



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Applied Developmental Psychology 25 (2004) 283–302

**Applied
Developmental
Psychology**

Promoting self-regulation through school-based martial arts training

Kimberley D. Lakes*, William T. Hoyt

Department of Counseling Psychology, University of Wisconsin-Madison, USA

Abstract

The impact of school-based Tae Kwon Do training on self-regulatory abilities was examined. A self-regulation framework including three domains (cognitive, affective, and physical) was presented. Children ($N = 207$) from kindergarten through Grade 5 were randomly assigned by homeroom class to either the intervention (martial arts) group or a comparison (traditional physical education) group. Outcomes were assessed using multidimensional, multimodal assessments. After a 3-month intervention, results indicated that the martial arts group demonstrated greater improvements than the comparison group in areas of cognitive self-regulation, affective self-regulation, prosocial behavior, classroom conduct, and performance on a mental math test. A significant Group \times Gender interaction was found for cognitive self-regulation and classroom conduct, with boys showing greater improvements than girls. Possible explanations of this interaction as well as implications for components of martial arts training for the development of self-regulation in school-age children are discussed.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Self-regulation; Martial arts; Self-discipline; Prosocial behavior; Classroom conduct; Gender differences; Character development; Attention

1. Introduction

In spite of the fact that there are numerous research-based programs for youth that are designed to reduce problem behaviors, there is a lack of research regarding the promotion of positive youth development (Larson, 2000). Baumeister (1997) describes one aspect of positive character development in his theory of self-regulation. According to this theory, effective self-regulation not only increases a person's capacity for success but also reduces self-destructive behavior. Baumeister defined self-

* Corresponding author. CUIDAR (Children's Hospital of Orange County/University of California-Irvine Initiative for the Development of Attention and Readiness), UCI Child Development Center, 19722 MacArthur Boulevard, Irvine, CA 92616, USA.

E-mail address: klakes@uci.edu (K.D. Lakes).

regulation as referring to “the processes by which the self alters its own responses, including thoughts, emotions, and behaviors” (p. 146).

One important basic form of self-regulation is delay of gratification. Mischel, Shoda, and Peake (1988) measured delay of gratification in 4- and 5-year-old children; more than 10 years later, they found that children who had shown the greatest capacity to delay gratification had the most successful outcomes later in their lives. This important finding supports the assertion that the development of self-regulation is an important aspect of positive child development. Furthermore, Posner and Rothbart (2000) suggested that self-regulation, or effortful control, is linked to empathy, aggression, and conscience, thus making it a key issue in child socialization. Posner and Rothbart argued that “understanding self-regulation is the single most crucial goal for advancing an understanding of development and psychopathology” (p. 427).

According to theorists (e.g., Carver & Scheier, 1981), there are two types of self-regulation failure: underregulation and misregulation. *Underregulation* occurs when the self fails to change its response to that which brings the best outcome. Underregulation is often manifested in procrastination, violence, and binge patterns. *Misregulation* involves efforts that do not bring about the best outcomes, perhaps because of a lack of understanding. Misregulation can contribute to drug and alcohol abuse, “choking” under pressure, and maladaptive goal setting (Baumeister, 1997). Understanding the processes involved in learning to self-regulate effectively may further our ability to design treatments or interventions that promote self-regulatory behavior and reduce pathological or problematic behaviors due to underregulation or misregulation. Posner and Rothbart (2000) argued that “understanding mechanisms of self-regulation in normal individuals will lead to advances in diagnosis, prevention, and possibly treatment of developmental problems like attention deficit disorder and learning disabilities” (p. 427).

Research on self-regulation makes clear that resources for self-regulation (colloquially referred to as “will power”) are limited, and that seemingly unrelated activities (e.g., resisting the temptation of a tempting food, persisting on difficult problem-solving tasks) draw on this same internal resource (Muraven & Baumeister, 2000). Muraven and Baumeister compared psychological will power to a physical muscle, raising the intriguing question whether this “muscle” can be strengthened, and what sorts of activities may promote strength of will and hence enhance self-regulatory abilities.

These findings from basic psychology research raise compelling questions for educators and applied psychologists. What is the role of children’s activities in strengthening or sapping resources needed for self-regulation? What effect does participation in positive (i.e., self-regulation promoting) activities have on outcomes that may be related to will power (e.g., behavior problems, persistence, self-confidence)? In this study, we investigated the effectiveness of an ancient system of self-discipline training—martial arts training—for promoting self-regulation in a nonclinical sample of elementary school children. We hypothesized that traditional martial arts training is a type of programming that incorporates techniques and experiences that are likely to foster self-regulation. For example, martial arts students learn to monitor their thoughts and actions using techniques, such as meditation and self-evaluation, and are rewarded for practice and self-discipline by attaining new levels of mastery. We hypothesized that these techniques would lead to an increase in self-regulation.

1.1. Martial arts as a system for teaching self-regulation

The martial arts historically have emphasized the importance of self-regulation, using terms such as self-control, body control, and discipline. The character training inherent in traditional martial arts teaches an individual to become more self-aware and to actively pursue character growth through the

constant evaluation of thoughts and actions and subsequent adaptation of thoughts and actions for the better. Fuller (1988) commented on the martial arts character training by stating, “From a psychotherapeutic viewpoint, the martial arts may be viewed as formalized, refined systems of human potential training which provide interesting practical models and mechanisms of psychological intervention” (p. 318). In this paper, we will focus on the mechanism of self-regulation as taught in the martial arts and on its impact on various aspects of an individual’s development.

The martial arts have existed for more than 3000 years, and today, there are hundreds of different styles. Richman and Rehberg (1986) suggested that the perseverance and growth of the martial arts over time might provide evidence of their significant physical and psychological benefits. Current research has linked empirical studies and theoretical reviews of the martial arts to self-regulation theory and provides a more thorough review of the processes involved in martial arts training that strengthen one’s capacity for self-regulation (Lakes, 2003). Broadly, outcomes of martial arts training can be classified into two domains: physical (i.e., physical skill and psychological effects related to physical appearance and ability) and psychological (i.e., generalized psychological benefits). In the physical domain, earlier researchers attributed outcomes, such as increased physical confidence (Finkenberg, 1990), improved self-perceptions of physical ability (Richman & Rehberg, 1986), and enhanced body image (Guthrie, 1995) to martial arts training.

Within the general psychological domain, there appear to be affective, cognitive, social, and behavioral benefits from martial arts training. Affective factors associated with martial arts training include higher self-esteem (e.g., Finkenberg, 1990; Richman & Rehberg, 1986; Trulson, 1986), a more positive response to physical challenges (Fuller, 1988), greater autonomy (Duthie, Hope, & Barker, 1978), emotional stability (Konzak & Boudreau, 1984), assertiveness (Konzak & Boudreau, 1984), and self-assurance or self-confidence (Duthie et al., 1978; Konzak & Boudreau, 1984). Cognitive factors influenced positively by martial arts training include concentration (Konzak & Boudreau, 1984) and a greater awareness of mental capacities as well as a cultivation of that potential (Seitz, Olson, Locke, & Quam, 1990). Social benefits of martial arts training include learning to be more respectful of others (Konzak & Boudreau, 1984).

It is important to note that the martial arts studies reported here were predominantly correlational and most did not control for the self-selection bias. In other words, it is possible that individuals in those studies selected martial arts training because it was consistent with their values and beliefs or that the students in the martial arts who did not have these qualities eventually dropped out of training and, therefore, were not among the research participants.

