring that domain were much more effective (d = 1.16), suggesting sole reliance on global self-concept is inappropriate for evaluating interventions designed to enhance a specific component of self-concept. Other moderators (e.g., feedback, experimental design, target population groups) also influenced effect sizes in ways useful to the design of new interventions. Methodologically, this research also demonstrates the use of both fixed and random effects models and incorporation of multiple outcomes from the same study.

As with many psychological constructs, self-concept suffers in that "everybody knows what it is", so that empirical researchers typically do not provide a theoretical definition of self-concept. Historically, self-concept research was dominated by a unidimensional perspective in which self-conceptions were seen to be relatively consistent and undifferentiated across social, academic, physical, and other domains (Byrne, 1984; Coopersmith, 1967; Wylie, 1989). Because these early researchers considered self-concept to be the sum total of an individual's self-perceptions (Marx & Winne, 1978), total self-concept or global self-esteem was typically represented by a single score.

CONSTRUCT DEFINITION OF SELF-CONCEPT

In their influential review of self-concept research and measurement, Shavelson, Hubner, and Stanton (1976) suggested that self-concept research had addressed the substantive "between network" issues of whether self-concept is related to other constructs (such as performance), before "within-network" problems of definition, measurement and construct interpretation had been resolved. Shavelson et al.

Correspondence should be addressed to Herbert W. Marsh, Department of Educational Studies, University of Oxford, 15 Norham Gardens, Oxford, OX2 6PY, England. E-mail: h.marsh@uws.edu.au

began by integrating various theoretical definitions of self-concept. Self-concept, according to their integrated definition, is a person's self-perceptions that are formed through experience with—and interpretations of—one's environment. These self-perceptions are influenced especially by evaluations by significant others, reinforcements, and attributions for one's own behavior. Shavelson et al. suggested that self-concept is both evaluative and descriptive, and that it has a multifaceted and hierarchical organization. Perceptions of personal behavior in specific situations are located at the base of the hierarchy, inferences about self in broader domains (e.g., social, physical, and academic) are at the middle of the hierarchy, while a global, general self-concept or self-esteem is found at the apex.

Although some researchers reserve the term *self-esteem* for the evaluative component of self-perception and use the term *self-concept* for descriptive components of self-perception, Shavelson et al. (1976; also see Marsh, 1993) argue that self-concept is both descriptive and evaluative. Thus, for example, statements such as "I am good at mathematics," "I can run a long way without stopping," and "I am good looking" all have both evaluative and descriptive components. Furthermore, typical application of the self-esteem construct in research based on instruments such as the Rosenberg scale (Rosenberg, 1979) emphasizes an overarching, general, or global construct that at least implicitly incorporates many (or all) specific components. In higher-order factor analysis studies (see Marsh & Hattie, 1996), global measures of self-esteem consistently correlated about .95 with the highest-order factor (representing the apex of the Shavelson et al. model) based on specific components of self-concept. It is important to note that multidimensional self-concept theorists do not deny the existence of an overarching, global self-concept or self-esteem. Indeed, multidimensional hierarchical models of self-concept integrate specific *and* global self-esteem dimensions of self-concept, such that global self-esteem is a component of the multidimensional self-concept structure (Marsh, Craven, & Martin, in press; Marsh, 1993). Hence, for the purposes of this study, we use the term "self-esteem" to refer to the global component of self-concept and to distinguish between this and specific components of self-concept (e.g., physical, social, academic).

Subsequent factor analytic research (e.g., Marsh, Byrne & Shavelson, 1988; Marsh & Hattie, 1996) showed that the hierarchical aspect of the multidimensional, hierarchical model proposed by Shavelson et al. (1976) was much weaker than originally hypothesized. In particular, correlations between specific components of self-concept were more differentiated and less highly correlated with each other than anticipated so that much of the variance in domain specific factors of self-concept could not be explained in terms of higher-order self-concept factors or self-esteem. Thus, for example, factor analysis of adolescent responses to a recent adaptation of the multidimensional SDQ III clearly supported the 17 self-concept factors that the instrument was designed to measure (Marsh, Trautwein, Lüdtke, Köller & Baumert, 2006). Further, the average correlation among the 17 self-concept factors—even after controlling for unreliability—was only .14, thus verifying the distinctiveness of these dimensions.

Consistent with a large amount of subsequent research in support of the Shavelson et al. (1976) model, Marsh and Craven (1997) argued for the importance of a multidimensional perspective of self-concept. They suggested that researchers should focus on specific components of self-concept most logically related to the goals of their particular research instead of, or in addition to, a global measure of self-esteem. Support for the construct validity of a multidimensional perspective of self-concept is particularly strong in educational psychology research, where academic outcomes are systematically related to academic components of self-concept, but nearly unrelated (or even negatively related) to nonacademic components of self-concept and to global self-esteem (Byrne, 1996; Marsh, 1993). For example, the pattern of domain specificity was clearly demonstrated in the Marsh, Trautwein et al. (2006) study, in which 17 components of self-concept were related to a set of eight academic outcomes (test scores, grades, and coursework selection in different school subjects). Although there were large, systematic relations between specific components of academic self-concept and corresponding academic outcomes (e.g., math self-concept correlated .77, .59, and .51 with math test scores, grades, and coursework selection, respectively), the academic outcomes were nearly unrelated to global self-esteem (*r*s ranging from –.03 to .05), as well as nine other nonacademic specific domains of self-concept.

We have focused specifically on a multidimensional perspective based on work stemming from Shavelson et al. (1976), but it is important to note that this perspective is also consistent with other self-concept theorists and theoretical work in related constructs. Bandura (1986), a proponent of domain specific approaches to self-beliefs, has criticized global self-concept measures. He has defined self-efficacy as the belief in one's capabilities to tackle and successfully execute even difficult tasks, and particularly the belief in one's ability to overcome barriers through one's own efforts. Self-efficacy typically is measured in a specific domain; this is consistent with our multidimensional perspective of self-concept. Whereas self-efficacy measures often refer to a more narrowly defined domain than self-concept measures, Marsh (1993) noted that this is not an inherent difference—self-concept can be defined in relation to very specific domains and self-efficacy can be defined more globally. Bandura (1986) emphasizes that self-esteem and self-concept—but not self-efficacy—are partly determined by "how well one's behavior matches personal standards of worthiness" (p. 410). In self-efficacy judgments, the focus of assessment is on the individual's capabilities in relation to the specific criterion items presented so that the influence of frame of reference effects is minimized (Marsh, Walker, & Debus, 1991). However, this distinction may not always be clear in measures that are given the labels of self-efficacy and self-concept. Importantly, in their comprehensive meta-analysis, Valentine, DuBois, and Cooper (2004) found no systematic difference in the ability of self-concept and self-efficacy measures to predict growth in academic achievement in longitudinal causal-ordering studies. In summary, a variety of theoretical perspectives emphasize the multidimensionality of self-belief constructs: the major focus of our meta-analysis.

The distinctive nature of the various domains of self-concept means that global self-esteem scales may conceal or distort self-judgment (Harter, 1990; Hattie, 1992). For example, if a child has a low reading self-concept and a high math self-concept, then a measure of global self-esteem is not useful as a diagnostic tool. This has ramifications for self-concept intervention research, as the under-use of multidimensional instruments may lead researchers to underestimate the effectiveness of their intervention (Bracken, 1996; Craven, Marsh, & Burnett, 2003).

THE IMPORTANCE OF SELF-CONCEPT ENHANCEMENT AND THE NEED FOR A META-ANALYTIC REVIEW

Self-concept is one of the oldest constructs in the social sciences (James, 1896) and widely acknowledged as an important construct, particularly in educational psychology (Marsh & Craven, 2006). Thus, enhancing self-concept is widely posited as a goal of education and has been associated with many educational benefits, such as improved academic achievement, persistence, coursework selection (Delugach, Bracken, Bracken, & Schicke, 1992; Marsh & Craven, 2006; Marsh & Yeung, 1997; Valentine et al., 2004), and approaches to learning (Burnett, Pillay, & Dart, 2003). Self-concept also mediates a host of other psychological and social outcomes as diverse as: healthy development of personal and social skills in children (Harter, 1990), coping skills (Shirk, 1988), social interaction (Gurney, 1986), general happiness (Harter, 1990), emotional adjustment and socialization (Donahue, Robins, Roberts, & John, 1993), cigarette smoking in teenage girls (Snow & Bruce, 2003), parent–adolescent relations (Barber, Hall, & Armistead, 2003), health-related physical activity in physical education classes (Marsh, Papaioannou, Theodorakis, 2006), and even winning gold medals at international swimming championships (Marsh & Perry, 2005). Indeed, according to the OECD, self-concept and related constructs are "closely tied to students' economic success and long-term health and wellbeing" (OECD, 2003, p. 9), which indicates the pervasive impact of positive self-concept development.

Acknowledgement of the importance of self-concept for manifold psychological, behavioral, and social outcomes has led to a proliferation of self-concept enhancement intervention research; unfortunately, optimal self-concept intervention strategies remain elusive. Traditional literature reviews have been unable to either elucidate or test promising directions in the self-concept enhancement literature because of the sheer volume and apparently contradictory nature of the literature (Marsh & Craven, 1997). A variety of research practices have been used in the self-concept enhancement literature, making comparisons across studies difficult. Furthermore, many of the older studies are based on outdated theoretical models, weak research designs, and poor measurement instruments (Hattie, 1992) and as such have not capitalized on recent advances in theory and research (Marsh & Hattie, 1996; Marsh, 1993, 1990). An important advance in methodology has been the construct validity approach to the study of self-concept intervention effects (Craven et al., 2003). Such an approach suggests that interventions should be evaluated as to their impact on: specific facets of self-concept directly related to the goals of the intervention, facets closely related to the target facets, and self-concept facets that are unrelated to the goals of the intervention (see Marsh, 1993; Marsh & Craven, 1997). In this article, these are referred to as target outcomes (e.g., math self-concept outcomes in an intervention intended to enhance math self-concept), target-related outcomes (e.g., academic self-concept might be logically related to math self-concept), and nontarget outcomes (e.g., social self-concept is not likely to be logically related to math self-concept).

While a number of meta-analyses have undertaken to test the impact of self-concept interventions, these meta-analyses have not applied this multidimensional, construct validity approach to the study of intervention effects. Indeed, the main meta-analytic studies in this area (i.e., Haney & Durlak, 1998; Hattie, 1992) have been based on effect sizes representing an average of all self-concept measures included in each study, regardless of whether or not the specific component of self-concept was relevant to the intervention. The justification for the practice was based primarily on an inability to analyze multiple outcomes in a statistically appropriate manner. However, this practice is antithetical to the primary goal of meta-analysis—to evaluate the generalizability of the findings. The essence of generalizability is the extent to which the results generalize over multiple outcomes, and this cannot be evaluated when results are based on an average of multiple outcome measures. This failure to account for the multidimensionality of the self-concept construct has contributed to the apparent contradictions, produced potentially misleading results, and has masked promising directions.

In our meta-analysis, we address this issue by formally introducing the notion of construct validity as a useful and necessary tool for meta-analysts and demonstrate how it can be used to interrogate data in relation to support for convergent and discriminant validity. The distinction between multidimensional and unidimensional approaches allows us to differentiate between outcomes in a way that was not possible in earlier meta-analyses of self-concept interventions. We aim to show that those outcomes most closely related to the aims of the intervention (target outcomes) have systematically larger effect sizes than would otherwise be apparent in an averaged effect size that included facets of self-concept not relevant to the intervention. In this way our research is an important contribution to methodological approaches to meta-analysis more generally and represents a blend of substantive and methodological contributions.

In addition to this innovative construct validity approach, we compare the results based upon fixed and random effects models to elucidate the nuances in these different statistical approaches. We focus on evaluating the effectiveness of self-concept interventions aimed at children up to the age of 18. The primary purpose of the study was to examine the impact of these interventions by identifying moderators and predictors of self-concept outcomes and to determine salient attributes of successful interventions that could be readily emulated to strengthen future research and practice. In summary, our overarching emphases are on the multidimensional structure of the self-concept construct consistent with recent advances in self-concept theory and the introduction of novel approaches to meta-analytic methods.

A MULTIDIMENSIONAL, CONSTRUCT VALIDITY APPROACH TO EVALUATING INTERVENTION STUDIES

The Shavelson et al. (1976) review proposed different ways of evaluating construct validity in self-concept research. Whereas most approaches to construct validity evaluate patterns of relations among different variables based on correlational data, Shavelson et al. (also see Marsh, 1993; Marsh, Debus, & Bornholt, 2005) emphasized that intervention studies with experimental and control groups are very useful in examining the construct validity of interpretations based on the intervention as well as interpretations of the self-concept responses. This multidimensional construct validity approach has been demonstrated in a series of studies based on the Outward Bound program. The Outward Bound standard course is a 26-day residential program based on physically and mentally demanding outdoor activities (Marsh, Richards, & Barnes, 1986a, 1986b). Consistent with the primarily nonacademic goals: (a) gains were significantly larger for the self-concept scales predicted a priori to be most relevant to the goals of the program, compared to less relevant self-concept scales; (b) the effect sizes were consistent across 27 different Outward Bound groups conducted by different instructors at different times and in different locations, and (c) the size and pattern of the gains were maintained over an 18-month follow-up period. In contrast, the Outward Bound bridging course is a 6-week residential program designed to produce significant gains in the academic domain for underachieving adolescent males through an integrated program of remedial teaching, normal schoolwork, and experiences likely to influence particularly academic self-concept (Marsh & Richards, 1988). Consistent with the primarily academic goals of the bridging course: (a) academic self-concept effects were substantial and significantly larger than nonacademic self-concept effects; and (b) there were also corresponding effects on reading and math achievement.

If global self-esteem alone had been measured in these two Outward Bound studies, the interventions would have been judged much weaker, and the richness of matches between specific intended goals and actual outcomes would have been lost. The juxtaposition of these two interventions and their contrasting predictions provides a powerful demonstration of the importance of a multidimensional perspective of self-concept and the usefulness of a construct validity approach in support of the multidimensional perspective. The inclusion of self-concept scales that are less relevant to the aims of a particular intervention also provides "control scales" to evaluate counter-interpretations of the intervention effects (e.g., placebo effects, halo effects, and demand characteristics; see Marsh, Richards, & Barnes, 1986a, 1986b). In the present investigation, we adapt this multidimensional construct validity approach to the evaluation of results of a meta-analysis of self-concept intervention studies.

DIRECT AND INDIRECT INTERVENTIONS

A conceptual issue of relevance to the present meta-analysis is the varied use of direct and indirect self-concept interventions. Some studies directly aim to enhance self-concept through activities designed to foster improved self-conceptions (e.g., the Developing Understanding of Self and Others package; see Morse, Bockoven & Bettesworth, 1988), while others indirectly aim to enhance self-concept through skill building (e.g., a reading ability peer-tutoring program). This raises the question of whether directly focusing on self-concept in the intervention will lead to larger increases in self-concept at post-test than indirect (skill building) interventions.

