October 2016

The Promise of Character Education in Middle School: A Meta-Analysis

Calvary R. Diggs  
University of Minnesota, diggs042@umn.edu

Patrick Akos  
University of North Carolina Chapel Hill, pakos@email.unc.edu

Follow this and additional works at: https://scholarworks.uvm.edu/mgreview

Part of the Curriculum and Social Inquiry Commons, and the Developmental Psychology Commons

Recommended Citation
Available at: https://scholarworks.uvm.edu/mgreview/vol2/iss2/4

This Research is brought to you for free and open access by the College of Education and Social Services at ScholarWorks @ UVM. It has been accepted for inclusion in Middle Grades Review by an authorized editor of ScholarWorks @ UVM. For more information, please contact donna.omalley@uvm.edu.
The Promise of Character Education in Middle School: A Meta-Analysis

Calvary R. Diggs (University of Minnesota)

Patrick Akos (University of North Carolina Chapel Hill)

Abstract

Early adolescence is a developmental stage characterized by changes in reasoning, social cognition, and desire for autonomy in youth aged 11-14 (or grades 6-8). This period is also associated with heightened impulsivity and risk-taking that has been linked to school-related challenges such as antisocial behaviors and declining grades. Character education, a particular brand of social-emotional practice, has been promulgated as a developmentally responsive program that can promote prosocial behavior and academic success by building upon existing developmental strengths. However, research findings to date are primarily informed by elementary school program outcomes. Due to this limitation, a meta-analytic review of recent research on middle school character education programs and interventions was completed. Findings demonstrate positive associations between character education and academic and behavioral success, as well as social and internal perceptions.

Introduction

What is Character Education

The term, character education, has a rich history with many socio-cultural influences. Character education was first conceptualized by Aristotle, and the term has continued to exist in a variety of forms. Examples range from values inculcation discussions, to cyber ethics, to community service programs utilizing performance character (Auciello, 2007; Ohler, 2010; Smith, 2013b). The character traits adopted and valued by particular constituents can vary across demographics, families, and school contexts such as by race, cultural background, ethnicity, socioeconomic status, or region (Boen, 2010). As a result, a diverse array of character education programs have been established and utilized in school settings – all with varied goals, outcomes, and ways of measuring success (U.S. Department of Education, 2007). In America, this has been the case since the 1920s when character education practices began in schools (Leming, 1993).

In relation to its long history (Leming, 1993), the study of character education has only recently garnered rigorous scientific investigation (Berkowitz & Bier, 2004). In the past 30 years, scientific research has explored character education’s relation to a variety of processes and outcomes.

Schools that effectively utilize character education have reported gains in students’ test performance, appreciation of education, understanding of content knowledge, and GPA (Benninga, Berkowitz, Kuehn, & Smith, 2006; Berkowitz & Bier, 2007; Brannon, 2008; Corrigan, Grove, Vincent, Chapman, & Walls, 2007; Krasmtsova, 2008; Park & Peterson, 2009; Skaggs & Bodenhorn, 2006). Previous research has found an association between character education programs and positive outcomes in student behavioral domains as well. Schools that effectively adopt character education programs have been found to have students who are more on-task. Furthermore, these schools also dispense less referrals, suspensions, and expulsions (Skaggs & Bodenhorn).

These positive outcomes have been associated with changes in teacher and student perceptions of social behavior and character traits. Programs often aim to influence the way that students act or view particular situations in the school settings (Smith, 2013a). In fact, findings support positive associations between character
education program implementation and prosocial behaviors, civic engagement, more respectful classrooms, and students who feel more safe at school (Caldarella, Shatzer, Gray, Young, & Young, 2011; Corrigan et al., 2007; Skaggs & Bodenhorn, 2006). Teachers, students, and school staff have reported more positive feelings towards class and school environments once character education programs were implemented (Skaggs & Bodenhorn). Following this trend, character education programs have been associated with a heightened sense of empathy and social competence for students (Brannon, 2008; Cleary, 2008).

Even in situations where character education programs did not have observable outcomes, teachers still believed in the practice and desired more training (Cleary, 2008). Furthermore, research suggests that character education influences teachers’ general and personal teaching efficacy and other school-related factors such as school climate (Benninga et al., 2006; Brannon, 2008; Corrigan et al., 2007).