In the present study, we assessed the impact of martial arts on multiple levels. Would martial arts training have some of the impacts cited above when applied to a broader population, controlling for self-selection? If martial arts training does have these effects, what mechanisms bring about these exciting changes? We hypothesized that martial arts training would have these physical and psychological benefits even when applied to a sample controlled for self-selection, and that the mechanism which brings about these changes is self-regulation.

1.2. Anticipated effect size

The mean effects in 68% of all medical, psychological, and educational interventions range between .19 and .75 (Lipsey & Wilson, 1993). Although a study like this one, that is, a physical education class that uses martial arts theory and techniques to teach self-regulation, has not been published in the

scientific literature, intervention and prevention researchers have studied numerous programs whose effects can be used to gain a perspective on common effect sizes. Hattie, Marsh, Neill, and Richards (1997) reviewed 96 studies of adventure education and outward bound programs and found that the programs had an average effect size (standardized mean difference) of $d = .34$. Hattie et al. argued that self-regulation appeared to be the theme that underlies the categories of greatest effects of the interventions (e.g., independence, $d = .47$; confidence, $d = .33$; and assertiveness, $d = .42$). These programs were intensive—the majority (72%) lasted between 20 and 26 days—and the nature of the comparison group (no-treatment or alternative treatment group) was not reported for these studies.

In primary prevention and educational school-based programs, Baker, Swisher, Nadinichek, and Popwicz (1984) reported a mean effect size of $d = .55$ among 41 outcome studies. An important characteristic of the studies reviewed in this meta-analysis was specificity; both interventions and measures targeted specific behavior or knowledge, resulting in stronger effect sizes. Moreover, Durlak and Wells (1997) reviewed 177 outcome studies of prevention programs focused on preventing behavioral and social problems; they reported an overall mean effect size of $d = .29$.

Program evaluations that assess the impact of a certain type of intervention on broad outcome variables often obtain smaller effect sizes. For example, mentoring programs, a very common form of school intervention and prevention, have an average effect size of $d = .14$ (among 59 studies) on a broad range of outcome variables (Dubois, Holloway, Valentine, & Cooper, 2002). Additional study considerations, such as student socioeconomic status and comparison group conditions, further contextualize anticipated effect size in the present study. Dubois et al. emphasized the impact of socioeconomic status on mentoring program effect sizes, reporting that the larger effect sizes ($d = .19$) were found with students from low socioeconomic backgrounds, and smaller effect sizes ($d = .11$) were found with students from higher socioeconomic backgrounds.

Because we used a comparison group that participated in physical education classes, which are known to have positive effects on children, we expected that our effect sizes would be smaller than those we might obtain using a wait-list control group. However, our previous experience as well as our review of the literature led us to believe that this martial arts training intervention would have substantial impact on self-regulatory abilities; therefore, we anticipated stronger effects on direct measures of self-regulation with more moderate effects on indirect measures of self-regulation. We hypothesized that in this sample of students of high socioeconomic status, this intervention would have effects between those expected in a mentoring program ($d = .11$) and in a school intervention program with greater specificity of intervention and effects ($d = .55$).

1.3. The present study

1.3.1. Leadership education through athletic development

To examine the utility of martial arts training for promoting self-regulation among children from kindergarten through Grade 5, we conducted an evaluation of the Leadership Education Through Athletic Development (LEAD) curriculum. LEAD is a program born out of the Moo Gong Ryu (Korean for “guardian of peace style”) martial arts system, designed by Senior Grand Master Joon Pyo Choi, who has studied and contributed to the martial arts for nearly 50 years and who has earned a ninth degree black belt. A primary goal of LEAD and most traditional martial arts is self-improvement, which is generally defined as progression to a higher level of personal character and physical and mental ability. Because of the many misconceptions about the martial arts that are perpetuated in society through media that

emphasize the more aggressive or violent aspects of martial arts and to separate the program from those misconceptions, Grand Master Choi's Moo Gong Ryu martial arts system was deliberately renamed LEAD before being introduced into the school curricula. The martial arts curriculum, instructional setting, and teaching methods were adapted for the school setting by the LEAD instructor (Pasquinilli, 2001).

1.3.2. Research hypotheses

In this study, we sought to confirm the following research hypotheses: (a) LEAD participants would demonstrate enhanced self-regulation in the physical, affective, and cognitive domains in response to a physical challenge relative to participants in the comparison group; (b) LEAD participants would exhibit enhanced self-confidence relative to comparison group participants; (c) LEAD participants would demonstrate amelioration of hyperactivity/inattention symptoms and other negative behavioral/interpersonal symptoms relative to participants in the comparison group; (d) LEAD participants would demonstrate greater improvement in self-regulation than comparison group participants when faced with a cognitive challenge; and (e) LEAD participants would show increased social responsibility when compared to the comparison group participants. Finally, we examined Group \times Gender interactions on all dependent measures. We did not hypothesize differences in a specific direction, but thought it prudent to examine whether gains differed by gender. Because the instructor was male, and because martial arts training may conform more closely to gender norms for boys than girls (e.g., male students predominate in martial arts classes in the United States), we thought it prudent to examine whether boys and girls experienced the LEAD program differently.

To our knowledge, this is the first empirical evaluation of a school-based martial arts program for children. In fact, few previous martial arts research studies included any child participant, and those that did included at-risk or delinquent children only. Another important aspect of this study was the control for self-selection; by using random assignment and a comparison group, we did not face the limitations previous researchers did in their studies of existing martial artists. In addition, this study includes a larger sample size than found in the previous research. Finally, this was one of the first studies to use observer ratings as a measurement tool in addition to previously used self-report measures.

2. Method

2.1. Participants

Participants in this study were 207 students from kindergarten through Grade 5 at a private lower school in a midsize Midwestern city. Two students left due to relocation and 12 students were absent during either the pre- or posttest periods. There were 94 boys and 99 girls in the final sample of 193 students (6% attrition).

The information from the demographic surveys sent to parents at the beginning of the year indicated that 83% of the students were Caucasian, 8% were Asian-American, 2% were African-American, less than 1% were Native American, 2% were identified as having other racial/ethnic backgrounds, and an additional 4% did not respond. Approximately 73% were from families with incomes of more than US\$100,000 per year, and 12% had family incomes between US\$75,000 and US\$100,000. Approximately 15% of the students had family incomes of less than US\$75,000 per year. Between 15% and 20% of the total student body received some form of financial aid from the school.

2.2. Procedures

During the 2000–2001 academic year, martial arts (LEAD) instruction was substituted for the standard physical education curriculum for two or three (depending on a weekly rotating schedule) of the four 45-min physical education periods each week. To evaluate the effectiveness of this experimental program, students were randomly divided by homeroom class into two groups: the first group (LEAD group) participated in the LEAD program two or three periods per week during the first 4 months of the school year, and the second group (comparison group) participated in standard physical education classes during these periods. A comparison of the two groups before and after the LEAD program (first semester) provided a stringent test of the efficacy of the LEAD program, over and above gains in self-regulation attributable to a conventional physical education curriculum.

In mid-August prior to the school year, the parents or guardians of all the students ($N = 208$) received a mailing that included information about the program and the research along with a consent form. Parents were instructed to return the consent form before or on the first day of school; 100% of the forms were returned by the end of the first day of school. Only one parent withheld consent completely; three others gave restricted consent (consent for all but one or two of the measures).

Students were pretested during the first 4 days of school. Testing took place during the physical education hour, with only a few modifications to the regular daily schedule. Immediately following the pretesting, the intervention began. The posttesting was conducted in late January after the LEAD participants had received twenty-six 45-min sessions of LEAD training.