Self-concept theory and research clearly supports a reciprocal relation between self-concept and performance (Marsh & Craven, 2006). Based on this reciprocal effects model, it follows that self-concept interventions can enhance both the targeted self-concept facets and related performance outcomes. Therefore, enhancing self-concept along with enhancing performance adds value beyond skill training alone. Particularly in educational research, there is clear evidence from longitudinal causal ordering studies (e.g., Marsh, Byrne, & Yeung, 1999; Marsh & Craven, 2006) and a recent meta-analysis of this research (Valentine et al., 2004) that prior levels of academic self-concept lead to higher levels of subsequent academic achievement beyond what can be explained by prior levels of academic achievement. Because this research demonstrates that there are reciprocal relations between academic self-concept and subsequent achievement, Marsh and Craven argue that researchers and practitioners should aim simultaneously to improve both academic self-concept and academic skills. For example, if practitioners enhance self-concepts without improving performance, then the gains in self-concept are likely to be short-lived. Conversely, if practitioners improve performance without also fostering participants' self-beliefs in their capabilities, then the performance gains are unlikely to be long lasting. Worse yet, interventions aimed at enhancing performance may unintentionally undermine self-concept in ways that will eventually undermine the short-term gains in performance. For example, Marsh and Peart (1988) found that both competitive and cooperative interventions led to positive short-term gains in physical fitness. Whereas the cooperative intervention also led to enhanced self-concept, the competitive intervention undermined the self-concept needed to sustain ongoing involvement in physical activity. An important implication for the present investigation is that self-concept can be enhanced directly by targeting a specific component of self-concept or indirectly by focusing on skill-development in a related area of achievement. A critical question addressed in our meta-analysis is the relative effectiveness of these different approaches.

Consistent with our emphasis on a multidimensional perspective of self-concept and our focus on academic components of self-concept in educational research, a key strength of the Valentine et al. (2004) meta-analysis was the incorporation both of global measures like self-esteem and more specific measures, such as academic self-concept. Results of this earlier meta-analysis provide clear support for the reciprocal effects model when self-beliefs are based on measures logically related to the corresponding measures of achievement (e.g., math self-concept and math achievement), but not when self-beliefs are based on global measures (like self-esteem). Interestingly, although Valentine et al. considered different domain-specific self-belief constructs (e.g., self-concept and self-efficacy), they found no differences between the different self-belief constructs. Thus, the domain specificity of the measures seemed to be the critical aspect, at least in terms of predicting growth in academic achievement and performance. Hence, recent advances in self-concept research (see review by Marsh & Craven, 2006) provide strong support for both a multidimensional perspective of self-concept and the reciprocal relations between specific components of self-concept and corresponding measures of achievement. As such, the present meta-analysis will examine the differential impact on self-concept of direct and indirect interventions within the context of a multidimensional approach.

PREVIOUS META-ANALYSES OF SELF-CONCEPT INTERVENTIONS AND HOW THEY DIFFER FROM OUR INVESTIGATION

Two previous meta-analyses of self-concept interventions set the stage for the present investigation. Hattie's (1992) meta-analysis of 89 studies conducted prior to 1983 focused primarily on how intervention type influenced the effectiveness of an intervention. Hattie found a mean effect size of .37 based on a total of 485 effect sizes. Hattie found no significant differences between studies that directly targeted self-concept, those that were indirect (e.g., academic achievement programs), those that were a combination of direct and indirect methods, and those for which self-concept was simply an incidental measure. Hattie also found that mean effect sizes (d) were higher for groups with previously diagnosed problems (d = .55) relative to groups without pre-existing problems (d = .26). Four studies reported 36 effect-sizes measured at delayed posttest. A significant difference was present between studies that were followed up (d = .16) and those without follow-up (d = .40). These results suggest that the effect size decreases over time, which some researchers believe may be explained by the dissipation of initial euphoria over time (Marsh et al., 1986a, 1986b) but is a typical finding in much educational intervention research.

In a subsequent meta-analysis of 102 self-concept intervention studies conducted prior to 1992, Haney and Durlak (1998) sought to elucidate interventions that led to significant improvements in self-concept, improvements in self-concept that were associated with other desirable outcomes, and factors that moderate outcome success. Of particular note, they found that treatments specifically focused on self-concept enhancement (direct self-concept enhancement studies) were significantly more effective than treatments aimed to impact on self-concept indirectly, through an associated construct such as social skills (indirect studies). The mean effect size for studies focused on enhancing self-concept was .57, while the mean effect size for interventions focusing on other outcomes was .10. Interventions were also categorized as to whether they were treating a pre-existing problem (treatment studies, d = .47) or maintaining "normal" levels of self-concept (preventive studies, d = .09). Nonrandomized designs resulted in significantly lower effect sizes (d = .04) than randomized studies (d = .38). Studies with control groups that received no treatment had significantly higher effect sizes (d = .34) than studies with placebo controls (groups that were provided with alternative, nontreatment programs designed to account for extraneous benefits as a result of interest or excitement induced by participation in a new program; d = .10). In contrast to Hattie (1992), Haney and Durlak found no significant change in self-concept at follow-up, although only five self-concept interventions conducted and reported follow-up results.

Importantly, there were key methodological limitations in both these previous meta-analyses of self-concept intervention research. Both the Hattie (1992) and Haney and Durlak (1998) studies used an implicit unidimensional perspective in which effect sizes for each study were averaged to form a single effect per study—an average across potentially diverse components of self-concept for those studies that had collected multiple dimensions of self-concept. This was partly due to the fact that most of the studies they synthesized were conducted before the multidimensional perspective of self-concept was popularized. This implicitly unidimensional approach, as argued above, is counterproductive in relation both to prevailing self-concept theory and to the general goal of meta-analysis of evaluating support for the generalizability of results across multiple outcomes and the construct validity of interpretations of the results. In contrast, our study will specifically focus on the multidimensionality of self-concept that is a critical feature of our construct validity approach to meta-analysis.

Furthermore, the present study includes a more in-depth analysis of the intervention literature through the consideration and analysis of moderators that affect the features of the interventions themselves, not just the research methods used (as in Haney & Durlak, 1998). Additionally, the study utilizes more sophisticated meta-analytic methods and presents both fixed and random effects meta-analytic models (discussed in more detail below), thereby addressing the concerns underpinning advances in meta-analytic techniques (see Raudenbush, 1994). Although we focus discussion of methodological approaches to meta-analyses on self-concept enhancement research, as noted previously, these issues have potentially important implications for meta-analytic research more generally.

Fixed and Random Effects Models of Meta-Analysis

Traditionally, meta-analyses have utilized fixed effects models, which have an underlying assumption that any random error in the effect size distribution is derived only from participant-level sampling error (Hedges & Vevea, 1998). However, some meta-analysts argue that this is an unrealistic assumption and that features of the studies themselves (such as treatment duration) are further sources of variance (Raudenbush, 1994). The random effects model attempts to account for variance beyond the participant-level error and is thus more generalizable to the larger population of intervention studies (Raudenbush, 1994).

When results are homogeneous (i.e., the dispersion of the effect sizes around their mean is no greater than that brought about by sampling error), fixed and random effects analyses produce similar results (Hedges & Vevea, 1998). A heterogeneous distribution implies that the variability of the effect sizes is larger than that expected by sampling error and that the differences between effect sizes have some other cause. If the assumption of homogeneity is not met, fixed effects models may underestimate and random effects models may overestimate the error variance (Overton, 1998). Further, the models support different inferences: The results of fixed effects models are only applicable to the set of studies used in the analysis, whereas random effects models allow inferences to be drawn to studies beyond the sample used (Cohn & Becker, 2003; Hedges & Pigott, 2001; Hedges & Vevea, 1998; Valentine et al., 2004). Given that the different models support different inferences about intervention success, an increasingly common practice in meta-analysis is to present analyses for both fixed and random effects models (e.g., see Kumkale & Albarracin, 2004; Shadish, Matt, Navarro, & Phillips, 2000; Valentine et al., 2004). While previous meta-analyses of the self-concept intervention literature have utilized only fixed effects analyses, this article follows the more defensible trend of presenting both fixed and random effects models.

THIS INVESTIGATION

This study was designed to critically analyze the impact of interventions on self-concept, with an overarching goal of extracting meaning from the self-concept enhancement literature. By approaching the intervention quandary from multiple angles (i.e., from the perspectives of intervention design, research design, and meta-analytic methodology), this study aims to:

1. Implement our new, multidimensional construct validity approach to meta-analyses of intervention effects by assessing whether multidimensional components of self-concept most logically related to the intervention (target outcomes) result in higher effect sizes than components of self-concept not logically related to the aims of the intervention (nontarget outcomes).
2. Assess the benefits of interventions directly targeting self-concept (direct interventions) in comparison to interventions that aim to improve self-concept through indirect means or for which self-concept is incidental to the aims of the study.
3. Critically evaluate specific strategies and characteristics of interventions (e.g., identifying who is the most effective administrator of interventions, comparing preventive and treatment interventions) to add to our understanding of how best to enhance self-concept.
4. Evaluate the impact of design features (in particular randomization and control group type) on observed effect size;
5. Investigate whether benefits achieved at post-test are maintained at follow-up testing.
6. Contrast the results based upon both random effects and fixed effects models in the analyses.

METHOD

Sample of Studies

The selection criteria used by Haney and Durlak (1998) were adopted, as far as possible, to facilitate comparisons with the earlier study. This meant that inclusion was restricted to studies published in English, that involved participants with a mean age of 18 or younger, and that included a control group drawn from the same population as the intervention group. To further maintain consistency with Haney and Durlak, studies were selected that included a measure of self-concept or another related self-concept construct (e.g., self-esteem, self-efficacy), which could be either a global measure (e.g., self-esteem) or a specific domain (e.g., academic self-concept).

However, we did diverge from the previous meta-analysis in several important ways. Whereas Haney and Durlak did not differentiate between indirect and incidental studies (see definitions below), we have expanded their selection criteria to include four main types of interventions:

1. Self-concept interventions in which both the aim and the procedure used were directly geared towards enhancing self-concept (direct studies).
2. Interventions with a principal focus on constructs other than self-concept (e.g., academic achievement), but which state an assumption that self-concept will be impacted upon and which include an outcome measure of self-concept (indirect studies).
3. Interventions in which self-concept is entirely incidental in the study. That is, self-concept enhancement was not an aim of the intervention, but the study contains an outcome measure of self-concept (incidental studies).
4. Combination interventions, which use both direct and indirect methods.

In addition, unlike the Haney and Durlak (1998) meta-analysis, we excluded studies using inferred self-concept measures (measures of one's self-concept that are actually assessed by someone else, such as a teacher or a parent). This decision follows from the original Shavelson et al. (1976) construct definition of self-concept in which they distinguished between self-concepts based on a person's own self-perceptions and inferred self-concepts that are based on inferences about the person made by another person. Consistent with this distinction, Marsh, Debus, and Bornholt (2005) argued that self-concept must be based on one's own self-perceptions and that responses by significant others represent a distinct construct that has theoretically important relations with self-concept (e.g., Marsh & Byrne, 1993; Marsh & Craven, 1991).

Ninety-six of the 102 studies used in the Haney and Durlak (1998) meta-analysis were thus included in the present investigation. Studies in the Hattie (1992) meta-analysis and the references of identified studies were also inspected for suitability. We included studies published between 1958 and 2000, whereas Hattie (1992) and Haney and Durlak (1998) only sampled up to 1983 and 1992, respectively. Additional studies were identified primarily through searches of Psychological Abstracts, Expanded Academic ASAP, Social Sciences Index, PsycINFO, and ERIC databases. These sources were searched using combinations of the key words *self-concept*, *self-esteem, self-efficacy, self-worth, self-attributions,* or *self-image*, and *youth, child, children, adolescent,* or *teenage*. The vast majority of studies reported scores for global self-esteem or domains of self-concept, with three studies reporting self-efficacy scores, zero reporting self-worth scores, zero reporting self-attributions, and three reporting self-image outcomes. The resulting sample consisted of 145 appropriate studies.

Coding Procedures

A code sheet and codebook were developed to capture information about potential moderator variables. The coding scheme was based on that used by Haney and Durlak (1998), but additional coding categories were devised to reflect the multidimensional perspective.

Two stages of coding reliability checks were conducted. The first stage consisted of extensive pilot testing with two coders, which entailed discussions over disparity in coding and subsequent amendments to the code sheet and codebook. One coder then coded all of the studies, while a second coder coded a random selection of 52 articles (excluding studies used in the pilot rounds). The double-coded items resulted in a mean agreement rate of 92.7% across all items. The percent agreement between the coders on items regarding the intervention's aim was 93.96% (SD = 7.69), while agreement rates for the relevance of the self-concept domain outcome were even higher for target (97.21%, SD = 6.50), target-related (97.79%, SD = 4.23), and nontarget (97.63%, SD = 2.49) outcomes.

Computation of Effect Sizes

Of the 145 studies located, some had multiple treatment groups, and many measured more than one self-concept domain. Consequently, coding was conducted for 200 interventions, 20 of which included follow-up testing. From these interventions, 501 unique effect sizes were generated: 460 of these effect sizes for post-test results, 41 effect sizes for follow-up testing.

Effect sizes were calculated for self-concept outcomes using the Comprehensive Meta-Analysis software program (version 1.0.23; Borenstein & Rothstein, 1999). The program calculates effect sizes from a variety of data input types, including means, standard deviations, p-values, and t-values. When group means and standard deviations were provided, the formula for the effect size (d) was $d = (M_t - M_c) / SD_{pre}$, where M_t = mean for the direct intervention group at posttest, M_c = mean for the control group at posttest, and SD_{pre} = pooled standard deviation. Some effect sizes were based on pretest–post-test differences between the treatment and control groups, such that the effect size represented a difference of differences. The 41 follow-up effect sizes were based on the difference between post-test and follow-up of comparisons between treatment and control groups. Alternative methods were used when other values were reported; for example, F-values were converted to t-values by taking the square-root of the F-value (see Lipsey & Wilson, 2001; Wolf, 1986). As in Haney and Durlak's (1998) meta-analysis, the effect size was set at zero if interventions reported nonsignificant findings, and no other applicable data were supplied. Once the unadjusted effect size was calculated, the raw effect sizes were imported into SPSS. In SPSS the effect sizes were adjusted using Hedges' correction for small sample bias and weighted by the inverse of their variance, which gives more weight to larger samples (Lipsey & Wilson, 2001).

Unit of Analysis

As most of the primary studies contributed more than one effect size to the analyses, it was necessary to adopt a shifting unit of analysis approach (Cooper, 1998). This approach minimizes violations of assumptions about the independence of effect sizes, while preserving as much of the data as possible (Cooper, 1998). For example, a study aimed at social self-concept might report both social self-concept (target outcome) and math self-concept (nontarget outcome). In this case initially the effect size for the two self-concept domains would be averaged to produce a single effect size for calculations involving the overall effect size for the sample. Thus, once the effect sizes were aggregated to produce one effect size for each intervention, the overall mean effect size for the post-test analyses was based on 145 effect sizes, and the overall mean effect size for the follow-up analyses was based on 20 effect sizes. However, in the moderator variable analyses, specific domains of self-concept targeted by the intervention would be examined separately such that for the example described, the study would contribute one effect size to each of the applicable levels ("target" and "nontarget") of that particular moderating variable. This means that the number of effect sizes in the homogeneity analyses depended upon the moderator variable being examined. For instances where there were multiple effect sizes for a particular moderator, and also to account for the fact that some of the 145 studies had more than one intervention utilizing the same control group, the weighting of any study that contributed more than one effect size was divided by the number of effect sizes for that study. This conservative approach reduces potential inflation of the contribution of multiple-effect size studies to the analyses, while still capturing the information for those outcomes.

Analyses

In this study both fixed and random effects models were used. An SPSS macro capable of computing both fixed and random effects models (see Lipsey & Wilson, 2001) was used to calculate the mean effect size and the 95% confidence interval across all studies. The confidence interval tests the precision of the estimate of the mean effect size, as well as allowing a test of significance of the effect size. A mean effect size was considered to be significant when its 95% confidence interval did not contain zero (Hunter & Schmidt, 1990). The confidence interval is expected to be smaller in fixed effects models (Hedges & Vevea, 1998). The homogeneity of within-group and between-group statistics, Q_W and Q_B, respectively, were calculated using an analog to ANOVA (Hedges & Olkin, 1985). The analog to ANOVA conducted in the homogeneity analyses tests the ability of a single variable (e.g., control group type) to explain the variability in the effect size distribution. Another SPSS macro (see Lipsey & Wilson, 2001) was used to test homogeneity of the sample.