Although applications of character education vary, at their core, the mission of character education programs is to influence the development of individual virtues. By proactively strengthening the social and moral decision-making abilities of youth, character education has the potential to work as a form of prevention and/or intervention. This becomes particularly important for the middle grades where early adolescents show increased risk for negative outcomes in health, academics, and social interactions. For example, risk for drug use, declining grades, and interpersonal violence increase during this time, in comparison to elementary school (e.g., Casey, Getz, & Galvan, 2008; U.S. Department of Education, 2011). As a result, character education programs become a topic of interest due to their potential to prepare students to successfully navigate the various challenges they may face in adolescence and beyond.

Describing character education. Knowing the scope and mission of character education programs, a fluid definition of character education is best (Berkowitz & Bier, 2007). To begin, character education is a program or intervention. A program or intervention is anything formally established to achieve a desired outcome. As an example, a fictitious middle school – let us call it Higher Potential Academy (HPA) – implements a program to increase the number of books that its students read.

Character traits are enduring patterns of behavior that can be generalized to a personality characteristic. Therefore, students at HPA do not forge parent signatures in their reading logs, and they read for four hours a day. These students are honest and diligent.

Students use or develop character traits for value decisions. Value decisions are the single situations and subsequent actions that lead to enduring patterns of behavior. For example, each time a student at HPA forgets to get a parent signature on their reading log or does not read, they are faced with a choice. They can forge the signature and the hours read, or they can accept a zero for the day. Over time, in the context of this situation, students could be called honest.

Value decisions can be either moral or nonmoral. A moral decision is one that relates to right and wrong (e.g., adhering to rules). A nonmoral decision does not (Lockwood, 2013). Thus, students at HPA may choose different books to read based on the genre they value; however, this has no moral implication. Students are not right or wrong for selecting mystery over nonfiction. There is no punishment and/or threat to their peers or society as a result of their decision.

There are two noteworthy aspects to be aware of when approaching character education. The first is that character education is positively inclined. Character traits have the ability to be used in negative, nonmoral ways. Continuing the previous example, a student at HPA Middle School may consistently forge the hours and signatures in their reading log by mastering their mother’s signature after spending hours completing a task that is of greater interest. In this regard, the student is dishonest but also still diligent. The assumption of character education research is that these learned traits are positively applied, and little to no attention is devoted to the acquisition of negative traits (unless in the context of developing positive traits to counteract them – i.e., a character education intervention). In this way, character education focuses on the development and maintenance of character traits appropriately applied despite the situation. This leads to the next noteworthy aspect: As alluded to earlier, value decisions can be moral or nonmoral. Students at HPA may
choose to read fiction or nonfiction, but this has no moral implications. However, whether they choose to forge a parent’s signature does. Character education focuses on situations in which values decisions do, in fact, have a moral basis. For example, a character education program may be initiated to reduce the number of students at HPA who forge their parent’s signatures while also promoting a schoolwide climate of honesty and diligence.

Why Character is Important to Study at the Middle Grades Level

Much of the practice and research literature refers to character education as an intervention, or a way to combat developmental challenges by building strengths and skills (Berkowitz & Bier, 2007; Brannon, 2008; Ohler, 2010). Character education is often used in the school setting to address particular issues and challenges that occur throughout development. In America, character education has been used this way in varying degrees and purposes since the 1920s (Leming, 1993).

A common belief held by character education advocates is that the practice is inextricably tied to teaching practices, thereby residing in school curriculum both implicitly and explicitly (Williams, 2000). In essence, what teachers and schools do matters. Character education is a way for schools to formally clarify and intentionally teach skills necessary for school and life success. Although easily said, the implementation of these programs is much more complex with many diverse approaches and mixed findings.

In the context of the rise in evidence-based practices and scientific rigor in education, more concrete data was needed (Williams, 2000). A key question was raised: does character education work and/or help students? If schools are to make formal efforts to engage in character education, they should know if it is a worthwhile use of resources.

This need led to the rise of formal research in character education. Despite positive findings and applications of character education, the majority of research and what is known about best practices is informed by elementary school character education programs. For example, Berkowitz and Bier (2007) included more studies of character education programs for elementary school children than middle or high school. There are considerable differences between the developmental and school contexts of elementary and middle school students, which means that there is likely a need for alternative practices and additional inquiry (Lockwood, 2013).

Character education programs can address a key developmental period: the middle school years. Both early adolescence and character education are a function of the start of autonomous decision making and values development (Thornburg, 1981). Values are freely chosen from alternatives, require consideration of costs and benefits, eventually prized, publically acknowledged, acted on repeatedly, and eventually internalized. When values are based on careful analysis and consideration of alternative possibilities, individuals are more likely to maintain that value, resist persuasion against that particular value, and are more likely to act consistently with the particular value (Lockwood, 2013). In sum, character education holds the potential to influence the value development process of middle school students. Character education could enrich the positive acquisition of character traits, morals, and prosocial behavior development of adolescents in a more meaningful way.