2.3. LEAD program

The LEAD program consisted of courses taught by a martial arts instructor who had held a black belt for more than 10 years and had nearly 10 years experience in instructing adults and children. He had taught children with diverse abilities and backgrounds and currently served as the director of a nonprofit Asian arts academy in Dayton. The instructor's methods and philosophies about teaching martial arts to children have been published in a succinct manual (Pasquinilli, 2001). The techniques taught to the children in the LEAD program included traditional Moo Gong Ryu techniques, such as blocks, kicks, and punches. Forms, a series of martial arts movements and techniques applied in an artistic arrangement of movement, were also emphasized. In addition, children learned board-breaking techniques, complete body-stretching techniques, and deep-breathing relaxation techniques. Finally, all of these techniques were taught in an environment characterized by respect, discipline, and self-control. Children were taught to use their techniques only to protect themselves and never to hurt another person except when absolutely necessary for self-defense.

In this intervention, children attended co-ed martial arts classes with a total of 14 to 16 students per class. Classes were organized based on homeroom classes: each homeroom class attended its own martial arts class two to three times per week. When students entered the gymnasium, all were wearing the standard martial arts uniform and all students began the classes at the same rank or belt level (beginner level, white belt). All classes began the same way: students were told to line up in a particular formation and stand at attention. Students then faced the instructor and bowed to demonstrate respect. At the start of each class, students spent a few minutes sitting in meditation. Students were instructed to clear their minds of thoughts and worries and to focus solely on their breathing. Deep-breathing techniques were taught and reinforced during meditation. Following meditation, students participated in

a class focused on teaching several techniques for the day. Common to the teaching of all techniques was the principle that to progress, the student should ask himself/herself three questions designed to promote self-monitoring: (1) Where am I? (2) What am I doing? (3) What should I be doing? After asking themselves these questions, students were told to correct their stances, behavior, or thoughts, thus regulating their behavior and thoughts in accordance with the expectations of the particular situation. The instructor emphasized that students were responsible for their own behavior, not only in the LEAD class but also in all aspects of their lives. At the conclusion of each class, students lined up and bowed to the instructor as the class was dismissed.

Children in the comparison group participated in a standard physical education curriculum. This curriculum included activities, such as stretching, running, and playing a variety of physical games. In addition, sports, such as basketball and soccer, were taught. The curriculum varied from class to class, providing the children with a broad range of experiences.

2.4. Instruments

The evaluation design incorporated a variety of evaluation modalities and multiple perspectives. Parents provided demographic data via a survey, which was returned directly to us. Trained evaluators from outside the school assessed the children using an observer rating measure and objective standardized measures. Finally, both the participants' academic classroom teachers and the participants themselves completed evaluation instruments.

2.4.1. Demographic survey

The parent survey included questions relating to family constellation, family income, and the racial or ethnic background of the child.

2.4.2. Self-regulation in response to challenge

This construct was measured using the Response to Challenge Scale (RCS). The RCS is an observer-rated measure of children's responding to a physical challenge that was developed for this study. The RCS includes 16 bipolar adjectives (e.g., vulnerable—invincible) rated on seven-point scales. Students completed a challenging obstacle course and were rated by seven independent raters who did not know the children and who did not know which group the children were in (LEAD or comparison). The first author selected RCS items after a thorough review of the self-regulation literature and pilot testing. A detailed analysis of the psychometric qualities of the RCS can be found in Lakes and Hoyt (2003).

The RCS items reflect three domains of self-regulation demonstrated in response to a challenge: physical, cognitive, and affective. The physical subscale (three items) focuses on level of physical control and skillfulness (e.g., awkward—skillful). The cognitive subscale (seven items) reflects ability to focus attention and efforts on the task at hand (e.g., distractible—focused). The affective subscale (six items) assesses self-confidence, emotional control, persistence, and will (e.g., quitting—persevering). Possible scores in all subscales ranged from 1 to 7, with higher scores indicating greater self-regulation in response to challenge. Subscale scores were computed as the mean of the relevant items, and higher scores indicate greater self-regulation on that dimension.

All participants were evaluated on the RCS by seven independent evaluators who were blind to the experimental conditions. Evaluators were primarily psychology students, both graduate and advanced undergraduate students, with the exception of one evaluator who had been educated in a human resources

field and had extensive experience with children. The first author trained these evaluators before the pretest and again before the posttest. Training took approximately 30 min at each testing occasion and consisted of a discussion of the meaning of the words included on the scale with examples of how those characteristics might be demonstrated in children. In addition, evaluators were told to rate the children on a developmentally appropriate level by anchoring their ratings on the first child of a given grade level, in order to rate all children of that grade category in comparison to their age-level peers.

Interrater reliabilities were computed using SPSS by analyzing scale scores and treating raters as the items in a multi-item scale. Reliability was estimated as the relative intraclass correlation coefficient for the composite score based on seven raters [i.e., ICC (3,7); Shrout & Fleiss, 1979]. Seven was the number of raters for virtually all of the LEAD participants at both pretest and posttest. Mean rater agreement averaged across the two assessment occasions was strong: ICC = .92, .93, and .91, respectively, for physical, affective, and cognitive self-regulation. Factor analysis of aggregate pretest scores showed that the cognitive and affective subscales had weak discriminant validity—the items on these subscales loaded on the same factor, and the subscales were highly correlated, $r = .90$. The physical items formed a separate factor that was somewhat more distinct ($r_s = .75$ and $.78$ with cognitive and affective subscales, respectively).

2.4.3. *Student strengths and difficulties (teacher rated)*

Two teachers, a homeroom teacher and another teacher with whom the children had regular contact (e.g., art and language teachers), rated each child in kindergarten through fifth grade using Goodman's (1997) Strengths and Difficulties Questionnaire, Teacher Version (SDQT). The SDQ is a behavioral rating scale with very similar parent and teacher versions that consists of five subscales with five items each for emotional symptoms, conduct problems, inattention/hyperactivity, peer problems, and prosocial behavior. Items are scored on a three-point Likert scale (0 = not true, 1 = somewhat true, and 2 = certainly true). Sample items include "often unhappy, downhearted, or tearful" (emotional symptoms), "often has temper tantrums or hot tempers" (conduct problems), "restless, overactive, cannot stay still for long" (inattention/hyperactivity), "rather solitary, tends to play alone" (peer problems), and "often volunteers to help others" (prosocial). Scores on all five of the subscales of the SDQ range from 0 to 10. Higher scores indicate more problem behaviors for four of the scales; on the prosocial subscale, a higher score indicates more positive behaviors. The subscale score is the sum of individual items on that subscale, and the SDQ total score is the sum of all subscale scores except the prosocial subscale items.

The SDQ has demonstrated good reliability and validity across several studies. In a general population sample of 900 parents (Smedje, Broman, Hetta, & van Knorring, 1999), the SDQ, Parent Version internal consistency reliabilities were .76 (total), .75 (inattention/hyperactivity), .70 (prosocial behavior), .61 (emotional symptoms), .54 (conduct problems), and .51 (peer problems). Smedje et al. (1999) reported a test–retest reliability of .96 over a 2-week period. For the present study, internal consistency reliabilities of the SDQ, Teacher Version were .80 (emotional symptoms), .69 (conduct), .88 (hyperactivity), .69 (peer problems), and .85 (prosocial behavior).