However, homogeneity tests do not allow the testing of multiple moderators simultaneously. This can be achieved through a weighted multiple regression (Lipsey & Wilson, 2001), which is designed to help resolve the problem of confounded moderators. This is important because such confounds can make it difficult to establish whether a moderator has a significant independent relation to the effect size (Lipsey, 2003). Random and fixed effects weighted regression analyses were conducted to establish predictive ability across the moderators. A two-step process for the regression analyses that has been used in previous meta-analyses (e.g., Stice & Shaw, 2004) was implemented. First, the variables that were deemed to be heterogeneous in the Q tests were included in a method of moments mixed effects (i.e., random effects) weighted regression using the SPSS macros (see Lipsey & Wilson, 2001). Categorical variables were dummy coded, with a reference category excluded from the analyses as is done in regular multiple regression, and the regression coefficients (B) produced in the regression analyses were automatically adjusted for the other variables in the regression model by the macro (Lipsey & Wilson, 2001). Second, variables that were significant in the univariate models were included in a multivariate model. Because of the potential for multicollinearity when all coding variables were included in the multivariate regression analysis, a backward elimination process was used, deleting nonsignificant predictors step-by-step until only significant predictors remained.

In summary, the analyses followed the following steps:

1. *Both* fixed and random effects homogeneity tests were conducted on all relevant variables.
2. If significant heterogeneity was present in step 1, the significantly heterogeneous variables were subjected to *both* fixed and random effects univariate regression. Variables *without* significant heterogeneity were not subjected to further analysis because there was no variance unaccounted for.
3. The categories found to be *significant* for each variable in the fixed univariate regressions (conducted in step 2) were then included in the fixed multivariate regression model; the categories found to be *significant* for each variable in the random univariate regressions were included in the random multivariate regression model.

RESULTS

General Descriptive Results

Participants. Responses were based on a total of 16,900 individual participants. Forty-one percent of the studies targeted the middle childhood age group (6–10 years), 32% targeted pre-adolescents (11–12 years), 22% targeted adolescents (13–18 years), and 4.5% targeted early childhood (3–5 years). Different studies reported age in different ways; some reported the mean age, others an age range, while still others only noted the grade or year in school of the participants. Where the mean age of the participants was not provided, the mean age of each intervention's sample was inferred by determining the midpoint of the range of ages given. In supplementary analyses, a meta-analytic regression was conducted to determine if age was a significant predictor of effect size. The results were found to be nonsignificant, suggesting that developmental trends are unlikely to have a significant impact upon effect size.

There was a great deal of missing data on the variables sex, SES, and ethnicity of the participants. For sex, 93 of the 145 studies reported the number of males, with a mean of 47 male participants (minimum 3–maximum 506); 83 reported the number of females, with a mean of 47 female participants (minimum 2–maximum 491). For SES, 55.2% of studies did not mention the SES of the participants, with 17.2% of studies drawing participants from low SES backgrounds, 15.9% from medium SES, 1.4% from high SES backgrounds, and 10.3% stating that participants were mixed in terms of SES. For ethnicity, 60.7% of studies did not report the ethnic constituencies of the sample, 14.5% of the samples were from the ethnic majority, 11% from an ethnic minority, and 13.8% from highly mixed ethnic groups.

TABLE 1
Descriptive Statistics of the Reviewed Studies at the Study Level

Variable	*n*	%
No. of interventions per study		
One	104	71.7
Two	29	20.0
Three	11	7.6
Five	1	0.7
Publication date		
1958	1	0.7
1960–1969	7	4.8
1970–1979	35	24.1
1980–1989	68	46.9
1990–1999	29	20.0
2000	5	3.4
Conducted follow-up testing—Yes	12	0.1
Rationale provided (≥1 could be selected)		
Self-concept theory	37	25.5
Nonself-concept theory	30	20.7
Previous research findings	121	83.4
Hypothesis generated by author	52	35.9
No rationale provided	8	5.5
Control group type		
No treatment (inactive)	83	57.2
Waitlist	8	5.5
Placebo	29	20.1
Other or combination	25	17.2
Randomization		
Random individual	75	51.0
≥15 groups randomly assigned	5	3.4
Other random	8	5.5
Intact groups randomly assigned	12	8.3
Intact groups without explicitly stated random assignment	23	15.9
Matching	10	6.9
Self-selection	4	2.8
Criterion-based assignment (e.g., low self-concept)	8	5.5
Reliability of self-concept measure		
Published measure	125	86.1
Unpublished, reliability reported	7	4.9
Unpublished, reliability not reported	13	9.0

Note. $N = 145$.

Studies. The key descriptive statistics of the 145 studies are summarized in Table 1, the treatment settings in Table 2, and the characteristics of the interventions (e.g., skills training, use of praise, etc.) in Table 3. The mean duration of the interventions was 12.7 weeks (SD = 16.03); the mean number of treatment sessions was 21.16 (SD = 29.74); and the mean length of each session was 62.23 minutes (SD = 51.70). As an auxiliary analysis, tests of the homogeneity of the effect sizes by decade were conducted to see if the decade in which the studies were published moderated effect size. These effects were found to be nonsignificant, indicating that there were no significant differences between studies published in different decades.

The effect sizes ranged from $d = -2.26$ to $d = 4.50$, where a positive d indicates that the treatment group scored higher than the controls. The mean effect size for the 145 studies was highly significant according to both the fixed-effects model ($d = .31$, $p < .001$; 95% confidence interval = .28 to .35) and the random effects model ($d = .51$, $p < .001$; 95% confidence interval = .38 to .64). The meta-analytic model fit statistics for the fixed effects model indicated heterogeneity Q (144, k = 145) = 1668.34, $p < .001$, suggesting that systematic effect size variability remained to be accounted for. The presence of significant heterogeneity suggests the need for caution in generalizing results of the fixed effects beyond the studies analyzed in the present meta-analysis, and against placing too great confidence in the results of the fixed effects model. This warning is applicable to all subsequent analyses because the effect size distribution in the fixed effects models failed to meet homogeneity assumptions in all the instances discussed below. Heterogeneity in the model also demonstrated that it was necessary to examine the moderators that influenced the effect sizes across different interventions.

TABLE 2
Setting in Which the Treatment Was Conducted

Treatment Setting	*Frequency*	*%*
Early childhood school	4	2.0
Primary school	104	52.0
Secondary school	44	22.0
Home	2	1.0
Mental health or psychology–psychiatry clinic	5	2.5
General hospital	1	0.5
Residential treatment centre (psychiatric or special school)	5	2.5
Combination of at least two of the above	7	3.5
Other	9	4.5
Unspecified	19	9.5

TABLE 3
Frequencies of the Primary Characteristics of the Intervention Programs

Characteristics of Intervention	*Mean Effect Size*	*Count*	*% Responses*
Assertiveness training	0.282	4	1.3
Cooperative learning	0.320	5	1.6
Discussion	0.521	20	6.5
Environmental change	0.043	15	4.9
General self-concept activities	0.648	18	5.8
Group counseling	0.438	63	20.5
Individual counseling	0.801	12	3.9
Modified curriculum (e.g., new math curricula)	0.492	7	2.3
Motivational training	0.442	2	0.6
Participation in a new project developed for study	0.514	18	5.8
Peer tutoring	0.239	4	1.3
Physically oriented (e.g., fitness programs)	0.336	13	4.2
Practice on established tasks	0.184	7	2.3
Praise–feedback	0.897	27	8.8
Relaxation training	0.585	7	2.3
Self-concept training in specific facets	2.344	2	0.6
Skills training (e.g., problem-solving, decision-making exercises, or both)	0.445	51	16.6
Social support restructure	0.417	11	3.6
Training in new tasks	0.334	22	7.1
Total responses	—	308	100.0

Note. Some—but not all—interventions concurrently implemented two of the above characteristics; for example, administering praise–feedback within the same program as skills training. As such, the total count or number of responses ($n = 308$) is greater than the number of interventions ($n = 200$). *Count* reflects how many times a particular treatment characteristic was utilized across treatment programs, whereas *% responses* indicates each characteristics' count as a percentage of the total count.

Mean Intervention Effect Sizes and Their Moderators

The following section presents a summary of the meta-analytic results for each examined moderator. The weighted mean effect size for the moderator categories are reported in Table 4, and the homogeneity statistics for each moderator are presented in Table 5. The results of the fixed and random effects multivariate regression analyses can be found in Table 6.

Match between intervention's aims and self-concept measurement instrument. In both the random and the fixed effects models homogeneity analyses, it was found that the greater the congruence between the aims of the intervention and the self-concept domains actually considered, the larger the mean effect size (see Table 4). Interventions that targeted a specific self-concept domain and measured that specific domain yielded the highest effect sizes of all (fixed effects $d = .76$; random effects $d = 1.16$). An example of this type of intervention is an intervention seeking to enhance reading self-concept and then administering a scale with a measure of reading self-concept. Interventions that did not adequately match the intervention to the outcome measures in terms of multidimensionality had substantially lower effect sizes. In particular, a specific self-concept dimension intervention (such as one to improve reading self-concept) assessed using a global measure of self-concept (a total score or a self-esteem scale) yielded fixed effects $d = .21$ and random effects $d = .41$. This is an example of an intervention that implicitly assumes a multidimensional conception of self-concept by targeting a specific self-concept domain other than global self-concept, but that is internally inconsistent by measuring the outcomes using a global (unidimensional) self-concept scale. In supplemental analyses we found no statistically significant differences in effect sizes for studies based on a global self-esteem scale (as in the widely-used Rosenberg, 1965, self-esteem instrument) and those based on a total score based on diverse self-concept items (as in the Coopersmith, 1967, instrument). This provides empirical support for the theoretically based decision to include both as indicators of a global measure of self-concept. However, the match between focus of the intervention on a particular self-concept dimension and the outcome did result in a significant difference between levels of this variable for both fixed and random effects models (see Q_B in Table 5).

The univariate regression analyses were significant for both fixed and random effects models (fixed effects Q_R (10) = 89.16, $p < .001$; random effects Q_R (10) = 20.48, $p < .05$), therefore the "match" moderator was included in both fixed and random effects multivariate regression analyses. After backward elimination in the fixed effects multivariate regression, 7 of the 11 categories were found to be significant predictors (see Table 6). For the random effects model, only the group of studies focusing on a specific self-concept domain that included a measure of that domain ($B = .78$, $SE = .17$, $p < .001$), and the group of studies in which a specific self-concept domain was targeted but measured using multiple multidimensional scales ($B = .43$, $SE = .18$, $p < .01$), were significant predictors of effect size (see Table 6).

TABLE 4
Mean Effect Sizes for Categories Within Each Moderator

Moderator	Fixed Effects *d*	Fixed Effects 95% CI	Random Effects *d*	Random Effects 95% CI	*k*
Match between intervention and outcome measure[a]					
Unidimensional–global conception of SC in intervention with global SC score	0.18	0.12, 0.25	0.35	0.13, 0.57	58
Unidimensional–global conception of SC in intervention with multiple multidimensional SC measurements	0.30	0.10, 0.49	0.37	–0.15, 0.89	10
Multidimensional conception of SC in intervention; intervention focused on single specific domain of SC (e.g., social SC) with specific domain measured	0.76	0.59, 0.92	1.16	0.74, 1.59	19
Multidimensional conception of SC in intervention; intervention focused on single specific domain of SC with multiple multidimensional SC measurements	0.69	0.50, 0.87	1.00	0.56, 1.44	17
Multidimensional conception of SC; intervention focused on specific domain of SC with unidimensional measurement	0.21	0.11, 0.31	0.41	0.09, 0.74	27
Multidimensional conception of SC in intervention; intervention focused on multiple domains of SC with specific facets measured but not clearly related to intervention	0.14	–0.02, 0.29	0.23	–0.22, 0.68	14
Multidimensional conception of SC in intervention; intervention focused on multiple domains of SC with specific facets measured and clearly related to intervention (whether other nontarget facets included)	0.44	0.37, 0.51	0.46	0.14, 0.78	24
SC as incidental to the intervention; only a measure of SC	0.23	0.08, 0.37	0.45	0.05, 0.85	18
Relevance of SC outcome					
Target SC domain outcome	0.49	0.44, 0.54	0.55	0.40, 0.70	121
Target-related SC domain outcome	0.11	0.05, 0.17	0.49	0.30, 0.68	72
Nontarget SC domain outcome	0.08	–0.02, 0.18	0.21	–0.04, 0.45	49
Focus of intervention on SC					
Intervention directly targets SC	0.88	0.79, 0.97	0.67	0.38, 0.96	31
Intervention indirectly targets SC	0.18	0.13, 0.23	0.48	0.33, 0.64	109
SC incidental to study	0.11	0.02, 0.20	0.27	0.01, 0.53	38
Combination of direct and indirect methods	0.25	0.16, 0.34	0.49	0.15, 0.83	22
Prevention versus treatment					
Preventive intervention	0.27	0.23, 0.31	0.42	0.26, 0.59	98
Treatment intervention	0.33	0.33, 0.48	0.54	0.36, 0.71	99
Combination preventive and treatment	0.21	–0.09, 0.51	0.31	–0.60, 1.23	3
Treatment characteristics					
Group counseling or discussion	0.33	0.22, 0.44	0.33	0.06, 0.60	28
Individual counseling with or without group counseling	0.48	0.31, 0.65	0.68	0.25, 1.11	11
Group counseling or discussion with skills training	–0.25	–0.36, –0.14	0.28	–0.02, 0.58	24
Practice or training	0.18	0.09, 0.27	0.42	0.19, 0.65	39
Praise–feedback	1.90	1.76, 2.04	1.13	0.71, 1.55	14
Physically oriented	0.13	0.01, 0.25	0.44	0.02, 0.86	11
SC activities	0.30	0.16, 0.45	0.67	0.29, 1.05	16
Social support or environmental restructure	0.30	0.14, 0.47	0.34	–0.12, 0.80	10
Praise with other features	0.86	0.61, 1.12	0.86	0.40, 1.31	14
Social support or environmental restructure with other features	0.16	0.06, 0.27	0.19	–0.20, 0.59	14
Peer tutoring–cooperative learning–modified curriculum	0.15	–0.04, 0.34	0.31	–0.21, 0.84	8
Other	0.24	0.13, 0.35	0.34	–0.07, 0.75	11
Treatment administrator					
Mental health professional	–0.39	–0.52, –0.25	0.23	–0.20, 0.67	15
Professional trainee	0.21	0.09, 0.33	0.41	0.03, 0.78	20
Teacher	0.64	0.57, 0.72	0.53	0.30, 0.76	53
School counselor	0.43	0.30, 0.56	0.53	0.20, 0.87	24
Experimenter–researcher	0.17	0.03, 0.30	0.53	0.12, 0.94	16
Other nonprofessionals	0.31	0.16, 0.46	0.35	–0.16, 0.86	9
Mixed administrators	0.22	0.14, 0.30	0.46	0.19, 0.72	38
Unspecified	0.24	0.15, 0.34	0.53	0.19, 0.87	25
Control group type					
Inactive control group	0.34	0.29, 0.38	0.47	0.32, 0.62	127
Waitlist control group	0.45	0.30, 0.60	0.60	0.27, 0.92	30
Placebo control group	0.27	0.18, 0.36	0.46	0.22, 0.71	52
Other control group	–0.02	–0.15, 0.10	0.12	–0.48, 0.73	7

(*continued*)

TABLE 4 (Continued)

Moderator	Fixed Effects *d*	Fixed Effects *95% CI*	Random Effects *d*	Random Effects *95% CI*	*k*
Group assignment procedure					
Random assignment: Individual	0.39	0.33, 0.46	0.58	0.42, 0.73	106
Random assignment: Group	1.04	0.92, 1.15	0.41	–0.12, 0.93	7
Random assignment: Other	0.24	–0.00, 0.49	0.43	–0.04, 0.90	14
Intact groups: Randomly assigned to conditions	0.32	0.20, 0.43	0.75	0.37, 1.12	15
Intact groups: Not randomly assigned to conditions	0.14	0.07, 0.21	0.35	0.08, 0.63	29
Nonrandom assignment: Matching	0.25	0.10, 0.41	0.18	–0.26, 0.62	12
Nonrandom assignment: Self-selection	–0.61	–0.76, –0.46	–0.31	–1.03, 0.41	4
Nonrandom assignment: Criterion	0.21	–0.03, 0.45	0.18	–0.28, 0.63	13

Note. k = number of cases in the particular category; SC = self-concept.
[a]Categories with small cell sizes pertaining to this variable are not presented here.