Early adolescents and middle school students are typically more cognitively and socially developed than their younger elementary counterparts. This is evident through a new phase of synaptic pruning and a developing prefrontal cortex among other neural pathways (Casey et al., 2008). Developmental systems theory frames the bi-directional influence between early adolescent neural activity and behavior that can be influenced by the environment (in this case character education) (Lerner, 2006). During this time, adolescents start to notice discrepancies, exceptions, and variability in decisions that are made by those around them and they begin to position themselves on issues based on prior exposure to morals and values. The dual timing of development and contextual change may make middle school the most influential window for the utilization of character education.

Character in Middle School and The Nature of Middle School Students

The middle school years (grades 6-8) and early adolescence (ages 11-14) have been associated with a rise in impulsivity and problem behaviors including delinquency, antisocial tendencies,
and risky behaviors such as substance abuse (Casey et al., 2008). This period is also associated with increased instances of violence and academic dishonesty that can continue on to secondary education (Stephens & Wangaard, 2013; U.S. Department of Education, 2011). Unfortunately, when these challenges are not properly addressed, they can continue to higher levels of education like secondary and postsecondary. One example of this is with academic dishonesty (Stephens & Wangaard).

Despite a variety of challenges present in this age group, there is an important caveat. Early adolescence and middle school can be a challenging developmental stage, but it is not inherently that way. During this time, cognitive and social capacities increase and early adolescents experience more freedom to explore their identity and peer-relationships (Ojanen, Sijtsma, Hawley, & Little, 2010; Roeser, Eccles, & Sameroff, 2000). The effect is that these years are ripe with developmental opportunity for students to branch out and grow.

The implication is that, in theory, character education in middle school could lead to an array of positive outcomes for students, despite an increase in developmental, peer-related, and environmental challenges. Programs could build on this period of student development and growth (Lockwood, 2013).

Character education during the middle school years has the potential to bolster emerging capacities of early adolescents. By simultaneously approaching social and decision-making challenges and nurturing developmental assets, character education takes on greater salience for not only prevention of issues but also the promotion of optimal development (Roeser et al., 2000). However, character education is an umbrella term under the category of social-emotional learning, so there are a variety of programs with varying results and measures of effectiveness. In order to better inform schools and educators concerning evidence for character education, attempts must be made to synthesize the body of existing literature.

Character Education in Middle School: A Meta-Analysis

A study synthesizing data on the effects of character education on middle school students does not presently exist. Berkowitz and Bier (2007) conducted a somewhat similar study concerning the effectiveness of character education programs; however, their focus was primarily describing elements of the existing body of character education in K-12. It did not address the magnitude of effects that character education programs have on students.

Yet, due to the developmental opportunity presented in middle school, we were curious about the effects of character education programs on these students explicitly. The current study conducts a meta-analysis to investigate the effects of character education on middle school students across academic, behavioral, and social and internal perception domains.

Before conducting the analysis, operational definitions of the various constructs were established.

Operational Definitions for Character Education

Character education. A survey of the literature and a close reading of Berkowitz and Bier’s (2007) definition of character education were used to develop the construct for the current study. Berkowitz and Bier’s definition was,

Character education targets a subset of child development. This subset (character development) is the composite of those psychological characteristics that enable and motivate the child to function as an effective moral agent... Character education includes educational initiatives intended to promote such development, and effective character education relies on strategies empirically demonstrated to effectively promote such development. (p. 30)

The reading led to our present definition of character education: Character education is a program or intervention that intentionally aims to influence the way in which students use or develop their character traits for value decisions.

Programs and interventions. Programs and interventions are approaches for developing character to use for future value decisions.

Value decision. Value decisions are the adoption and utilization of moral beliefs that
have implications for a student’s actions and interactions with others.

**Character traits.** Character traits are enduring patterns of behavior based on value decisions over time. The 24 character strengths from Peterson and Seligman’s (2004) classification of character strengths was used to establish these characteristics a fluid definition for what counts as ‘character.’ Peterson and Seligman’s is a well-known and comprehensive review of the research in character strengths (positively applied character traits).

These traits exist within six overarching domains of (1) wisdom and knowledge, (2) courage, (3) humanity, (4) justice, (5) temperance, and (6) transcendence. The specific character traits were creativity, curiosity, judgment, love of learning, perspective, bravery, perseverance, honesty, zest, love, kindness, social intelligence, teamwork, fairness, leadership, forgiveness, humility, prudence, self-regulation, appreciation of beauty and excellence, gratitude, hope, humor, and spirituality. In this way, the meta-analysis was able to accommodate a wide range of studies while adhering to specific constructs and bodies of research.