The SDQ is a relatively brief measure of child behavior that overlaps in content with the longer Child Behavior Checklist (CBCL; Achenbach, 1991). Goodman and Scott (1999) reported the following correlations between the SDQ and CBCL on related subscales: .87 (total), .84 (externalizing/conduct problems), .71 (hyperactivity), .74 (Internalizing/emotional symptoms), and .59 (social/peer problems). Goodman and Scott indicated that the two measures discriminated equally well between children.

Goodman and Scott (1999) also reported acceptable and equivalent correlations with maternal interviews for a high-risk sample for both SDQ and CBCL, with one exception. Goodman and Scott found that the correlation between the inattention/hyperactivity subscales and interview ratings was significantly higher for the SDQ ($r = .43$) than for the CBCL ($r = .15$).

To enhance reliability of behavioral ratings, we aggregated ratings from two teachers on each subscale. However, interrater reliability between the two teacher–observers was poor: mean ICC (3,2) across the two rating occasions was .19, .64, .71, .33, and .16, for emotional, conduct, hyperactivity, peer problems, and prosocial, respectively. Thus, consensus among teachers was very weak except for conduct and hyperactivity subscales, for which there was modest agreement between the two raters.

2.4.4. Freedom from distractibility

This measure is a subscale of the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) and includes the Arithmetic and Digit Span subtests. Our primary interest in this measure stemmed from indications that these subscales define a factor that has attention-concentration aspects (Kaufman, 1975) and was useful in evaluating change in a previous attention-training intervention (Kerns, Eso, & Thomson, 1999). However, some investigators have suggested that the subscale also assesses sequencing ability (Bannatyne, 1974), short-term and auditory memory (Cohen, 1957), numerical ability (Lindsey, 1967), and executive processes (Wielkiewicz, 1990). In addition, Lutey (1977) indicated that the Freedom from Distractibility (FD) subtests were the most susceptible to test anxiety.

The correlation between the Arithmetic and Digit Span subtests for children age 6 in the standardization sample was $r = .50$ (Wechsler, 1991). Correlations between the Arithmetic and Digit Span subtests for our sample were .66 for the pretest and .63 for the posttest. The possible range of scores for the Arithmetic and Digit Span subtests is 0–30. We standardized the two subtest scores and summed them to compute FD scores.

The same evaluators who rated the children on the RCS administered this measure. Evaluators received 30–60 min (based on the individual's previous experience with the measure) of training prior to each testing occasion. The training consisted of the standard WISC-III instructions for the two subtests, as well as the opportunity to give several practice administrations and to receive feedback on one's administration of the subtests.

2.4.5. Self-esteem

The Coopersmith (1967) Self-esteem Inventory (SEI) is a 58-item self-report inventory used in this study with fourth and fifth grade students ($n = 66$). Students responded to items by indicating “like me” or “unlike me.” Sample items include, “there are lots of things about myself I'd change if I could” and “I often feel upset in school.” Items were scored as either 0 or 1 and summed to compute a total score, with higher scores indicating higher self-esteem.

Self-esteem is thought to be related to mental health in general, and it may impact a person across a number of domains. Coopersmith (1967) presented evidence that SEI scores were significantly related to academic achievement, resistance to group pressures, and willingness to express unpopular opinions. These are qualities that have important implications for children, specifically their ability to succeed in school and to withstand negative peer pressure.

The SEI is psychometrically sound and has been used widely with children. Spatz and Johnston (1973) reported an internal consistency reliability estimate of .81 for 100 fifth grade students. Kimball (1972) administered the SEI to a group of students including 2909 fourth and fifth graders and obtained

internal consistencies of .92 for fourth graders and .87 for fifth graders. The internal consistency reliability coefficient we obtained for the current sample was .84 for both the pre- and posttests.

SEI scores also appear to remain relatively stable over time. Fullerton (1972) tested 104 children in Grades 5 and 6 with a 12-month period between the test and retest. A coefficient of .64 was obtained. Coopersmith (1967) used a shorter interval of 5 weeks and found the test–retest reliability coefficient to be .88 for 50 children in Grade 5.

2.5. Power analysis

Power is an important consideration in interpreting the findings of this study. With an n of 100 per group (or $N = 200$), our power to detect an effect size of $d = .11$ (comparable to those in other studies, such as mentoring programs) would be 13%. To detect an effect size of $d = .40$ with $N = 200$, our power would be 80%. Because of the possibility of a small effect size due to the strong comparison group, we elected to use an α level of $p = .05$ for all analyses, to avoid further compromising the statistical power of our evaluation. Such a strategy increases the experimentwise Type-I error rate, but may be advisable when small effects are of practical importance (Cohen, 1988).

3. Results

3.1. Overview of analyses

Outcomes for LEAD and comparison group participants were compared using factorial analysis of covariance (ANCOVA), with group (LEAD or comparison group) and gender as the independent variables. Posttest scores were the dependent variables (3 RCS scores, 5 SDQT scores, 3 WISC-III scores, and 1 SEI score) in 12 separate ANCOVA analyses, with pretest scores as the covariate in each analysis. To give a sense of the practical significance of our findings, we report mean residualized change scores for students in the LEAD and comparison groups in Table 1. Change scores were the unstandardized residuals after posttest scores for each measure were regressed onto corresponding pretest scores. Effect sizes are reported as standardized mean differences (d) in these change scores, with positive values of d indicating more favorable outcomes for the LEAD participants for all measures except SDQ inattention/hyperactivity, emotional symptoms, conduct problems, and peer problems. On these SDQ subscales, a negative d indicates more favorable outcomes for LEAD participants. Residualized change scores and effect sizes for (a) the full sample, (b) boys only, and (c) girls only are presented in Table 1.

3.2. Self-regulation in response to a challenge

The RCS results were the strongest indicator of positive change in this study. Children in the martial arts group showed greater self-regulation in response to a challenge than children in the comparison group for all three dimensions of self-regulation [$F_s(1,174) = 11.18, 7.38, \text{ and } 3.93, p_s < .05$ for cognitive, affective and physical self-regulation, respectively]. The means and effect sizes (standardized mean differences in residualized change scores) are reported in Table 1 for each measure of self-regulation.

Table 1

Means (and *SDs*) for change scores by group and gender, with standardized mean differences (*d*) between groups for the full sample (*N* = 193), for boys (*n* = 94), and for girls (*n* = 99)

Instrument	Combined (girls and boys)			Boys			Girls		
	LEAD change	Comp change	<i>d</i>	LEAD change	Comp change	<i>d</i>	LEAD change	Comp change	<i>d</i>
RCS Cognitive ^a	.22 (0.86)	-.26 (1.09)	.49**	.35 (0.72)	-.46 (1.3)	.80**	.12 (0.95)	-.09 (0.82)	.24
Affective	.19 (0.95)	-.22 (1.01)	.42**	.43 (0.83)	-.21 (1.0)	.70**	-.04 (1.0)	-.24 (1.0)	.20
Physical	.13 (0.99)	-.15 (1.0)	.28*	.07 (0.98)	-.37 (0.97)	.45**	.16 (0.99)	.04 (0.98)	.12
SDQT Emotional	-.07 (0.85)	.08 (0.95)	-.17	-.03 (0.86)	-.02 (0.84)	-.01	-.12 (0.86)	.16 (1.0)	-.30
Conduct ^b	-.07 (0.60)	.09 (0.82)	-.23 ⁺	-.14 (0.67)	.26 (1.09)	-.45**	-.02 (0.51)	-.07 (0.43)	.11
Hyperactivity	-.15 (1.71)	.17 (1.51)	-.20	.06 (1.96)	.49 (1.67)	-.24	-.36 (1.4)	-.11 (1.3)	-.19
Peer problems	.09 (0.92)	-.11 (0.89)	.22	.15 (0.85)	.02 (0.95)	.14	.07 (1.0)	-.23 (0.83)	.33
Prosocial	.18 (1.33)	-.20 (1.3)	.29*	.17 (1.2)	-.59 (1.4)	.58**	.12 (1.45)	.15 (1.1)	-.02
WISC III-FD	.02 (0.92)	-.03 (0.78)	.06	.07 (0.92)	.09 (0.85)	-.02	-.02 (.94)	-.13 (.72)	.13
Arithmetic	.08 (0.52)	-.09 (0.49)	.34*	.13 (.45)	.08 (.54)	.42*	.03 (.57)	-.10 (.46)	.25
Digit span	-.04 (0.73)	.04 (0.59)	-.12	-.02 (.79)	.15 (.61)	-.24	-.05 (.67)	-.06 (.57)	.02
Self-esteem	.12 (0.92)	-.21 (1.09)	.33	.07 (0.89)	-.21 (1.08)	.28	.20 (0.98)	-.21 (1.16)	.38