TABLE 5
Homogeneity Results for Fixed and Random Effects Models

Variable	Fixed Effects Q_B	Fixed Effects Q_W	Random Effects Q_B	Random Effects Q_W
Multidimensional match between intervention and outcome measure	89.16***	1635.08***	20.48*	193.88
Relevance of the self-concept outcome	108.94***	1632.83***	5.53	258.58
Focus of the intervention on self-concept	206.79***	1516.45***	4.11	229.14
Prevention versus treatment	10.14**	1713.11***	0.99	219.18
Treatment characteristics	622.53***	1100.72***	19.88*	255.77***
Treatment administrator	201.55***	1521.70***	2.07	221.20
Type of control group	31.92***	1693.65***	1.82	223.11
Group assignment procedure	340.58***	1382.67***	12.49	233.13*

*$p < .05$. **$p < .01$. ***$p < .001$.

In the fixed effects multivariate regression analyses, it can be seen that interventions focusing on multiple domains of self-concept were negative predictors of effect size relative to other characteristics included in the multiple regression. This result may seem contradictory to the homogeneity analyses (Table 4), which revealed that the effect sizes of these categories were significantly positive, but less positive than other categories representing the match between the intervention and the outcome measure (e.g., a multidimensional conception of self-concept in the intervention that focused on a specific self-concept domain that was measured). That is, interventions that targeted multiple domains of self-concept were associated with a less positive increase in self-concept at post-test, even when multidimensional measures of self-concept are used. However, the results of the homogeneity analyses (Table 4) and multivariate regression analyses (Table 6) are not necessarily inconsistent. These findings are likely due to the fact that the effect sizes included in these categories would be a mixture of target, target-related, and nontarget self-concept outcomes. For instance, a study targeting both math and verbal self-concept might use the Self-Description Questionnaire, which measures 11 different self-concept domains. Obviously, only 2 of those 11 domains would be target self-concept domains, and yet for this analyses the effect sizes would be grouped together to form the mean effect size of all 11 scores, thereby losing valuable information. This further emphasizes the need to target and measure specific self-concept domains distinctly. The finding that the negative *B* value for multiple self-concept interventions measured using a global scale was lower than using a multidimensional scale supports this assertion. This argument also touches at the nerve of the meta-analytic assumption of independence of data points. By aggregating the data to minimize such violations, we risk losing insight into the complexity in the data (Cooper, 1998). Nevertheless, because the effect of interventions focusing on multiple domains of self-concept was not significant in the random effects model, the generalizability of this finding should be viewed cautiously.

Relevance of self-concept domain outcome. Both fixed and random effects models found that facets of self-concept that were unrelated to the intervention (nontarget outcomes) had a lower effect size (fixed effects $d = .08$; random effects $d = .21$) than target outcomes (fixed ef-

TABLE 6
Fixed and Random Effects Multivariate Regression Coefficients and Standard Errors Predicting Effect Size Remaining After Backwards Elimination

Variable	*B*	*SE*
Fixed Effects Regression Analyses		
Unidimensional–global conception of SC in intervention with global SC score	–0.26	.05
Unidimensional–global conception of SC in intervention with multidimensional SC measured and referred to, but treated, as a global score	–0.77	.21
Multidimensional conception of SC in intervention; intervention focused on single specific domain of SC (e.g., social SC) with specific domain measured	0.54	.09
Multidimensional conception of SC in intervention; intervention focused on single specific domain of SC with multiple multidimensional SC measurements	0.21	.10
Multidimensional conception of SC in intervention; intervention focused on multiple domains of SC with specific facets measured and but not clearly related to intervention	–0.25	.08
Multidimensional conception of SC in intervention; intervention focused on multiple domains of SC with specific facets measured and clearly related to intervention (whether other nontarget facets included)	–0.27	.06
Multidimensional conception of SC in intervention; intervention focused on multiple domains of SC with multidimensional SC measured but treated as a global score (e.g., averaged, totaled)	–0.57	.14
Target SC domain outcome	0.32	.07
Target-related SC domain outcome	0.35	.07
Intervention directly targets SC	0.29	.07
Intervention indirectly targets SC	–0.16	.05
Preventive intervention	–0.10	.05
Treatment characteristics: Individual counseling with or without group counseling	0.42	.09
Treatment characteristics: Praise–feedback	1.30	.08
Treatment characteristics: Praise with other features	0.51	.12
Treatment characteristics: Other	0.21	.08
Treatment administrator: Mental health professional	–0.20	.10
Treatment administrator: School counselor	0.19	.07
Treatment administrator: Mixed	–0.23	.06
Placebo control group	–0.30	.05
Random assignment: Individual	0.80	.12
Random assignment: Group	0.91	.13
Random assignment: Other	0.96	.17
Intact groups: Randomly assigned to conditions	0.66	.14
Intact groups: Not randomly assigned to conditions	0.72	.12
Nonrandom assignment: Criterion based	0.42	.16
Nonrandom assignment: Matching	0.31	.14
Random effects regression analysis		
Intervention focused on specific domain of SC, with specific domain measured	0.78	.17
Intervention focused on specific domain of SC with multiple multidimensional SC measurements	0.43	.18
Treatment characteristics: Praise–feedback	0.59	.19

Note. Regression coefficients are unstandardized and are the amount of change in the effect size associated with one unit change in the predictor. All variables presented in the table are significant at the $\alpha = .05$ level. SC = self-concept.

fects $d = .49$; random effects $d = .55$; see Table 4). Q_B was significant in both models, indicating significant differences between the categories (see Table 5).

The univariate regression analyses showed that the relevance of self-concept domain outcome was a significant predictor of effect size in the fixed effects model (fixed effects Q_B (2) = 108.94, $p < .001$), but not in the random effects model (Q_B (2) = 5.53, $p > .05$). As such, this moderator was only included in the fixed effects multivariate regression analyses. For the fixed effects multivariate regression, both target outcomes and target-related outcomes were significant ($p < .001$; see Table 6).

Focus of the intervention on self-concept. The focus of the intervention on self-concept had a similar pattern of results in the two models (see Table 4). Direct interventions, in which the intervention implemented was directly aimed at increasing self-concept, exhibited the largest mean effect size (fixed effects $d = .88$; random effects $d = .67$). Indirect interventions that aimed at increasing self-concept through non–self-concept programs, such as social skills training, had a mean effect size of .18 in the fixed effects and .48 in the random effects models. Interventions that had a combination of features (direct and indirect) had a fixed effects d of .25 and a random effects d of .49, while interventions for which self-concept was an incidental variable had the lowest effect sizes in both models (fixed effects $d = .11$, random effects $d = .27$). The fixed effects analyses indicated significant between-group differences for different levels of this variable, but between-group differences were not significant for the random effects model (see Table 5).

In the fixed effects univariate regression, this moderator accounted for significant variability in the effect sizes (Q_R (3) = 206.79, $p < .001$) and so was included in the fixed effects multivariate regression. In the fixed effects multivariate regression, direct studies ($B = .29$, $SE = .07$) and indirect studies ($B = –.16$, $SE = .05$) were significant predictors of effect size (see Table 6). Incidental studies were not significant predictors of effect size in the fixed effects multivariate regression and were removed during backward elimination. The random effects model univariate regression was not significant (Q_R (3) = 4.11, $p > .05$), and so the variable was not included in the random effects multivariate regression.

Preventive versus treatment interventions. In both random and fixed effects models, interventions whose participants had pre-diagnosed problems (fixed effects $d = .33$ and random effects $d = .54$) yielded larger effect sizes than prevention interventions (fixed effects $d = .27$ and random effects $d = .42$; see Table 4). Q_B was significant for the fixed effects analyses, but not the random effects model.

In the fixed effects univariate regression, the moderator explained significant variance in the effect sizes (Q_R (2) = 10.14, $p < .01$), and so it was included in the multivariate regression. In the fixed effects multivariate regression (see Table 6), prevention interventions were significant predictors of effect size ($B = –.10$, $SE = .05$, $p < .05$). For the random effects univariate regression, between-group differences were not statistically significant (Q_R (2) = .99, $p > .05$), and so this moderator was not included in the random effects multivariate regression.

Treatment characteristics. The interventions were also examined on the basis of features of the treatment. Table 3 shows the frequencies of the primary characteristics of the intervention programs. Across the 200 interventions, 20 different types of intervention features were coded, and more than one type could be selected for each intervention. For example, an intervention could incorporate both peer tutoring and the introduction of a new curriculum. As a result, there were numerous combinations of intervention characteristics, making cell sizes for each distinct type very small. For the purpose of analysis, categories reflecting the most commonly recurring characteristics or combinations of characteristics were devised, which resulted in 12 broad categories (see Table 4).

Differences between groups associated with this variable were significant for both random and fixed effects models (see Q_B values in Table 5). The most common form of intervention focused on practice or training in an established task, which had a fixed effects d of .18, and a random effects d of .42. In both fixed- and random-effects models, interventions that primarily utilized praise and/or feedback had the highest mean effect size (fixed effects $d = 1.90$, random effects $d = 1.13$). The 27 interventions categorized as being principally a praise-based or feedback-based used a variety of methods; these were attributional feedback, goal feedback, general praise, self-praise, contingent praise on target facets, and noncontingent praise on target facets. Several of these 27 interventions reported the use of more than one of these methods. In a supplementary analysis of the praise and feedback interventions that used one of these methods exclusively ($n = 14$), interventions using attributional feedback yielded the highest mean effect sizes ($d = 1.52$, $n = 3$), followed by contingent praise ($d = .76$, $n = 4$), goal feedback ($d = .63$, $n = 4$), and noncontingent praise ($d = .31$, $n = 3$).

The fixed effects univariate regression indicated that the moderator "treatment characteristics" was a significant predictor of effect size ($Q_R(11) = 622.53$, $p < .001$) and was thus included in the fixed effects multivariate regression. Four of the 12 categories were found to be significant predictors when included in the backward multivariate regression (see Table 6). The largest predictor category of this moderator was for studies utilizing praise/feedback ($B = 1.30$, $SE = .08$, $p < .001$). The random effects univariate regression also yielded a positive Q_R-value ($Q_R(11) = 19.88$, $p < .05$) and was therefore included in the random effects multivariate regression analyses. After backward deletion in the random effects multivariate regression, only the group of studies from this moderator that used praise/feedback techniques remained ($B = .59$, $SE = .19$, $p < .01$; see Table 6).

Treatment administrator. Differences between groups associated with treatment administrator type were statistically significant for the fixed-effects model, but not for the random-effects model (see Q_B values in Table 5). For the fixed effects model, the highest mean effect size was for the group of studies with teachers administering the intervention ($d = .64$), followed by school counselors ($d = .43$; see Table 4). The lowest mean effect size in the fixed effects model was for studies with a mental health professional administering the intervention (e.g., psychiatrist, psychologist, social worker, etc.; $d = –.39$). For the random effects model, studies with teachers, school counselors, the researchers themselves, and "unspecified" treatment administrators had approximately the same mean effect size ($d = .53$).

In the fixed effects univariate regression, a significant model ($Q_R(7) = 201.55$, $p < .001$) meant that this moderator was included in the multivariate regression model. In the fixed effects multivariate regression, only the categories of mental health professional ($B = –.20$, $SE = .10$), school counselor ($B = .19$, $SE = .07$), and mixed treatment administrators ($B = –.23$, $SE = .06$) were significant predictors of effect size (see Table 6). For the random effects univariate regression, a nonsignificant model ($Q_R(7) = 2.07$, $p > .05$) precluded the moderator from the random effects multivariate regression.

Control group type. Differences between groups associated with "control group type" were statistically significant for the fixed effects model, but not for the random effects model (see Q_B values in Table 5). In the fixed effects analysis, the waitlist control group had the highest mean effect size (.45), followed by inactive controls (d = .34) and then placebo control groups (d = .27). Although the ordering of the groups was similar in the random effects model, the differences between groups were not significant (see Table 4).

The model for the fixed effects univariate regression was significant ($Q_R(3) = 31.92$, $p < .001$), leading to its inclusion in the fixed effects multivariate regression. Of the categories in this moderator, only placebo controls were significant predictors of effect size, with a negative influence (B = 150>.30, SE = .05; see Table 6). The univariate regression for the random effects model was not significant ($Q_R(3) = 1.82$, $p >$ .05), so this moderator was not included in the random effects multivariate regression.

Group assignment procedure. There were a number of types of random assignment methods (e.g., randomly assigned individuals and randomly assigned groups) and several different nonrandom assignment techniques (e.g., matching group, and self-selection); these are summarized in Table 4. While the random assignment strategies consistently produced higher mean effect sizes than nonrandom assignment procedures, these differences were statistically significant for the fixed-effects model but not for the random-effects model (see Q_B values in Table 5).

In the fixed effects univariate regression, this moderator was a significant predictor of effect size ($Q_R(7) = 340.58$, $p <$ .001). In the fixed effects multivariate regression, 7 of the 8 categories for the group assignment procedure were found to be significant predictors of effect size, with the strongest predictive category being 'other' types of random assignment (B = .96, SE = .17, $p < .001$; see Table 6). Because the random effects univariate regression yielded a nonsignificant model, ($Q_R(7) = 12.49$, $p > .05$), this moderator was not included in the random effects multivariate regression.

Multivariate Regression Models

The categories that reached significance in the fixed and random effects multivariate regression are discussed in the section above. However, this analysis does not tell us much about the model as a whole. For the fixed effects multivariate regression, R^2 = .432. In other words, 43.2% of the variance in the effect sizes was explained by the final model produced after backward elimination. The model, however, had significant residual (error) variance ($Q_E(230) = 1152.06$, $p < .001$), indicating error beyond the subject-level sampling error, which remains unaccounted for across effect sizes. This suggests that a random effects model regression should be used (Lipsey & Wilson, 2001).

The random effects multivariate regression produced an R^2 of .110. As in the fixed effects multivariate regression, Q_E was significant ($Q_E(196) = 294.05$, $p < .001$), indicating that additional variables are needed to explain differences in the effect sizes.