Residualized change scores are residuals from regression of posttest scores on pretest scores for each measure. Actual *Ns* vary across measures due to missing data. *d* = (mean treatment group – mean comparison group)/*SD*_{pooled}. *SD* = standard deviation (in parentheses). LEAD = Leadership Education through Athletic Development (martial arts group); Comp = comparison group. RCS = Response to Challenge Scale. SDQT = Strengths and Difficulties Questionnaire-Teacher version. WISC-III FD = Wechsler Intelligence Scale for Children-Third Edition—Freedom from Distractibility Scale. Self-Esteem = Coopersmith SEI.

^a Significant Gender × Intervention interaction, *p* < .05.

^b Significant Gender × Intervention interaction, *p* < .01.

⁺ *p* < .10. **p* < .05. ***p* < .01.

These group differences were qualified by a significant Group × Gender interaction for cognitive self-regulation and the interaction approached significance for affective and physical self-regulation as well [*F*(1,174) = 4.43, 2.17, and 1.97; *p*s < .05, .15, and .17, respectively]. As shown in Table 1, the effect sizes for boys were generally high, especially for cognitive and affective self-regulation, while effect sizes for girls were low to moderate at best. As the means in Table 1 show, boys in the LEAD group improved more than girls in the LEAD group, although differential gains were significant only on the measure of cognitive self-regulation.

3.3. Strengths and difficulties

Of the five behavioral dimensions evaluated on the SDQT, the main effect for group was statistically significant only for the prosocial subscale [*F*(1,191) = 4.27, *p* < .05] with a relatively small effect size, indicating a greater improvement for the LEAD group relative to students in the comparison condition. In addition, the main effect on the conduct problems subscale approached significance [*F*(1,191) = 2.81, *p* < .10] with a small effect size, suggesting a trend toward the reduction of teacher-reported conduct problems among LEAD participants. These main effects were qualified by a Group × Gender interaction that was significant for conduct problems [*F*(1,191) = 6.63, *p* < .05] and approached significance for prosocial behavior [*F*(1,191) = 3.09, *p* < .10]. On the conduct problems subscale, the effect size was moderate to large for boys, and low for girls (see Table 1). For the prosocial subscale,

there was a large effect size for boys, but a low effect for girls. The mean residualized change scores in Table 1 show that teachers observed fewer conduct problems and marginally higher levels of prosocial behaviors for boys, but not for girls, who participated in LEAD.

3.4. Freedom from distractibility

The Group (LEAD or comparison group) \times Gender ANCOVA conducted on FD total scores showed that the main effect for group was not significant [$F(1,181) = 0.13, p > .10$] and the effect size was very small. However, analysis of the subtest scores demonstrated statistically significant differences between the LEAD group and comparison group in the posttest on the Arithmetic subtest [$F(1,181) = 5.14, p < .05$] with a moderate effect size. There was no evidence of a Group \times Gender interaction on the FD scale or on either subscale.

3.5. Self-esteem

Self-esteem was measured only for fourth and fifth grade participants ($n = 64$). Although the effect size indicates a small to moderate difference favoring the LEAD group, the ANCOVA main effect for group was not significant [$F(1,59) = 1.71, p > .10$]. There was no evidence of a Group \times Gender interaction for self-esteem.

3.6. Post hoc analyses

3.6.1. Gender differences at pretest

To examine whether Group \times Gender interactions were due to floor or ceiling effects on some scales, we compared boys' and girls' preintervention scores on all measures using a series of t tests. The means and results of these analyses, presented in Table 2, show that boys initially scored significantly higher than girls on Arithmetic, RCS-cognitive, RCS-affective, SDQT-hyperactivity, SDQT-conduct, and SDQT-total, and significantly lower on SDQT-prosocial. Higher scores on the SDQT-hyperactivity, conduct, and total subscales indicate more problem behaviors in these areas. A lower score on the prosocial subscale indicates fewer positive social behaviors for boys, on the pretest measures, than girls. Examination of pretest means by gender indicated a potential floor effect for girls on SDQT-conduct, with a mean for girls of .49 ($SD = .91$) on a scale ranging from 0 to 10.

3.6.2. Developmental differences in gains

We conducted post hoc analyses to test for differences in gains by developmental level, that is, grade level. To test for an interaction between grade level and group treatment, we used hierarchical regression analyses on those dependent variables that resulted in a significant main effect for group in our primary analyses (i.e., RCS cognitive, affective and physical subscales, SDQT prosocial subscale, and WISC-III Arithmetic subtest). In these analyses, we entered the covariate (participants' pretest scores) as Block 1, the main effects (group and grade) as Block 2, and the Group \times Grade interaction (product term) as Block 3. The interaction term was significant only for affective self-regulation [$F(1,174) = 5.19, p = .02$]. To interpret this interaction, we computed standardized mean differences (comparing residualized change scores for LEAD and comparison groups) within three developmental subgroups. Results

Table 2
Mean (and *SD*) pretest scores of girls and boys

Instrument/variable	Participant gender					
	Girls		Boys		<i>t</i>	<i>p</i>
	Mean	<i>SD</i>	Mean	<i>SD</i>		
RCS/cognitive	5.44	0.59	5.66	0.52	−2.66	.01**
RCS/affective	5.06	0.78	5.54	0.69	−4.44	.00**
RCS/physical	5.21	0.72	5.23	0.97	−0.23	.82
FD-total	−0.28	1.70	0.31	1.90	−2.25	.03*
Arithmetic	12.61	3.95	14.30	4.66	−2.75	.01**
Digit span	12.18	3.23	12.79	3.67	−1.24	.22
SDQT/emotional	0.97	1.18	0.96	1.52	0.03	.97
SDQT/conduct	0.49	0.91	0.90	1.16	−2.82	.01**
SDQT/hyperactivity	1.82	1.74	3.32	2.63	−4.76	.00**
SDQT/peer problems	1.07	1.10	1.10	1.34	−0.16	.87
SDQT/prosocial	7.78	1.73	7.12	1.90	2.55	.01**
Self-esteem	0.73	0.11	0.74	0.14	−0.384	.70

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

showed that d s = .00, .43, and .83 for younger (kindergarten and first grade), middle (second and third grade), and older (fourth and fifth grade) elementary students, respectively. These results suggest a trend for the LEAD intervention to be more effective in promoting affective self-regulation with older students. As these were post hoc analyses, results should be interpreted cautiously.