Stability of Effects Over Time

Twelve studies comprising 20 interventions reported both post-test and follow-up outcomes (N = 20 effect sizes). The mean effect size between post-test and follow-up was calculated to determine whether effect sizes systematically decreased, increased, or remained the same in the period following the treatment. Both the fixed effects and random effects models produced nonsignificant results (fixed effects d = .08, SE = .08, $p > .05$; random effects d = .01, SE = .11, $p >$ .05). This suggests that the benefits attained at post-test did not significantly decrease or increase between post-test and follow-up testing. However, there was a moderately positive correlation (r = .44) between effect size and time lapsed between post-test and follow-up (3 weeks to 14 months), suggesting a possible sleeper effect in which the size of the intervention grew larger over time.

DISCUSSION

Our meta-analysis advances knowledge about self-concept and its enhancement in valuable ways. It is promising to see the overall positive effectiveness of the interventions and that these effects do not systematically diminish over time. Further, while the results of the regression analyses suggest that there is still unaccounted for variance, our results are a significant advance on previous meta-analyses, as a number of moderators of effect size were found. The most notable outcome, as emphasized in our multidimensional construct validity approach, was the support for the multidimensional perspective of self-concept. This was reflected in the finding that intervention effects were substantially larger for facets of self-concept that were logically related to the intervention than for unrelated facets of self-concept. More generally, our findings may help to maximize the benefits of future interventions and are discussed below, where we place particular emphasis on the random effects results that are more generalizable to other research than results based upon fixed effects.

Theoretical Issues That Make a Difference: A Multidimensional Perspective of Self-Concept Enhancement

From the perspective of self-concept theory, the main intent and the most important contribution of our meta-analysis is to provide further support for the multidimensional perspective of self-concept. This is particularly relevant in that a unidimensional perspective—reliance on a single global self-esteem scale or a total averaged score across diverse components of self-concept—continues to be widely used in intervention research (Marsh & Craven, 1997). We argue that this has systematically undermined estimates of the effect of an intervention on self-concept. Unlike other meta-analyses that consider global self-concept, self-esteem, or related self-belief constructs, our meta-analysis incorporates a multidimensional perspective.

The most pronounced finding in this study is that the largest mean effect size of all the moderator categories was for interventions aimed at enhancing a specific self-concept facet and that also measured that specific self-concept domain (d = 1.16; from the moderator "match between intervention's aims and self-concept measurement instrument"). Given that the overall (random effects) mean effect size was d = .51, this is indeed a substantial finding, demonstrating that effective self-concept enhancement interventions need to account for the multidimensionality of the self-concept construct. The multivariate regression analyses also provided support, whereby the strongest positive predictor was an intervention with a well-matched specific self-concept scale (B = .78, SE = .17). Interventions targeting a specific self-concept domain, but which used multiple multidimensional scales, yielded moderate predictive regression coefficients in the random effects multivariate regression analyses (B = .43, SE = .18). Further, in the homogeneity analyses, interventions that measured target self-concept domains had higher effect sizes in comparison to studies where facets of self-concept measured were not clearly related to the goals of the intervention. These findings imply that using unidimensional scales for assessment may result in inaccurate evaluation of the impact of interventions on self-concept and an underestimation of the actual benefits of the intervention.

Support for our multidimensional construct validity approach has both theoretical and practical implications. In terms of theory, the finding that distinct domains of self-concept were differentially affected by the interventions supports the multidimensional structure of self-concept. As mentioned above, it was found that target domain outcomes (e.g., math self-concept in a study targeting math self-concept) were higher than target-related self-concept domains (e.g., general academic self-concept), but the difference was not large (.55 and .49 for target and target-related domains, respectively). However, these effects were both much larger than the effects on nontarget self-concept domains (e.g., social self-concept), which had a mean effect size of .21. This suggests that there may be some correlation between gains in related self-concept domains but very weak relations between gains in unrelated domains—which provides strong support for a prediction of Shavelson et al.'s (1976) hierarchical model of self-concept. We do, however, offer a caveat to the interpretations of these results. A growing body of research shows that different specific components of academic self-concept are very distinct. Hence, for example, verbal and math self-concepts are nearly uncorrelated (for further discussion, see Marsh, 1986; see also Marsh & Hau, 2003, for a discussion of the internal–external frame of reference effect). Hence, we would not expect that an intervention designed to enhance math self-concept would *necessarily* have a positive effect on verbal self-concept. Analysis of this issue was beyond the scope of the present meta-analysis due to the small number of studies reporting specific subject self-concept domains, with even fewer studies reporting more than one specific subject self-concept domain. This highlights the need for more intervention studies that are evaluated with a systematic profile of multidimensional self-concept scales in different domains to provide more fine-grained tests of the domain specificity of interventions on different self-concept domains.

From a practical standpoint, the findings from this study provide strong support for the usefulness of the multidimensional, construct validity approach to the study of intervention effects when undertaking both primary self-concept enhancement studies and meta-analyses synthesizing the self-concept literature. We therefore advocate that researchers should measure both target and nontarget facets of self-concept when evaluating the impact of self-concept enhancement interventions to fully explore and demonstrate the tangible benefits of the intervention. This requires the use of multidimensional measures of self-concept when assessing intervention success, with particular attention to including scales that specifically measure the domain being targeted. The importance of the match between self-concept measures and the intervention is also related to earlier warnings by Pajares (1996; but also see Marsh, Roche, Pajares, & Miller, 1997) about the match between self-efficacy constructs and validity criteria such as test scores to which they are related. Indeed, personality researchers have long recommended a context-specific measurement approach, particularly when dealing with self-reports (Mischel, 1977). Additionally, these results suggest that developing interventions that target specific facets of self-concept could strengthen self-concept enhancement intervention design and practice and produce greater effects. Thus, we strongly urge the use of context-relevant multidimensional measures of self-concept and specific self-concept domain targeted interventions.

Methodological Issues That Make a Difference: Optimizing Evaluation of Self-Concept Enhancements

In our meta-analysis, we aimed to shed light upon the way in which self-concept intervention research is conducted. Indeed, it would have been of limited value to examine only the conceptual and practical features of the interventions, without taking into account the design of the programs in terms of randomization and control groups. Features of intervention research methodology have long been implicated in masking or over-emphasizing the effectiveness of interventions (Lipsey & Wilson, 2001; Strein, 1988).

Random assignment. Failure to randomly assign participants to treatment conditions has been suggested as a source of inaccurate research findings because outcomes may be biased due to potentially differential levels of self-concept in the groups at pre-test—a form of selection bias (Matt & Cook, 1994; Wilson & Lipsey, 2001). However, in a synthesis of 319 meta-analyses of psychological, behavioral, and educational intervention research, Wilson and Lipsey (2001) found that randomized designs did not differ significantly from nonrandomized designs, although the variation between meta-analyses was quite large. In our study we found that random assignment to groups was a highly significant predictor in the fixed-effect model but that there was so much heterogeneity in the results from different studies that the effect was not statistically significant in the more conservative random effects model.

Whereas "conventional wisdom" suggests that the effect sizes should be smaller (less positive) when true random assignment is used, we found that the effect sizes were larger when true random assignment was used. The apparent explanation is that the most effective interventions were based on students with a pre-existing problem who were disadvantaged in some way. In this situation a lack of true random assignment is likely to result in comparison "control" groups that are less disadvantaged than intervention groups, resulting in a negatively biased estimate of the intervention effects. The finding that this effect was not statistically significant for the random effects model suggests that this bias in the nonrandom assignment of students to comparison groups was not consistent over studies. Because of the critical implications of this potential bias, it is imperative for researchers to fully describe procedures used to assign participants to groups—particularly when there is not true random assignment at the individual level.

Placebo control groups. Another key finding emanating from the present investigation was that placebo groups had an (albeit nonsignificant) tendency toward lower effect sizes than waitlist controls. These results suggest that so-called placebo controls might have a "real" effect on self-concept beyond that which might be expected from true "no treatment" control groups. Thus, for example, Marsh et al. (1986a, 1986b) argued that comparisons with placebo controls in self-concept research should be made cautiously because conditions that are like the intervention, but are unrelated to the intended effects of the intervention, are unlikely to exist or may not be feasible to construct. They recommended that an alternative control for placebo-like effects was a construct validity approach that included different components of self-concept that varied systematically in their relevance to the intervention. Support for the construct validity of the intervention requires that the intervention effects are substantially larger for more relevant (target) domains than less relevant (nontarget) domains—a pattern of results found in our meta-analysis. Marsh et al. also suggested that maintenance of intervention effects over post-test at follow-up testing also runs counter to the construct validity of counter-interpretations of results in relation to placebo effects. Although this is clearly beyond the scope of the present investigation, we suggest that studies using placebo controls should also include no-treatment control groups. In this way they could critically evaluate differences in the design of conditions in the intervention and placebo control groups and test the construct validity of interpretations in relation to multiple dimensions of self-concept and follow-up testing.

Intervention Strategies That Make a Difference: Optimizing Self-Concept Enhancements

Use of praise and feedback. The most common intervention was for practice or training in certain tasks ($n = 82$), which produced a mean effect size slightly lower than the overall mean ($d = .42$). However, we found that interventions emphasizing praise and/or feedback yielded the highest mean effect size ($d = 1.13$). This was also one of only three categories that were left in the random effects multivariate regression after backward elimination and was a strong predictor of positive effect size. Furthermore, this finding has been supported by traditional literature reviews (e.g., Craven et al., 2003). Given the low cost of interventions based on the appropriate use of praise and feedback, and the relative ease with which such interventions can be introduced, this finding has important practical implications for intervention delivery.

To elaborate on this finding, we conducted supplementary analyses in which praise/feedback interventions were divided into specific subtypes. Specifically, we found that interventions administering noncontingent praise were substantially less effective than those utilizing attributional feedback, goal feedback, or contingent praise. This result is consistent with findings such as those by Dweck (see Kamins & Dweck, 1999; Mueller & Dweck, 1998), who also notes that not all forms of praise and feedback foster a positive self-worth, typically citing person or trait-related praise (noncontingent praise) as being a cause of helplessness, with the potential to undermine self-worth. This has been discussed in the context of a comparison with ability or effort contingent praise (Kamins & Dweck, 1999; Mueller & Dweck, 1998) and has been explained from a motivational perspective (e.g., Dweck & Legget, 1988).

It is also relevant to note that the second highest effect size was for interventions that used praise or feedback in conjunction with some other feature, such as skills training. Importantly, these results support the synergy of self-concept interventions with skills acquisition. About half of the skill/feedback manipulations were explicitly described as contingent upon performance on a particular task, so it is not surprising that effects of praise/feedback with or without a specific skills training component were similarly effective. In conjunction with the suggestions of reciprocal effects theorists (see Marsh & Craven, 2005), it could be argued that schools need to simultaneously focus upon enhancing skills and self-concept, using praise and feedback strategies to address the latter (e.g., Craven et al., 2003; Craven, Marsh, & Debus, 1991). In contrast, skills training interventions without praise/feedback were less effective. These results have important implications for practice, in that interventions aiming to enhance self-concept are best served by using praise or feedback strategies because skill training alone is less effective.

Focus of intervention on self-concept. In our meta-analysis, we sought to explicate effective strategies in the design and implementation of interventions in the hope of identifying optimal self-concept enhancement strategies. We found that studies that directly target self-concept yielded higher mean effect-sizes than those targeting other constructs ($d = .67$ for direct interventions, $d = .48$ for indirect, and $d = .27$ for incidental interventions), which is consistent with Haney and Durlak's (1998) meta-analytic findings. However, in contrast to the Haney and Durlak (1998) meta-analysis, these differences were not found to be statistically significant because the groups in the present meta-analysis were found to be heterogeneous. The difference between Haney and Durlak's finding and the present result is possibly due to our emphasis on a random effects model, as the effects of this variable were statistically significant for our fixed effects model, which was similar to that used by Haney and Durlak. Further, we categorized the studies differently to the Haney and Durlak meta-analysis, adopting Hattie's (1992) use of four categories. As in the present meta-analysis, Hattie found no significant difference between direct, indirect, incidental, and combination interventions. These findings emphasize the need to address self-concept development rather than simply building one's skills. Thus, given that higher effect sizes were found for direct interventions in both meta-analyses, and in conjunction with the findings from the "match" moderator discussed above, it is reasonable to suggest that "optimal" interventions should ideally be focusing directly on enhancing self-concept in specific areas relevant to the goals of the intervention.

However, although direct interventions are clearly preferred if self-concept enhancement is desired, we offer two qualifications. First, rarely do administrators seek to enhance self-concept only. Enhancing a range of other outcomes (such as achievement) might also be desired. Consistent with the reciprocal effects model (Marsh & Craven, 2006), it is suggested that a combination of direct and indirect methods would be preferred if the goal of an intervention is to improve both self-concept and related performance outcomes (Marsh, Byrne, & Yeung, 1999). Thus, for example, if the goals of an intervention are to simultaneously enhance both math self-concept *and* math achievement, then it makes sense to target both self-concept and achievement. Second, the long-term impact of interventions is likely to be stronger if both performance and self-concept are increased. It is not clear whether enhancing self-concept alone can be sustained over time in the absence of actual improvement in ability. As such, we suggest that interventions using a combination of direct and indirect methods can be beneficial in both enhancing non–self-concept outcomes and for the maintenance of intervention success after post-test.

In summary, whereas the optimal way to improve self-concept over the short-term is to focus interventions directly on self-concept enhancement, interventions that combine direct self-concept enhancement in concert with performance enhancement, coupled with appropriate feedback and praise, are likely to be advantageous when the goals of the intervention are to improve both self-concept and performance (Marsh & Craven, 2006). An important area of future research is to combine research showing the reciprocal effects of self-concept and performance (Marsh & Craven, 2006) with research into the most effective self-concept enhancement interventions (Craven et al., 2003) to better integrate the important implications from both areas of research.

Nature of participants. We found that interventions targeting initially disadvantaged participants (i.e., those diagnosed with pre-existing problems such as low self-esteem, behavioral problems, learning disabilities, etc.) were more effective than preventive interventions, which is consistent with the findings of previous self-concept meta-analyses (Haney & Durlak, 1998; Hattie, 1992). This finding demonstrates that the potential to increase self-concept is much larger for groups who are "at risk". In contrast, groups without pre-existing problems are already functioning at reasonably high levels of self-concept and thus do not benefit nearly as much from self-concept enhancement interventions. This is a particularly important finding for children and adolescents in school settings where the gap between advantaged and disadvantaged students grows continually larger with age and with time spent in formal schooling (see Marsh & Kleitman, 2002 for further discussion of the social inequality gap reduction model and the widely supported Mathews effect in which smaller differences in young children grow larger as they grow older). By implication, this increasing gap between advantaged and disadvantaged students suggests that our schools are failing their most needy students. Hence, it is gratifying to find that self-concept interventions are particularly beneficial for disadvantaged students. Given the reciprocal relation that self-concept shares with academic achievement and its impact on a plethora of desirable outcome measures, it would seem useful to ensure that "at risk" groups experience simultaneously self-concept enhancement and skills interventions to assist in closing the inequality gap. From a practical perspective this might be one of the most exciting and important implications of our study.

Enhancement administrator. There has been much controversy in educational psychology since Hattie (1992) found in a meta-analysis of self-concept interventions that teachers were the least effective treatment administrators. The present findings, however, do not support those results. The random effects between group homogeneity was nonsignificant, which indicates there was no significant difference between the different categories. A number of the categories (teachers included) had mean effect sizes slightly higher than the overall mean but were indistinguishable from each other. Thus, intervention administrator had no statistically significant impact on intervention success. The difference between the current meta-analysis and Hattie's (1992) results may stem from the different samples considered. Hattie also included adults in the sample, which means that a broader categorization of teacher was used than was applied here. The lack of discrepancy between groups in the current investigation can be seen as a positive outcome, as it suggests that successful interventions are not limited to those conducted by researchers, and thus self-concept enhancement interventions can be successfully implemented by a diversity of administrators in a range of contexts. These results demonstrate the potent practicality of enhancing self-concept in diverse settings.