4. Discussion

4.1. Differences between groups

The results of this study support the hypothesis that participating in school-based martial arts training would result in improved self-regulatory skills. The results indicate that LEAD participants made greater gains, in comparison with participants in standard physical education classes, in all three areas of self-regulation, with the greatest relative gains occurring in cognitive and affective self-regulation. Significant gains were also observed for LEAD participants over the other children in prosocial behavior. The differences in reduction of conduct problems and in attention scores on the intellectually challenging (math) task posttest also suggested that the LEAD program might influence behaviors in these domains. Finally, although nonsignificant because of the smaller size of this subsample, our findings suggest the possibility of gains in self-esteem among the fourth and fifth grade LEAD participants that are similar in magnitude to those just reported. Future studies of LEAD or similar interventions should also track the influence of such programs on self-esteem.

Although participants in the present study were relatively low risk (e.g., they had low scores on the SDQT checklist of problem behaviors), effect sizes on both positive and negative measures of functioning were strong relative to those typically observed in low-risk populations (Dubois et al., 2002). LEAD appears to be an effective approach to both alleviating problem behaviors and developing positive capacities in children. Favorable effects on participant self-regulation were observed in a

variety of contexts (physical tasks, intellectual tasks, and social settings) using measures derived from multiple sources (self-ratings, teacher ratings, task performance, and behavioral observations). Moreover, the observed effect sizes were relatively large despite the limited LEAD hours (approximately 28 h of LEAD were received by the intervention group throughout the entire evaluation period) and despite the strong comparison group (the comparison group participated in a physical education class, an experience that is known to have positive effects on children; [Council of Physical Education for Children, 2001](#)).

Although we did not analyze specific components of LEAD to determine specific mechanisms of change in this study, we hypothesized that the changes were, at least in part, the result of self-regulation training, inherent in traditional martial arts training and explicitly taught in the LEAD intervention. As noted earlier, in every LEAD class, children were instructed to ask themselves “Where am I?” “What am I doing?” and “What should I be doing?” The purpose of the three questions is to teach children to self-monitor and increase self-regulation. After asking these questions, children were instructed to make a decision to modify their thoughts and behavior to be congruent with the answer to the third question (i.e., What should I be doing?). Asking, “Where am I?” can help a child become oriented to her current context (e.g., “I am in school” or “I am in martial arts class”). The second two questions direct a child to select a target behavior, compare it to his/her current behavior, and prepare to modify his/her behavior. It is likely that there are other aspects of LEAD training, including its use of incremental increases in level of challenge and reward for achieving higher levels of competence that contributed to positive changes. Future studies could benefit from the use of component analyses to further investigate mechanisms of change.

Results of this study have important implications for both practitioners and researchers. Practitioners working with children who have weak self-regulatory skills can apply the described self-monitoring strategy (three questions and a decision to self-correct) to therapeutic and educational settings to promote self-regulatory improvement. In addition, as martial arts classes proved to be a context conducive to teaching self-regulation, therapists and educators should consider implementing similar programs in schools or community settings. Several implications for researchers should be noted as well. For example, researchers might further examine self-regulation training using these procedures (i.e., three questions and a decision to self-correct) in contexts other than martial arts to evaluate the efficacy of the technique in other contexts with other populations. In addition, the use of component analyses in martial arts interventions would help identify and further clarify additional mechanisms contributing to positive changes.

4.2. Differences between boys and girls

Gender analyses demonstrated that both girls and boys in the LEAD group showed benefits relative to the comparison group; however, the effects for boys were numerically larger than the effects for girls on many scales, and significantly so on several. Measures that assessed observer-rated cognitive factors (e.g., focus, concentration, and attention) and teacher-rated conduct factors (e.g., obedience to adults, aggression toward other children, and anger) were particularly susceptible to differential effects of the programs on boys and girls. On these measures, boys showed greater improvements in the LEAD versus comparison groups than girls did. This pattern has important implications given the current literature and national concern regarding the prevalence of attention and conduct problems in boys and longer term outcomes associated with such problems (e.g., [Arnold, 1997](#); [Campbell, Shaw, & Gilliom, 2000](#); [Gaub](#)

& Carlson, 1997; Nagin & Tremblay, 1999). Our results suggest that this intervention is especially well suited to promoting self-regulation in cognitive and behavioral areas with boys and may greatly improve school behaviors, such as attention and classroom conduct.

4.2.1. *Artifactual explanations*

The Group \times Gender interaction raises important questions concerning the self-regulatory gains for girls in martial arts-based interventions. Although findings for girls were in the expected direction (indicating gains relative to the comparison group), they were significantly smaller than the gains for boys on some measures. This may be due in part to the fact that there were significant differences between boys and girls at the pretest. On both measures that detected gender interactions, the means of boys and girls at the pretest were significantly different. The mean for girls at pretest was lower, indicating fewer problem behaviors at the outset of the study, than the mean for boys on the SDQ-T conduct subscale. Thus, the girls' failure to improve much on the SDQT conduct subscale may be attributable in part to floor effects; that is, girls did not have much room for a decrease in problem behaviors. In other words, on the SDQT, boys had more room for improvement, and this could explain the greater effect for boys.

On the cognitive and affective self-regulation subscales of the RCS, boys were rated more highly at the pretest, suggesting that this measure may have tapped into boys' strengths. It is possible that the setting in which this measure was used (performance on an obstacle course) was more anxiety provoking, uncomfortable, or unusual (i.e., challenging) for girls, or simply that those ratings reflect boys' greater physical competence and confidence. In any case, the Group \times Gender interaction on RCS cognitive was clearly not an artifact of ceiling effects: If anything, we would expect this artifact to favor girls, as they had greater room for improvement on this subscale. Therefore, it is possible that additional substantive explanations may account for the differences in effect sizes for boys and girls. Several are discussed below.

4.2.2. *Substantive explanations for gender differences*

Substantive explanations for gender differences favoring boys in this study include (a) differential reactions to martial arts training for boys and girls, (b) differential reactions to the (male) instructor, and (c) differential reactions to the co-ed nature of the classes. First, it is possible that a significant Gender \times Intervention Type interaction exists. That is, it may take more time for girls to fully engage in martial arts, possibly due in part to the fact that certain aspects of the martial arts may contradict gender role socialization for girls. However, [Konzak and Boudreau \(1984\)](#) argued that martial arts training may move participants beyond traditional gender stereotypes and cited evidence that training is associated with both increased sensitivity in males and assertiveness in females. Therefore, we are not suggesting that martial arts training is less beneficial for girls than for boys. Benefits received by girls may be manifested differently (i.e., benefits may be more internal than external) than those obtained by boys and some of these gains may not be as readily evident in girls as they were in boys in this study (i.e., a longer period of intervention may be required to detect certain changes in girls). However, it is also possible that martial arts training is especially conducive to promoting change in boys. Future studies could investigate gender differences in attitudes toward and perceptions of martial arts training to further clarify this issue.

A second possible explanation is that gender differences in gains for LEAD participants may be less likely to occur if instructors of both genders are present. Perhaps a male instructor and role model is

less effective for girls than for boys in this context. The effectiveness of martial arts training is due in part to its inclusion of optimal challenge levels. Elementary school girls may have a harder time envisioning themselves accomplishing a task or learning a skill (especially skills that are physical) when the person demonstrating or teaching the skill is male. Other researchers (e.g., [Bussey & Bandura, 1984](#)) have indicated that young children prefer to emulate same-sex models in social learning. Having an opposite-gender instructor or role model may increase the level of challenge, making a task overly, rather than optimally, challenging. This would reduce the effectiveness of the intervention for girls; perhaps having a female instructor may have helped the girls in this study to benefit from the LEAD experience.