It is interesting to note that the fixed effects homogeneity model indicated that there was a statistically significant difference between the treatment administrator groups (Table 4). Further, the corresponding multivariate regression analysis (Table 6) disturbingly showed that the effects of mental health professionals and mixed administrators were actually negative. This could suggest a negative impact of these administrator types on intervention effectiveness. However, it is possible that there is some confounding of administrator effectiveness with the type of intervention delivered. For instance, mental health professionals were typically associated with treatments requiring more technical expertise, such as those involving electromyography relaxation training, than praise-based interventions. Also, the level of familiarity with the participants could confound the treatment administrator type, as some administrators (e.g., teachers) might have more contact with the participants. Thus, future research might look at the interaction effects of such possible moderators. Importantly, because these results were not significant in the random effects model, their generalizability should be viewed cautiously.

Stability of effects over time. The present meta-analysis found that benefits attained at post-test were maintained at follow-up. This finding is consistent with Haney and Durlak's (1998) meta-analysis, although it contradicts Hattie's (1992) findings. Both of the earlier meta-analyses had very small samples of studies reporting follow-up effects, so the present finding is likely to be more indicative of actual effects. This is promising as it suggests that intervention effects are maintained in the moderate long-term. Nevertheless, it is recommended that future researchers conduct follow-up testing to gain a better understanding of the long-term benefits of self-concept interventions. Particularly for studies that simultaneously target self-concept and performance, it is important to establish whether growth in performance supports the long-term growth of self-concept as posited in the reciprocal effects model (Marsh & Craven, 2006). Although clearly relevant for self-concept research, related issues are important for educational research more generally in that the effects of most interventions are found to diminish over time after the intervention has concluded.

Strengths, Limitations, and Directions for Future Research

Detail of included studies. Meta-analysis is limited by the information provided by the studies synthesized. In this meta-analysis, a large number of studies with nonsignificant results did not report enough information to calculate an accurate estimate of effect. We used the procedure of setting nonsignificant effects at zero, which is typical in research syntheses but which may lead to a downward bias in effect sizes (Lipsey & Wilson, 2001). Hence, for example, when the average effect size is positive overall, it is more likely that nonsignificant effects are positive than negative. Unfortunately, there is no way to deal with missing effect sizes that is universally agreed upon (Pigott, 1994), although it is quite clear that not including such studies is a poor compromise.

Also problematic for the present meta-analysis was the lack of detail regarding the age of the participants reported in the studies. With the majority of studies not reporting a directly comparable age indicator (such as mean age), we were forced to infer the mean age by determining the midpoint of the range of ages reported. This could seriously undermine the accuracy of the analyses regarding the predictive value of developmental trends in self-concept intervention success. It seems likely that developmental trends would be a significant predictor of effect size and yet were not found to be in the present analyses. Whereas using a single value of age to characterize a study might be reasonable when participants are homogeneous in terms of age, much potentially valuable developmental information is lost when age within a particular study is heterogeneous. Particularly when the effects of the intervention and other critical variables interact with age, it is important that primary researchers provide sufficient information so that these within-study age effects can be incorporated into meta-analyses as well as the between-study age differences emphasized in the present investigation. Ideally, this would involve reporting separate means (and standard deviations) for each age group, but this is rarely done. Researchers are strongly encouraged to report detailed demographic information such as age, sex, and SES—and how these effects interact with the intervention—to more accurately describe the results and to facilitate comparisons with other research.

Another key recommendation for future self-concept researchers would be to systematically conduct follow-up testing. Follow-up studies showed no statistically significant increase nor decrease in effect sizes over the follow-up period. However, despite the small number of follow-up studies available, the effect size for follow-up studies was positively related to length of time between the end of the intervention and follow-up testing. This suggests that there may be a sleeper effect such that effects grow larger over time. Because of the small number of follow-up studies, it was not possible to explore moderators of this effect to determine what design characteristics maximized long-term growth. However, the possibility clearly warrants further consideration in this area of research, and we suggest that researchers should include long-term follow-ups in their intervention studies.

Fixed versus random effects. In our meta-analysis we presented both fixed and random effects analyses. Fixed effects analyses attempt to determine whether similar results would be produced to those found in the present set of studies, if these studies were conducted again in an identical manner but using different participants drawn from the same population. On the other hand, random effects analyses ask whether a hypothetical set of new studies from the same population as the current 145 studies would produce similar results, despite using different methods and different participants drawn from the same population.

The increased generalizability afforded by the random effects model has led numerous meta-analytic researchers to conclude that random effects models are more appropriate (e.g., Erez, Bloom, & Wells, 1996; Hunter & Schmidt, 2000). Given that the self-concept literature is so varied in terms of the programs implemented, meaning that each study is not simply an exact replication of those before it, we argue that the random effects model is more appropriate for the self-concept literature. This contention is further supported by the homogeneity analyses, which consistently revealed that the data was heterogeneous. Rejection of homogeneity suggested that generalizing the findings of the fixed effects model beyond the current 145 reports may not be warranted. Because the aim of this article was to elucidate features of interventions that are the most effective in producing self-concept enhancements, the ability to infer from this sample of studies to the greater population is essential, thus making the random effects model more suitable.

Meta-analysis: A multidimensional, construct validation perspective. Historically, a focus of meta-analysis has been on the synthesis of research in which results from a large number of studies are aggregated into a single average effect size. However, particularly in education, psychology, and the social sciences more generally, it is typically unrealistic to treat different studies included in a meta-analysis as simple replications of one study. Indeed, there is a bias against even publishing "simple" replication studies in these disciplines. Hence, typically there is considerable variation among the primary studies in the participants, outcome measures, design features, and so forth. One of the particularly valuable features of meta-analysis is the ability to evaluate how different features of the studies, participants, and design influence the effect sizes. This extremely valuable and heuristic aspect also opens up a Pandora's box of problems regarding how to interpret the results of these various features that influence the observed effect sizes. Even when all of the studies included in a meta-analysis are based on true experimental designs with random assignment (which is typically not the case), differences between the studies are clearly "correlational" in nature so that "causal" interpretations must be offered as tentative hypotheses whose validity must be tested in relation to competing interpretations. From the perspective of our construct validity approach, we argue that the meta-analysts have an obligation to systematically evaluate the construct validity of their interpretations.

Although there are numerous strategies meta-analysts can use to interrogate the construct validity of the interpretations, one of the most universally accepted approaches is to consider the construct validity of results in relation to a multidimensional profile of outcome measures. As applied in the present investigation, the multidimensional profile consisted of different domains of self-concept. In relation to traditional construct validity terminology, we sought support for the convergent and discriminant validity of the interpretations. Support for convergent validity requires that the self-concept domains most logically related to the intended aims of an intervention have statistically significant and substantial effect sizes. Support for discriminant validity requires that self-concept domains least logically related to the aims of an intervention have substantially smaller effect sizes than those most logically related. One of the most striking features of the present investigation is the strength of support for both convergent and discriminant validity of the interpretations of the results. Whereas the success of our multidimensional construct validity approach provided particularly strong support for the multidimensional perspective, we contend that this approach—or variations of it—provides an important new set of strategies for meta-analysts to use more generally.

CONCLUSION

The use of sophisticated meta-analytic techniques to elucidate the results has proven to be an invaluable tool for synthesizing and critically analyzing the available body of self-concept intervention studies. The random effects mean effect size of .51 is promising: The positive direction of the effect size means that children and adolescents are benefiting from self-concept enhancement interventions. Further, the findings point to new ways in which we can improve self-concept interventions. For instance, such interventions may demonstrate improved efficacy if they incorporate appropriate praise and/or feedback strategies into the program, especially if these strategies are contingent upon performance that is attributional in nature and goal-relevant. The current study also indicates that targeting children and adolescents with diagnosed problems may be more valuable than instigating preventive programs. However, the most distinct finding was in regard to the multidimensionality of self-concept. The results suggest that interventions need to focus on specific dimensions of self-concept and then assess the effects of the intervention in relation to that particular self-concept domain instead of, or in addition to, other specific and global components of self-concept. Rather than using blanket interventions that try to improve all aspects of the children and adolescents' self-evaluations at once, researchers should focus on domain-specific programs to ensure the goals of the programs are truly met. In terms of program evaluation, using multidimensional measures is important because the current results indicate that using global scales of self-concept may underestimate intervention benefits. As a result, it is possible that successful intervention features have not been detected because inadequate scales cannot detect significant gains resulting from the treatments. Given the importance of having a high self-concept and the money and time invested into implementing enhancement programs, it seems timely for researchers and practitioners to capitalize on developments in multidimensional self-concept theory and instrumentation to create an array of interventions that target domain specific facets of self-concept.

REFERENCES

*Akande, A., Ntshanga, P., & Nyanhette, P. (1995). Communication: Its effects on the self-concept of children: A South African perspective. *Early Child Development and Care, 105*, 69–76.

*Allen, G. J., Chinsky, J. M., Larcen, S. W., Lochman, J. E., & Selinger, H. V. (1976). *Community psychology and the schools: A behaviorally oriented multilevel preventive approach.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

*Altmann, H. A., & Firnesz, K. M. (1973). A role-playing approach to influencing behavioral change and self-esteem. *Elementary School Guidance and Counseling, 7*(4), 276–281.

*Amerikaner, M., & Summerlin, M. L. (1982). Group counseling with learning disabled children: Effects of social skills and relaxation training on self-concept and classroom behavior. *Journal of Learning Disabilities, 15*(6), 340–343.

*Anderson, E. C. (1967). Counseling and consultation versus teacher-consultation in the elementary school. *Elementary School Guidance and Counseling, 2*, 276–285.

*Arborelius, E., & Bremberg, S. (1998). "It is your decision!"—Behavioural effects of a student-centered health education model at school for adolescents. *Journal of Adolescence, 11*(4), 287–297.

*Arnheim, D. D., & Sinclair, W. A. (1974). The effect of a motor development program on selected factors in motor ability, personality, self-awareness and vision. *American Corrective Therapy Journal, 6,* 167–171.

*Ascione, F. R., & Borg, W. R. (1983). A teacher-training program to enhance mainstreamed, handicapped pupils' self-concepts. *Journal of School Psychology, 21*(4), 297–309.

*Baker, S. B., Thomas, R. N., & Munson, W. W. (1983). Effects of cognitive restructuring and structured group discussion as primary prevention strategies. *The School Counselor, 31*, 26–33.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Barber, C. N., Hall, J., & Armistead, L. (2003). Parent-adolescent relationship and adolescent psychological functioning among African-American female adolescents: Self-esteem as a mediator. *Journal of Child & Family Studies, 12,* 361–374.

*Barrett, P. M. (1999). Adolescent self-esteem and cognitive skills training: A school-based intervention. *Journal of Child and Family Studies, 8*(2), 217–227.

*Bates, M. (April, 1968). A test of group counseling. *Personnel and Guidance Journal,* 749–753.

*Bhola, P. (1999). Brief supportive psychotherapy with emotionally disturbed adolescent girls. *NIMHANS Journal, 17*(2), 131–147.

*Biermann, K. L., & Furman, W. (1984). The effects of social skills training and peer involvement on the social adjustment of pre-adolescents. *Child Development, 55,* 151–162.

*Bijstra, J. O., & Jackson, S. (1998). Social skills training with early adolescents: Effects on social skills, well-being, self-esteem and coping. *European Journal of Psychology of Education, 13*(4), 569–583.

*Bleck, R. T., & Bleck, B. L. (1982). The disruptive child's playgroup. *Elementary School Guidance and Counseling, 17*, 137–141.

*Bockoven, J., & Morse, L. A. (1986). A comparative study of the efficacy of the DUSO and DUSO-R on children's social skills and self-esteem. *Elementary School Guidance and Counseling, 20*(4), 290–296.

Borenstein, M. T., & Rothstein, H. (1999). *Comprehensive meta-analysis: A computer program for research synthesis* (Version 1.0.23) [Computer software]. Englewood Cliffs, NJ: Biostat.

*Bornstein, M. T., Bornstein, P. H., & Walters, H. A. (1988). Children of divorce: Empirical evaluation of a group-treatment program. *Journal of Clinical Child Psychology, 3,* 248–254.

Bracken, B. A. (1996). Clinical applications of a context-dependent, multidimensional model of self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept* (pp. 463–503). New York: Wiley.

*Bruce, P. (1958). Relationship of self-acceptance to other variables with sixth grade children oriented in self-understanding. *The Journal of Educational Psychology, 49*(5), 229–238.

*Bruya, L. D. (1977). Effect of selected movement skills on positive self-concept. *Perceptual and Motor Skills, 45*, 252–254.

*Burke, D. M., & Van de Streek, L. (1989). Children of divorce: An application of Hammond's group counseling for children. *Elementary School Guidance and Counseling, 24,* 112–118.

*Burnett, P. C. (1983). A self-concept enhancement program for children in the regular classroom. *Elementary School Guidance and Counseling, 18*(2), 101–108.

Burnett, P. C., Pillay, H., & Dart, B. C. (2003). The influences of conceptions of learning and learner self-concept on high school students' approaches to learning. *School Psychology International, 24,* 54–66.

Byrne, B. M. (1984). The general/academic self-concept nomological network: A review of construct validation research. *Review of Educational Research, 54*, 427–456.

Byrne, B. M. (1996). *Measuring self-concept across the life span: Issues and instrumentation.* Washington, DC: American Psychological Association.

*Calsyn, R. J., Pennell, C., & Harter, M. (1984). Are affective education programs more effective with girls than with boys? *Elementary School Guidance and Counseling, 19,* 132–140.

*Calsyn, R. J., Quicke, J., & Harris, S. (1980). Do improved communication skills lead to increased self-esteem? *Elementary School Guidance and Counseling*, 15, 48–55.

*Cangelosi, A., Gressard, C. F., & Mines, R. A. (1980). The effects of a rational thinking group on self-concepts in adolescents. *The School Counselor, 27,* 357–361.

*Cherry, V. R., & Belgrave, F. Z. (1998). NTU: An Africentric approach to substance abuse prevention among African American youth. *Journal of Primary Prevention, 18*(3), 319–339.

*Christopher, F. S., & Roosa, M. W. (1990). An evaluation of an adolescent pregnancy prevention program: Is "just say no" enough? *Family Relations, 39,* 68–72.

*Clements, B. E. (1966). Transitional adolescents, anxiety, and group counseling. *Personnel and Guidance Journal, 45,* 67–71.

Cohn, L. D., & Becker, B. J. (2003). How meta-analysis increases statistical power. *Psychological Methods, 8*, 243–253.

*Comer, J. P. (1985). The Yale-New Haven primary prevention project: A follow-up study. *Journal of the American Academy of Child Psychiatry, 24,* 154–160.

Cooper, H. M. (1998). *Integrating research: A guide for literature reviews* (3rd ed.). Newbury Park, CA: Sage.

Coopersmith, S. A. (1967). *The antecedents of self-esteem*. San Francisco: Freeman.

*Costas, M., & Landreth, G. (1999). Filial therapy with nonoffending parents of children who have been sexually abused. *International Journal of Play Therapy, 8*(1), 43–66.

Craven, R. G., Marsh, H. W., & 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, & D. McInernery (Eds.), *International advances in self research* (Vol. 1, pp. 91–126). Greenwich, CT: Information Age.

*Craven, R. G., Marsh, H. W., & Debus, R. L. (1991). Effects of internally focused feedback and attributional feedback on enhancement of academic self-concept. *Journal of Educational Psychology, 83*(1), 17–27.

*Crow, M. L. (1969). A comparison of three group counseling techniques with sixth graders. *Elementary School Guidance and Counseling, 6*, 37–41.

Delugach, R. R., Bracken, B. A., Bracken, M. J., & Schicke, M. C. (1992). Self-concept: Multidimensional construct exploration. *Psychology in the Schools, 29*, 213–223.

*Denkowski, K. M., Denkowski, G. C., & Omizo, M. M. (1983). The effects of EMG-assisted relaxation training on the academic performance, locus of control, and self–esteem of hyperactive boys. *Biofeedback and Self-Regulation, 8*(3), 363–375.

*Dicken, C., Bryson, R., & Kass, N. (1977). Companionship therapy: A replication in experimental community psychology. *Journal of Consulting and Clinical Psychology, 45*(4), 637–646.

*Dillard, J. M., Warrior-Benjamin, J., & Perrin, D. W. (1977). Efficacy of test-wiseness on test anxiety and reading achievement among black youth. *Psychological Reports, 41*, 1135–1140.

Donahue, E. M., Robins, R. W., Roberts, B. W., & John, O. P. (1993). The divided self: Concurrent and longitudinal effects of psychological adjustment and social roles on self-concept differentiation. *Journal of Personality* and Social Psychology, 64, 834–846.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273. *Adolescence, 29*(116), 935–939.

*Eggert, L. L., Thompson, E. A. (1995). Reducing suicide potential among high-risk youth: Tests of a school-based prevention program. *Suicide and Life-Threatening Behavior, 25*(2), 276–296.

*Eitan, T., Amir, Y., & Rich, Y. (1992). Social and academic treatments in mixed-ethnic classes and change in student self-concept. *British Journal of Educational Psychology, 62*(3), 364–374.

*Eldridge, M. S., Barcikowskik, R. S., & Witmer, J. M. (1972). Effects of DUSO on the self-concepts of second grade students. *Elementary School Guidance and Counseling, 7*(4), 256–260.

*Eldridge, M. S., Witmer, J. M., Barcikowski, R., & Bauer, L. (1977). The effects of a group guidance program on the self-concepts of EMR children. *Measurement and Evaluation in Guidance, 9,* 184–191.

*Elias, M. J. (1983). Improving coping skills of emotionally disturbed boys through television-based social problem solving. *American Journal of Orthopsychiatry, 53*(1), 61–72.

*Emmons, R. D., & Nystul, M. S. (1994). The effects of a prenatal course including prep for effective family living of self-esteem and parenting attitudes of adolescents: A brief report. *Adolescence, 29,* 935–938.

Erez, A., Bloom, M. C., & Wells, M. T. (1996). Using random rather than fixed effects models in meta-analysis: Implications for situational specificity and validity generalization. *Personnel Psychology, 49*, 275–306.

*Esters, P., & Levant, R. F. (1983). The effects of two parent counseling programs on rural low-achieving children. *The School Counselor, 31*, 159–166.

*Felker, D. W., Stanwyck, D. J., & Kay, R. S. (1973). The effects of a teacher program in self-concept enhancement on pupils' self-concept, anxiety, and intellectual achievement responsibility. *Journal of Educational Research, 66*(10), 443–447.

*Felner, R. D., Ginter, M., & Primavera, J. (1982). Primary prevention during school transitions: Social support and environmental structure. *American Journal of Community Psychology, 10*, 277–290.

*Fertman, C. I., & Chubb, N. H. (1992). The effects of a psychoeducational program on adolescents' activity involvement, self-esteem, and locus of control. *Adolescence, 27*(107), 517–526.

*Flaherty, M. A. (1992). The effects of a holistic creativity program on the self-concept and creativity of third graders. *Journal of Creative Behavior, 26*(3), 165–171.

*Flowers, J., & Martson, A. (1972). Modification of low self-confidence in elementary school children. *Journal of Educational Research, 66*, 30–35.

*Foster, G. D., Wadden, T. A., & Brownell, K. D. (1985). Peer-led program for the treatment and prevention of obesity in the schools. *Journal of Consulting and Clinical Psychology, 53*(4), 538–540.

*Franklin, D. (1986). A comparison of the effectiveness of values clarification presented as a personal computer program versus a traditional therapy group: A pilot study. *Occupational Therapy in Mental Health, 6*(3), 39–52.

*Fryrear, J. L., Nuell, L. R., & Ridley, S. D. (1974). Photographic self-concept enhancement of male juvenile delinquents. *Journal of Consulting and Clinical Psychology, 42*(6), 915.

*Fullerton, S. (1973). Self-concept changes of junior high students. *Journal of Counseling Psychology, 20,* 493–494.

*Galvin, M. (1983). Making systematic problem solving work with children. *The School Counselor, 31*(2), 130–136.

*Gesten, E. L., Rains, M. H., Rapkin, B. D., Weissberg, R. P., de Apocada, R. F., Cowen, E. L. et al. (1982). Training children in social problem-solving competencies: A first and second look. *American Journal of Community Psychology, 10*, 95–115.

*Giannotti, T. J., & Doyle, R. E. (1982). The effectiveness of parental training on learning disabled children and their parents. *Elementary School Guidance and Counseling, 17,* 131–135.

*Goni, A., & Zulaika, L. (2000). Relationships between physical education classes and the enhancement of fifth grade pupils' self-concept. *Perceptual and Motor Skills, 91*(1), 246–250.

*Graybill, D. (1986). A multiple-outcome evaluation of training parents in active listening. *Psychological Reports, 59,* 1171–1185.

*Graybill, D., & Gabel, H. (1982). A school-based counseling program for parents of learning disabled children using paraprofessional co-counselors. *Psychological Reports, 50,* 729–730

*Graybill, D., & Gabel, H. (1981). Evaluation and comparison of reflective and behavioral groups counseling for parents. *Journal of Counseling and Psychotherapy, 4*(1), 19–37.

Gurney, P. (1986). Self-esteem in the classroom: Theoretical perspectives and assessment issues. *School Psychology International, 7*, 199–209.

*Hains, A. A. (1994). The effectiveness of a school-based, cognitive-behavioral stress management program with adolescents reporting high and low levels of emotional arousal. *The School Counselor, 42*(2), 114–125.

*Hains, A. A., & Szyjakowski, M. (1990). A cognitive stress reduction intervention program for adolescents. *Journal of Counseling Psychology, 37*(1), 79–84.

Haney, P., & Durlak, J. A. (1998). Changing self-esteem in children and adolescents: A meta-analytic review. *Journal of Clinical Child Psychology*, *27*(4), 423–433.

*Harris, S. R. (1976). Rational-emotive education and the human development program: A guidance study. *Elementary School Guidance and Counseling*, *11*, 113–122.

Harter, S. (1990). Processes underlying adolescent self-concept formation. In R. Montemayor, G. Adams, & T. Gullotta (Eds.), *From childhood to adolescence: A transitional period?* (pp. 205–239). Thousand Oaks, CA: Sage.

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

*Hauserman, N., Miller, J. S., & Bond, F. T. (1976). Influence of decision-making on attitudes, creativity, motor skills and self-concept in elementary children. *Psychological Record, 26*(1), 111–116.

*Hay, I., Byrne, M., & Butler, C. (2000). Evaluation of conflict-resolution and problem-solving programme to enhance adolescents' self-concept. *British Journal of Guidance and Counseling, 28*(1), 101–113.

*Hayes, E. J., Cunningham, G. K., & Robinson, J. B. (1977). Counseling focus: Are parents necessary? *Elementary School Guidance and Counseling*, *12*, 8–14.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for the meta-analysis.* New York: Academic.

Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods, 6*, 203–217.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*, 486–504.

*Hudgins, E. W. (1979). Examining the effectiveness of affective education. *Psychology in the Schools*, *16*(4), 581–585.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, *8*, 275–292.

*Jackson, M. D., & Brown, D. (1986). Use of systematic training for effective parenting (STEP) with elementary school parents. *The School Counselor*, *34*, 100–105.

*Jackson, R. M., Cleveland, J. C., & Merenda, P. F. (1968–69). The effects of early identification and counseling of underachievers. *Journal of School Psychology*, *7*, 42–49.

James, W. (1896). *Talks to teachers.* New York: Norton.

*Jason, L., & Burrows, B. (1983). Transition training for high school seniors. *Cognitive Therapy and Research, 7*(1), 79–92.

*Jason, L. A., Betts, D., Johnson, J., Smith, S., Krueckeberg, S., & Cradock, M. (1989). An evaluation of an orientation plus tutoring school-based prevention program. *Professional School Psychology*, *4*, 273–284.

*Jones, M. B., & Offord, D. R. (1989). Reduction of antisocial behavior in poor children by nonschool skill-development. *Journal of Child Psychiatry*, *30*(5), 737–750.

*Jorgensen, S. R. (1991). Project Taking Charge: An evaluation of an adolescent pregnancy prevention program. *Family Relations: Interdisciplinary Journal of Applied Family Studies, 40*(4), 373–380.

*Kahn, J. S., Kehle, T. J., Jenson, W. R., Clark, E. (1990). Comparison of cognitive-behavioral, relaxation, and self-modeling interventions for depression among middle-school students. *School Psychology Review*, *19*, 196–211.

Kamins, M., & Dweck, C. S. (1999). Person vs. process praise and criticism: Implications for contingent self-worth and coping. *Developmental Psychology, 35*, 835–847.

*Karnes, M. B. (1983). Short-term effects of early programming for the young gifted handicapped child. *Exceptional Children, 50*(2), 103–109.

*Kendall, P. C., & Braswell, L. (1982). Cognitive-behavioral self-control therapy for children: Components analysis. *Journal of Consulting and Clinical Psychology*, *50*, 672–689.

*Knaus, W., & Bokor, S. (1975). The effect of rational-emotive education lessons on anxiety and self-concept in sixth grade students. *Rational Living, 10*(2), 7–10.

*Krop, H., Calhoon, B., & Verrier, R. (1971). Modification of the "self-concept" of emotionally disturbed children by covert reinforcement. *Behavior Therapy*, *2*, 201–204.

Kumkale, G. T., & Albarracin, D. (2004). The sleeper effect in persuasion: A meta-analytic review. *Psychological Bulletin, 130*, 143–172.

*Landry, R. G., Schilson, E., & Pardew, E. M., (1974). Self concept enhancement in a preschool program. *The Journal of Experimental Education, 42*(4), 39–43.

*Levant, R. F., & Slattery, S. C. (1982). Systematic skills training for foster parents. *Journal of Clinical Child Psychology*, *11*, 138–143.

*Lieberman, L. D., Gray, H., Wier, M., Fiorentino, R., & Maloney, P. (2000). Long-term outcomes of an abstinence-based small-group pregnancy prevention program in New York City schools. *Family Planning Perspectives, 32*(5), 237–245.

Lipsey, M. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals of the American Academy, 587*, 69–81.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

*Lochman, J. E. (1992). Cognitive-behavioral intervention with aggressive boys: Three-year follow-up and preventive effects. *Journal of Consulting and Clinical Psychology, 60*(3), 426–432.

*Lochman, J. E., Lampron, L. B., Gemmer, T. C., Harris, S. R., & Wyckoff, G. M. (1989). Teacher consultations and cognitive-behavioral interventions with aggressive boys. *Psychology in the Schools*, *26*, 179–188.

*Loffredo, D. A., Omizo, M. M., & Hammett, V. L. (1984). Group relaxation training and parental involvement with hyperactive boys. *Journal of Learning Disabilities*, *17*(4), 210–213.

*Ludwig, D. J., & Maehr, M. L. (1967). Changes in self-concept and stated behavioural preferences. *Child Development, 38*, 453–467.

*Malgady, R. G., & Rogler, L. H. (1990). Hero/heroine modeling for Puerto Rican adolescents: A preventive mental health intervention. *Journal of Consulting and Clinical Psychology, 58*(4), 469–474.

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,* 1224–1236.

Marsh, H. W. (1990). A multidimensional, hierarchical self-concept: Theoretical and empirical justification. *Educational Psychology Review, 2,* 77–172.

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., & Byrne, B. M. (1993). Do we see ourselves as others infer: A comparison of self-other agreement on multiple dimensions of self-concept from two continents. *Australian Journal of Psychology, 45*, 49–58.

Marsh, H. W., Byrne, B. M., & Shavelson, R. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology, 80,* 366–380.

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and revised recommendations. *Educational Psychologist, 34,* 155–167.

Marsh, H. W., & Craven, R. (1991). Self-other agreement on multiple dimensions of preadolescent self-concept: Inferences by teachers, mothers, and fathers. *Journal of Educational Psychology, 83,* 393–404.

Marsh, H. W., & Craven, R. (1997). Academic self-concept: Beyond the dustbowl. In G. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement and adjustment* (pp. 131–198). Orlando, FL: Academic Press.

Marsh, H. W., & Craven, R. G. (2005). A reciprocal effects model of the causal ordering of self-concept and achievement: New support for the benefits of enhancing self-concept. In H. W. Marsh, R. G. Craven, & D. McInerney (Eds.), *International advances in self research, vol. 2*. (pp. 15–52). Greenwich, CT: Information Age.

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,* 95–180.

Marsh, H. W., Craven, R. G., & 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., Debus, R., & Bornholt, L. (2005). Validating young children's self-concept responses: Methodological ways and means to understand their responses. In D. M. Teti (Ed.), *Handbook of research methods in developmental psychology* (pp. 138–160). New York: Blackwell.

Marsh, H. W., & Hattie, J. (1996). Theoretical perspectives on the structure of self-concept. In B. A. Bracken (Ed.), *The handbook of self-concept* (pp. 38–90). New York: Wiley.

Marsh, H. W., & Hau, K. T. (2003). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal-external frame of reference predictions across 26 countries. *Journal of Educational Psychology, 96*, 56–67.

Marsh, H. W., & Kleitman, S. (2002). Extracurricular school activities: The good, the bad, and the nonlinear. *Harvard Educational Review*, *72,* 464–502.

Marsh, H. W., Papaioannou, A., & Theodorakis, Y. (2006). Causal ordering of physical self-concept and exercise behavior: Reciprocal effects model and the influence of physical education teachers. *Health Psychology, 25,* 316–328.

Marsh, H. W., & Peart, N. (1988). Competitive and cooperative physical fitness training programs for girls: Effects on physical fitness and on multidimensional self-concepts. *Journal of Sport and Exercise Psychology, 10,* 390–407.

Marsh, H. W., & Perry, C. (2005). Does a positive self-concept contribute to winning gold medals in elite swimming? The causal ordering of elite athlete self-concept and championship performances. *Journal of Sport and Exercise Psychology, 27*, 71–91.

Marsh, H. W., & Richards, G. (1988) The Outward Bound Bridging Course for low achieving high-school males: Effect on academic achievement and multidimensional self-concepts. *Australian Journal of Psychology, 40,* 281–298.

Marsh, H. W., Richards, G., & Barnes, J. (1986a). Multidimensional self-concepts: The effect of participation in an Outward Bound program. *Journal of Personality and Social Psychology, 45,* 173–187.

Marsh, H. W., Richards, G., & Barnes, J. (1986b). Multidimensional self-concepts: A long-term follow-up of the effect of participation in an Outward Bound program. *Personality and Social Psychology Bulletin, 12,* 475–492.

Marsh, H. W., Roche, L. A., Pajares, F., & Miller, D. (1997). Item-specific efficacy judgments in mathematical problem solving: The downside of standing too close to trees in a forest. *Contemporary Educational Psychology, 22,* 363–377.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of Personality, 74*, 403–456.

Marsh, H. W., Walker, R., & Debus, R. (1991). Subject-specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology, 16*, 331–345.