Moreover, participating in a co-ed martial arts class may have affected the girls differently than the boys, perhaps causing them to be less zealous in their participation. In their meta-analysis of adventure education programs, [Hattie et al. \(1997\)](#) noted that single-gender groups had greater mean effect sizes than co-ed groups. Because the majority of the single-gender groups were Australian groups, who had higher effect sizes overall, the effects were confounded. The authors reported that they were unable to contrast the effects from male and female participants crossed with single-gender and co-ed programs because the data were not available. Hattie et al.'s findings support the assertion that the co-ed nature of the LEAD classes may have impacted the gains received by girls, and further research is needed to contrast the effects of single-gender and co-ed martial arts classes on girls.

Therefore, we believe that it is premature to conclude that LEAD is inherently less effective for girls than it is for boys. Each of the possible explanations for the gender differences is plausible. It is clear that more research is needed to address the effects of martial arts interventions by gender. In a 1997 report by the [President's Council on Physical Fitness and Sport \(1997\)](#), the importance of physical exercise for the physical and mental health of girls was emphasized and it was recommended that girls should be encouraged to participate in physical activities and sports at an early age. Future studies should be conducted to test the hypotheses we have discussed; for example, future studies could compare outcomes from co-ed martial arts interventions to same gender interventions or could examine outcomes of interventions that match the gender of the instructor to the gender of the students.

4.3. Developmental effects

Post hoc analyses suggested that there might be differential gains in affective self-regulation based on a child's developmental level, or age. Gains for participants in kindergarten/first grade were small, suggesting either that these children did not experience the benefits of intervention to the same degree, or that there are developmental processes involved that limit the degree to which they might benefit from this particular intervention. Children appeared to benefit more at older ages, with moderate gains by second and third graders and large gains by fourth and fifth graders. These results suggest that LEAD had an impressive impact on affective self-regulation in fourth and fifth graders, with a smaller impact on younger children. The trend toward greater gains in older children was present for cognitive and physical self-regulation as well, although the interaction was not significant. However, these results must be considered cautiously, as the analyses were conducted post hoc. To further investigate intervention impact on gains in self-regulation by developmental level, future researchers should include age or developmental level in their hypotheses. Future studies should clarify at what age students begin to profit from the more structured discipline of martial arts training, as compared with

the less structured, but presumably equally physically active setting of a conventional physical education class.

4.4. Implications

Few physical intervention studies have attempted to reach an entire school of children, as was the case in this study. Most sport programs have self-selected participants, many of whom may have characteristics that predispose them to benefit from sport interventions. In this study, we applied an intervention program, LEAD, to an entire elementary school, randomly selecting those who would participate in LEAD. This is an important distinction between this study and other studies of physical interventions, such as the adventure education studies cited by Hattie et al. (1997). Although the participants in the adventure education studies received benefits in areas associated with self-regulation, we cannot rule out the possibility that the individuals who self-selected to participate in those programs did so because they already possessed a motivation to improve themselves and a propensity toward self-regulation. Perhaps the most exciting implication of the present study is its impact on self-regulation across students, including those who would not have elected on their own to participate in such a program.

There are important distinctions to be made between martial arts training and other sport interventions. First, martial arts programs may be especially effective with a broad range of school-age children because martial arts success is linked with mastery orientation, or the belief that success is the result of effort (King & Williams, 1997). Martial arts philosophy emphasizes effort and determination more than “natural” ability, making it an attractive and feasible activity for children not naturally drawn to sports or for children who do not view themselves as “athletic.” Because progress in the martial arts is based on improvement, not just the achievement of a particular standard, martial arts training may be more amenable to children who struggle to achieve the required standard in other sports. Second, martial arts students receive regular feedback on their progress, detailing areas in which they need further improvement and areas in which they are already improving. Their performance is not evaluated using comparisons to others, but rather using an assessment of their previous performance and individual progress they have made. Third, the emphasis on self-regulation as a core value is not found in all sport interventions. Using techniques, such as meditation and self-monitoring, instructors teach students to be more self-aware and to own the responsibility for correcting their thoughts and behaviors. Self-regulation is valued, modeled, taught, and reinforced in martial arts training. For example, although a student testing for a beginner belt level will be required to meditate for only 1 h, he or she will observe instructors and masters meditating for periods from 12 to 24 h, modeling a significant amount of self-regulation. Finally, both children who are sports minded and those who are not are often very attracted to martial arts, suggesting that this is an activity that can appeal to a wide range of children. This is something that cannot be said of all sport programs, some of which appeal only to a small number of children.

4.5. Qualifications regarding martial arts training

However, because not every martial arts program is the same, it is critical to evaluate training programs individually. Trulson (1986) compared “modern” martial arts classes with “traditional” classes (such as classes in our study) and found that they were not equal in their results. He identified a number of factors present in the traditional classes that did not exist equally in the modern classes. First,

the instructor in traditional classes was a role model of the highest character. Second, the physical conditioning included an integration of psychological and philosophical training. Instructors emphasized respect, humility, confidence, responsibility, honesty, perseverance, and honor. Finally, there was a strong emphasis on self-control and on only using the techniques for self-defense.

These differential factors appear to be critical components of martial arts training. Trulson (1986) divided adolescents who were identified as juvenile delinquents into three groups (traditional Tae Kwon Do, modern martial arts, and a wait-list control), taught by the same instructor. The traditional classes met the aforementioned criteria, whereas the modern classes were solely focused on martial arts as a competitive sport. The boys in the traditional Tae Kwon Do group showed an increase in value orthodoxy, social ability, self-esteem, and a decrease in aggressiveness and anxiety. The boys in the modern martial arts group showed a greater tendency toward juvenile delinquency, increased aggressiveness, decreased self-esteem, and decreased social ability. The wait-list control group remained unchanged. Therefore, to promote positive youth development using a martial arts program, it is very important to utilize a traditional martial arts program.

4.6. Future directions and conclusions

Finally, in future evaluation plans, researchers should attempt to study the long-term effects of participation in programs, such as LEAD. Future studies should include male and female instructors and should attempt to determine the effectiveness of each with boys and girls. Future studies also should attempt to evaluate the effectiveness of martial arts training on self-regulatory abilities of special populations of children, such as those with attention deficit hyperactivity disorder. In addition, effectiveness should be assessed with children of diverse backgrounds, in terms of socioeconomic status and racial or ethnic background. Future research should attempt to evaluate internalizing behaviors, especially in girls. This could be accomplished by including more self-report measures or structured interviews for girls. In addition, it may be important to examine differences in manifestations of self-regulation between girls and boys. Using a broad range of assessment instruments and methodologies appears to be necessary to better understand the effects of martial arts training on boys and girls. In addition to behavioral, parent and teacher report, and self-report measures, future studies could include neuroimaging to further advance our knowledge of the brain systems involved in the development of self-regulation. Finally, future studies should adapt the self-regulation strategies emphasized in LEAD to other settings to evaluate their utility outside of the martial arts.

Results from this study have exciting and important implications for educators and mental health professionals interested in promoting positive youth development. Children can be taught in ways that increase self-regulatory abilities, and this teaching can take a form that is highly attractive to a broad range of children. The interventions can be interesting, enjoyable, challenging, and rewarding, while still accomplishing the objective of enhancing self-regulatory abilities. Because self-regulation is linked to later success in life, it is an aspect of positive child development that is worthy of future consideration.