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

*Martinek, T. J., Cheffers, J. T. F., & Zaichkowsky, L. D. (1978). Physical activity, motor development and self-concept: Race and age differences. *Perceptual and Motor Skills, 46*, 147–154.

Marx, R., & Winne, P. H. (1978). Construct interpretations of three self-concept inventories*. American Educational Research Journal, 15,* 99–108.

Matt, G. E., & Cook, T. D. (1994). Threats to the validity of research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 503–520). New York: Russell Sage Foundation.

*Meredith, R., & Benninga, J. S. (1979). Counseling with parents to benefit children. *Elementary School Guidance and Counseling*, *14*, 36–42.

*Mezzano, J. (1968). Group counseling with low-motivated male high school students—comparative effects of two uses of counselor time. *The Journal of Educational Research*, *61*, 222–224.

*Michel, D. E., & Farrell, D. M. (1973). Music and self-esteem: Disadvantaged problem boys in an all-black elementary school. *Journal of Research in Music Education, 21*(1), 80–84.

*Michelson, L., Mannarino, A. P., Marchione, K. E., Stern, M., Figueroa, J., & Beck, S. (1983). A comparative outcome study of behavioral social-skills training, interpersonal problem-solving and non-directive control treatments with child psychiatric outpatients. *Behavior Research and Therapy*, *21*(5), 545–556.

*Miller, R. L., Brickman, P., & Nolan, D. (1975). Attribution versus persuasion as a means for modifying behavior. *Journal of Personality and Social Psychology, 31*(3), 430–441.

Mischel, W. (1997). On the future of personality measurement. *American Psychologist, 32*, 246–254.

*Morse, C. L., Bockoven, J., & Bettesworth, A. (1988). Effects of DUSO-2 and DUSO-2 revised on children's social skills and self-esteem. (Follow-up to Bockoven & Morse, 1986). *Elementary School Guidance and Counseling*, *22*, 199–205.

*Morse, C. L., Bockoven, J., & Harman, M. A. (1987). DUSO-R and ACCEPTS: The differential effects of two social skills curricula on children's social skills and self-esteem. *Child Study Journal*, *17*(4), 287–289.

Mueller, C. M., & Dweck, C. S. (1998). Intelligence praise can undermine motivation and performance*. Journal of Personality and Social Psychology, 75*, 33–52.

OECD (2003). *Student engagement at school: A sense of belonging and participation*. Paris: OECD.

*Nelson, G., & Carson, P. (1988). Evaluation of a social problem-solving skills program for third- and forth-grade students. *American Journal of Community Psychology*, *16*(1), 79–99.

*Oldfield, D. (1986). The effects of the relaxation response on self-concept and acting out behaviors. *Elementary School Guidance and Counseling*, *20*(4), 255–260.

*Omizo, M. M., & Cubberly, W. E. (1983). The effects of reality therapy classroom meetings on self-concept and locus of control among learning disabled children. *The Exceptional Child*, *30*(3), 201–209.

*Omizo, M. M., Cubberly, W. E., & Omizo, S. A. (1985). The effects of rational-emotive education groups on self-concept and locus of control among learning disabled children. *The Exceptional Child*, *32*(1), 13–19.

*Omizo, M. M., & Omizo, S. A. (1987a). Group counseling with children of divorce: New findings. *Elementary School Guidance & Counseling*, *22*, 46–52.

*Omizo, M. M., & Omizo, S. A. (1987b). The effects of group counseling on classroom behaviour and self-concept among elementary school learning disabled children. *The Exceptional Child*, *34*(1), 57–64.

*Omizo, M. M., Omizo, S. A., & D'Andrea, M. J. (1992). Promoting wellness among elementary school children. *Journal of Counseling and Development, 71*(2), 194–198.

Overton, R. C. (1998). A comparison of fixed effects and mixed (random effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods, 3*, 354–379.

*Page, R. C., & Chandler, J. (1994). Effects of group counseling on ninth-grade at-risk students. *Journal of Mental Health Counseling, 16*(3), 340–351.

Pajares, F. (1996, April). *Assessing self-efficacy beliefs and academic outcomes: The case for specificity and correspondence.* Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY.

*Parker, H. C. (1974). Contingency management and concomitant changes in elementary-school students' self-concepts. *Psychology in the Schools*, *11*, 70–79.

*Pedro-Carroll, J. L., & Cowen, E. L. (1985). The children of divorce intervention program: An investigation of the efficacy of a school-based prevention program. *Journal of Consulting and Clinical Psychology*, *53*, 603–611.

*Percy, L. E., Dziuban, C. D., & Martin, J. B. (1981). Analysis of effects of distance running on self-concepts of elementary students. *Perceptual and Motor Skills, 52*, 42.

*Perkins, J. A., & Wicas, E. A. (1971). Group counseling bright underachievers and their mothers. *Journal of Counseling Psychology*, *18*, 273–278.

*Phillips, R. H. (1984). Increasing positive self-referent statements to improve self-esteem in low-income elementary school children. *Journal of School Psychology*, 22, 155–163.

*Piercy, F., & Lee, R. (1976). Effects of a dual treatment approach on the rehabilitation of habitual juvenile delinquents. *Rehabilitation Counseling Bulletin, 19*(3), 482–491.

Pigott, T. D. (1994). Methods for handling missing data in research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 163–175). New York: Russell Sage Foundation.

*Pihl, R. O., Parkes, M., Drake, H., & Vrana, F. (1980). The intervention of a modulator with learning disabled children. *Journal of Clinical Psychology*, *36*, 972.

*Platzer, W. S. (1976). Effects of perceptual motor training on gross-motor skill and self-concept of young children. *American Journal of Occupational Therapy, 30*(7), 422–428.

*Prewitt-Diaz, J. O., & Golub, L. S. (1979). Reading and self-concept of Hispanic and non-Hispanic students. *Journal of Instructional Psychology, 7*(4), 127–136.

Raudenbush, S. W. (1994). Random effects model. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301–321). New York: Russell Sage Foundation.

*Rawson, H. E., & Cassady, J. C. (1995). Effects of therapeutic intervention on self-concepts of children with learning disabilities. *Child and Adolescent Social Work Journal, 12*(1), 19–31.

*Reinert, D. (1999). Group intervention for children of recovering alcoholic parents. *Alcoholism Treatment Quarterly, 17*(4), 15–27.

*Reynolds, W. M., & Coats, K. I. (1986). A comparison of cognitive-behavioral therapy and relaxation training for the treatment of depression in adolescents. *Journal of Consulting and Clinical Psychology*, *54*, 653–660.

*Rhodes, J. E., Haight, W. L., & Briggs, E. C. (1999). The influence of mentoring on the peer relationships of foster youth in relative and nonrelative care. *Journal of Research on Adolescence, 9*(2), 185–201.

*Rollin, S. A., Rubin, R., Hardy-Blake, B. (1994). Project K.I.C.K., a school-based drug education research project: Peers, parents and kids. *Journal of Alcohol and Drug Education, 39*(3), 75–86.

*Rollin, S. A., Rubin, R., & Marcil, R. (1995). Project KICK: A school-based drug education health promotion research project. *Counseling Psychology Quarterly, 8*(4), 345–359.

*Roosa, M. W., & Christopher, F. S. (1990). Evaluation of an abstinence-only adolescent pregnancy prevention program: A replication. *Family Relations*, *39*, 363–367.

*Roosa, M. W., Gensheimer, L. K., Short, J. L., Ayers, T. S., & Shell, R. (1989). A preventive intervention for children in alcoholic families: Results of a pilot study. *Family Relations*, *38*, 295–300.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Rosenberg, M. (1979). *Conceiving the self.* New York: Basic Books.

*Rotheram, M. J., Armstrong, M., & Booraem, C. (1982). Assertiveness training in fourth- and fifth-grade children. *American Journal of Community Psychology*, *10*, 567–582.

*Royse, D. (1998). Scouting and Girl Scout curriculum as interventions: Effects on adolescents' self-esteem. *Adolescence, 33*(129), 159–168.

*Schaps, E., Moskowitz, J. M., Condon, J. W., & Malvin, J. H. (1984). A process and outcome evaluation of an affective teacher training primary prevention program. *Journal of Alcohol and Drug Education, 29*(3), 35–64.

*Schempp, P. G., Cheffers, J. T., & Zaichkowsky, L. D. (1983). Influence of decision-making on attitudes, creativity, motor skills and self-concept in elementary children. *Research Quarterly for Exercise and Sport, 54*(2), 183–189.

*Schmieding, O. (1966). An investigation of efficacy of counseling and guidance procedures with failing junior high students. *The School Counselor*, *14*, 74–80.

*Schulman, J. L., Ford, R. C., & Busk, P. (1973). A classroom program to improve self-concept. *Psychology in the Schools*, *10*, 481–486.

*Schunk, D. H. (1983b). Reward contingencies and the development of children's skills and self-efficacy. *Journal of Educational Psychology*, *75*(4), 511–518.

*Schunk, D. H. (1983a). Ability versus effort attributional feedback: Differential effects on self-efficacy and achievement. *Journal of Educational Psychology*, *75*, 848–856.

*Scovern, A. W., Bukstel, L. H., Kilmann, P. R., Laval, R. A., Busemeyer, J., & Smith, V. (1980). Effects of parent counseling on the family system. *Journal of Counseling Psychology*, *27*, 268–275.

Shadish, W. R., Matt, G. E., Navarro, A. M., & Phillips, G. (2000). The effects of psychological therapies under clinically representative conditions: A meta-analysis. *Psychological Bulletin, 126*, 512–529.

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

*Shearn, D. F., & Randolph, D. L. (1978). Effects of reality therapy methods applied in the classroom. *Psychology in the Schools*, *15*, 79–83.

*Shechtman, Z., Gilat, I., Fos, L., & Flasher, A. (1996). Brief group therapy with low-achieving elementary school children. *Journal of Counseling Psychology, 43*(4), 376–382.

*Sheridan, J. T., Baker, S. B., & deLissovoy, V. (1984). Structured group counseling and explicit bibliotherapy as in-school strategies for preventing problems in youth of changing families. *The School Counselor*, *32*, 134–141.

Shirk, S. R. (1988). What is child psychotherapy? In S. R. Shirk (Ed.), *Cognitive development and child psychotherapy* (pp. 1–16). New York: Plenum Press.

*Simmons, C. H., & Parsons, R. J. (1983). Developing internality and perceived competence: The empowerment of adolescent girls. *Adolescence*, *18*(72), 917–922.

*Skitka, L. J., & Frazier, M. (1995). Ameliorating the effects of parental divorce: Do small group interventions work? *Journal of Divorce and Remarriage, 24*(3–4), 159–179.

*Smoll, F. L., Smith, R. E., Barnett, N. P., & Everett, J. J. (1993). Enhancement of children's self-esteem through social support training of youth sport coaches. *Journal of Applied Psychology, 78*(4), 602–610.

Snow, P. C., & Bruce, D. D. (2003). Cigarette smoking in teenage girls: Exploring the role of peer reputations, self-concept and coping. *Health Education Research, 18,* 439–452.

*Somerville, D. E., & Leach, D. J. (1988). Direct or indirect instruction? An evaluation of three types of intervention programme for assisting students with specific reading difficulties. *Educational Research, 30*(1), 46–53.

*Sorsdahl, S. N., & Sanche, R. P. (1985). The effects of classroom meetings on self-concept and behavior. *Elementary School Guidance and Counseling*, 20, 49–56.

*Springer, D. W. (2000). Effects of a solution-focused mutual aid group for Hispanic children of incarcerated parents. *Child & Adolescent Social Work Journal, 17*(6), 431–442.

*Stacey, S., & Rust, J. O. (1985). Evaluating the effectiveness of the DUSO-1 (revised) program. *Elementary School Guidance & Counseling*, *20*, 84–90.

*Stafford, W. B., & Hill, J. D. (1989). Planned program to foster positive self-concepts in kindergarten children. *Elementary School Guidance and Counseling*, *24*, 47–57.

*Stark, K. D., Reynolds, W. M., & Kaslow, N. J. (1987). A comparison of the relative efficacy of self-control therapy and a behavioral problem-solving therapy for depression in children. *Journal of Abnormal Child Psychology*, *15*(1), 91–113.

*Stevens, R., & Pihl, R. O. (1982). The remediation of the student at-risk for failure. *Journal of Clinical Psychology*, *38*(2), 298–301.

Stice, E., & Shaw, H. (2004). Eating disorder prevention programs: A meta-analytic review. *Psychological Bulletin, 130*, 206–227.

*Stilwell, W. E., Brown, P. W., & Barclay, J. R. (1973). Effects of three classroom management methods on classroom interaction of fifth graders. *Psychology in the Schools*, *10*, 365–372.

*Stolberg, A. L., & Garrison, K. M. (1985). Evaluating a primary prevention program for children of divorce. *American Journal of Community Psychology*, *13*(2), 111–123.

*Stoner, S., & Fiorillo, M. (1976). A program for self-concept improvement and weight reduction for overweight adolescent females. *Psychology, 13*, 30–35.

Strein, W. (1988). Classroom-based elementary school affective education programs: A critical review. *Psychology in the Schools, 25*, 288–296.

*Summerlin, M. L., Hammet, V. L., & Payne, M. L. (1983). The effect of magic circle participation on a child's self-concept. *The School Counselor*, *31*, 49–52.
*Thomson-Rountree, P., & Woodruff, A. E. (1982). An examination of project AWARE: The effects on children's attitudes towards themselves, others, and school. *Journal of School Psychology*, *20*(1), 20–30.
*Turner, S., & Scherman, A. (1996). Big brothers: Impact on little brothers' self-concepts and behaviors. *Adolescence, 31*(124), 875–883.
Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relations between self-beliefs and academic achievement: A systematic review. *Educational Psychologist, 39,* 111–133.
*Vander Kolk, C. J. (1973). Paraprofessionals as psychotherapeutic agents with moderately disturbed children. *Psychology in the Schools*, *10*, 238–242.
*Waksman, S. A. (1984a). A controlled evaluation of assertion training with adolescents. *Adolescence*, *19*, 277–282.
*Waksman, S. A. (1984b). Assertion training with adolescents. *Adolescence*, *19*, 123–130.
*Walters, S. T., & Martin, J. E. (2000). Does aerobic exercise really enhance self-esteem in children? A prospective evaluation in 3rd–5th graders. *Journal of Sport Behavior, 23*(1), 51–60.
*Wanat, P. E., (1983). Social skills: An awareness program with learning disabled adolescents. *Journal of Learning Disabilities*, *16*, 35.
*Warger, C. L., & Kleman, D. (1986). Developing positive self-concepts in institutionalized children with severe behavior disorders. *Child Welfare*, *65*(2), 165–176.
*Washington, K. R. (1976). The effects of systematic reinforcement and a self-awareness program on the self-concept of black preschool children. *Child Study Journal, 6*(4), 199–208.
*White, K., & Allen, R. (1971). Art counseling in an educational setting: Self-concept change among pre-adolescent boys. *Journal of School Psychology, 9*(2), 218–225.
Wilson, D. B., & Lipsey, M. W. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods, 6*, 413–429.
*Wilson, N. H., & Rotter, J. C. (1986). Anxiety management training and study skills counseling for students on self-esteem and test anxiety performance. *The School Counselor*, *34*, 18.
*Wirth, S. (1977). Effects of a multifaceted reading program on self-concept. *Elementary School Guidance and Counseling*, 12, 33–40.
Wolf, F. M. (1986). Meta-analysis: Quantitative methods for research synthesis. Beverly Hills, CA: Sage.
Wylie, R. C. (1989). *Measures of self-concept.* Lincoln: University of Nebraska Press.