Acknowledgements

Funding for this study was provided by a grant from the Asian Arts Center to both authors as well as the Virginia Horne Henry Wisconsin Distinguished Graduate Fellowship awarded to Kimberley D. Lakes.

The authors thank Senior Grand Master Joon Pyo Choi for allowing us to conduct research with the Moo Gong Ryu martial arts system and Grand Master Ho Kim for his review of the manuscript. We thank Dr. Kathy Mecoli for her support and administrative assistance and the research assistants who collected the data: Kendra Cassidy, Amy Combs, Deanna Cook, Renee Gentry, Molly Fairchild, Amanda Fetters, Linda Knauer, Andy Olsen, Kate Russet, Diane Smith, Carolyn Welch, April Watson, and Jenny Wood.

References

- Achenbach, T. M. (1991). *Manual for the child behavior checklist/4-18 and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.
- Arnold, D. H. (1997). Co-occurrence of externalizing behavior problems and emergent academic difficulties in young high-risk boys: A preliminary evaluation of patterns and mechanisms. *Journal of Applied Developmental Psychology*, *18*, 317–330.
- Baker, S. B., Swisher, J. D., Nadenichek, P. E., & Popwicz, C. L. (1984). Measured effects of primary prevention strategies. *Personnel and Guidance Journal*, *62*, 459–464.
- Bannatyne, A. (1974). Diagnosis: A note on recategorization of the WISC scaled scores. *Journal of Learning Disabilities*, *7*, 272–274.
- Baumeister, R. F. (1997). Esteem threat, self-regulatory breakdown, and emotional distress as factors in self-defeating behavior. *Review of General Psychology*, *1*, 145–174.
- Bussey, K., & Bandura, A. (1984). Influence of gender constancy and social power on sex-linked modeling. *Journal of Personality and Social Psychology*, *47*, 1292–1302.
- Campbell, S. B., Shaw, D. S., & Gilliom, M. (2000). Early externalizing behavior problems: Toddlers and preschoolers at risk for later maladjustment. *Development and Psychopathology*, *12*, 467–488.
- Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York: Springer-Verlag.
- Cohen, J. (1957). The factorial structure of the WAIS between early adulthood and old age. *Journal of Consulting Psychology*, *21*, 283–290.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Coopersmith, S. (1967). *The antecedents of self-esteem*. Palo Alto, CA: Consulting Psychologists Press.
- Council of Physical Education for Children (2001). *Physical education is critical to a complete education: A position paper from the National Association for Sport and Physical Education* (Retrieved December 15, 2003 from http://www.aahperd.org/naspe/pdf_files/pos_papers/pe_critical.pdf).
- Dubois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology*, *30*, 157–197.
- Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology*, *25*, 115–152.
- Duthie, R. B., Hope, L., & Barker, D. G. (1978). Selected personality traits of martial artists as measured by the adjective checklist. *Perceptual and Motor Skills*, *47*, 71–76.
- Finkenber, M. E. (1990). Effect of participation in taekwondo on college women's self-concept. *Perceptual and Motor Skills*, *71*, 891–894.
- Fuller, J. R. (1988). Martial arts and psychological health. *British Journal of Medical Psychology*, *61*, 317–328.
- Fullerton, W. S. (1972). *Self-disclosure, self-esteem, and risk taking: A study of their convergent and discriminant validity in elementary school children*. Doctoral dissertation, University of California-Berkeley.
- Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*, 1036–1045.
- Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*, 581–586.
- Goodman, R., & Scott, S. (1999). Comparing the strengths and difficulties questionnaire and the child behavior checklist: Is small beautiful? *Journal of Abnormal Child Psychology*, *27*, 17–24.

- Guthrie, S. R. (1995). Liberating the Amazon: Feminism and the martial arts. *Women and Therapy*, 16, 107–119.
- Hattie, J., Marsh, H. W., Neill, J. T., & Richards, G. E. (1997). Adventure education and outward bound: Out-of-class experiences that make a lasting difference. *Review of Educational Research*, 67, 43–87.
- Kaufman, A. S. (1975). Factor analysis of the WISC-R at 11 age levels between 6 1/2 and 16 1/2 years. *Journal of Consulting and Clinical Psychology*, 43, 135–147.
- Kerns, K. A., Eso, K., & Thomson, J. (1999). Investigation of a direct intervention for improving attention in young children with ADHD. *Developmental Neuropsychology*, 16, 273–295.
- Kimball, O. M. (1972). *Development of norms for the Coopersmith SEI grades four through eight*. Doctoral dissertation, Northern Illinois University. Dissertation Abstracts International, 34: 1131–1132, 1973.
- King, L. A., & Williams, T. A. (1997). Goal orientation and performance in martial arts. *Journal of Sport Behavior*, 20, 397–411.
- Konzak, B., & Boudreau, F. (1984). Martial arts training and mental health: An exercise in self-help. *Canada's Mental Health*, 32, 2–8.
- Lakes, K. D. (2003). *The martial arts as a context for promoting the development of self-regulation*. Manuscript in preparation (Available from Kimberley Lakes, PhD, Children's Hospital of Orange County, 455 S. Main Street, Orange, CA 92868).
- Lakes, K. D., & Hoyt, W. T. (2003). *The response to challenge scale (RCS): Technical Manual* (Available from Kimberley Lakes, PhD, Children's Hospital of Orange County, 455 S. Main Street, Orange, CA 92868).
- Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist*, 55, 170–183.
- Lindsey, J. M. (1967). *The factorial organization of intelligence in children as related to the variables of age, sex, and subculture*. Doctoral Dissertation, University of Georgia. Dissertation Abstracts International, 27: 3664–3665, 1967.
- Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, 48, 1181–1209.
- Lutey, C. (1977). *Individual intelligence testing: A manual and sourcebook* (2nd ed.). Greeley, CO: Carol L. Lutey.
- Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology*, 54, 687–696.
- Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, 126, 247–259.
- Nagin, D., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development*, 70, 1181–1196.
- Pasquinilli, M. (2001). *The child whisperer*. Dayton, OH: Asian Arts Center.
- Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, 12, 427–441.
- President's Council on Physical Fitness and Sport (1997, May). *Physical activity and sport in the lives of girls: Physical and mental health dimensions from an interdisciplinary approach*. Washington, DC: Author.
- Richman, C. L., & Rehberg, H. (1986). The development of self-esteem through the martial arts. *International Journal of Sport Psychology*, 17, 234–239.
- Seitz, F. C., Olson, G. D., Locke, B., & Quam, R. (1990). The martial arts and mental health: The challenge of managing energy. *Perceptual and Motor Skills*, 70, 459–464.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 233–244.
- Smedje, H., Broman, J. -E., Hetta, J., & von Knorring, A. -L. (1999). Psychometric properties of a Swedish version of the strengths and difficulties questionnaire. *European Child and Adolescent Psychiatry*, 8, 63–70.
- Spatz, K., & Johnston, J. O. (1973). Internal consistency of the Coopersmith Self-Esteem Inventory. *Educational and Psychological Measurement*, 33, 875–876.
- Trulson, M. E. (1986). Martial arts training: A novel "cure" for juvenile delinquency. *Human Relations*, 39, 1131–1140.
- Wechsler, D. (1991). *Wechsler intelligence scale for children-third edition: Manual*. San Antonio, TX: Psychological.
- Wielkiewicz, R. M. (1990). Interpreting low scores on the WISC-R third factor: It's more than distractibility. *Psychological Assessment*. *Journal of Consulting and Clinical Psychology*, 2, 91–97.