**Operational Definitions for Descriptive Factors**

**Motivations for the study.** Variability impacts the degree to which educators and education researchers can generalize knowledge concerning the effects of character education programs. Because of the variability of character education programs, we wanted to describe the studies in this meta-analysis in the context of their quantitative effects but also their qualitative elements. This manifested in the development of descriptive categories, which were constructed from past research on character education programs. There were two categories: Developmental Opportunity and elements of Elements of Program Effectiveness. Developmental Opportunity was based completely on Lockwood’s (2013) theory of developmental character education, and elements of Elements of Program Effectiveness were based on Berkowitz & Bier’s (2007) findings concerning common elements of effective character education programs.

**Developmental opportunity.** Lockwood (2013) explored theoretical approaches to developmentally appropriate character education programs. His reasoning was that character education programs should not generalize to students of all ages. This resulted in his proposal of the theory of developmental character education. Under his theory,

Developmental character education is any school-oriented program designed to shape the moral and value understandings and commitments of young people in ways that positively influence their behaviors and engender ethically worthwhile relationships with others and society. (p. 69)

As the complexity of subjects increases with age, maturity, and knowledge, so should character education. Under the theory of developmental character education, students in elementary school should be taught foundational character education skills such as the vocabulary of character and understanding why particular character traits are worthwhile. In the middle and high school grades, they should be taught the nuances of values decisions and explore how to navigate these issues. An example of this could be resisting drug use even though a student’s peer group engages in it.

The curriculum and instruction of these programs intentionally addresses significant developmental differences between young children and adolescents. Lockwood’s (2013) efforts in synthesizing and integrating the literature of both domains culminated in a variety a recommendations concerning how character education can best target students across their educational careers. The section of his recommendations concerning middle school character education programs was integrated into the coding process. Furthermore, these elements relate to other constructs and evidence within the literature (examples: Benson, 2007; Benson, Scales, & Syvertsen, 2011; Kohlberg & Hersh, 1977; Ojanen et al., 2010; Williams, 2000). For the purposes of this study, we renamed it Developmental Opportunity; however, the elements provided in Lockwood’s recommendations are synonymous with the descriptors in this category.

These descriptive categories were meant to provide further details associated with program variability and explore how these theoretically grounded elements manifest in program implementation. In this way, the results of our analysis will not only provide information regarding how much an effect character
education programs have but also some descriptive information regarding the types of programs associated with particular results.

The following elements were hypothesized to be related to developmental opportunity: (1) Moral distinction is when programs establish that there is a clear boundary between right and wrong; (2) Real-world application of issues is when programs tie character development practices to practical situations and/or current events; (3) Explores nuances of choices is when students are taught that character traits can be used both positively and negatively and are taught to critically evaluate the advantages and disadvantages of potential choices; (4) Perspective-taking is when programs acknowledge the developmental capacities of early adolescents such as theory of mind or increased focus on social perceptions; (5) Moral fortitude is present in programs that teach students to engage in perspective-taking without necessarily changing their personal perspective; (6) Personal value reflection resembles values clarification in the sense that it is providing students with opportunities to identify what is important for them and how it impacts others; (7) Acknowledgement of differing peer values is teaching an understanding that peers may make different choices and have different values; and (8) Teachers model character education practices is an element of a program when the school's teachers explicitly engage in the character education practices.

**Elements of effective programs.** While developmental opportunity category focused on the age group of the students, elements of effective programs focuses on specific components of the character education initiative. Essentially, elements of effective programs can be defined as components of the character education programs related to program goals (ex., prevent students from drug abuse), and how they are established (ex., developed by the school or purchased) and implemented (ex., a monthly assembly discussing a character trait vs. daily lessons and discussions with service-learning opportunities). These elements can be thought of as contextual and situational factors associated with a character education program (and not the students themselves).

Berkowitz and Bier's (2007) study originally provided universal themes concerning elements of effective character education programs. Although they did not address middle school character education directly, they are common elements of all character education programs – including middle school. We were interested in how Berkowitz and Bier's elements of effective programs could be used to describe the studies included in the meta-analysis.

Berkowitz & Bier’s (2007) themes of effective character education programs were used to identify potential elements of effective programs: (1) Multifaceted targets at least two levels, (ex., classroom, school-wide, community); (2) Curriculum integration occurs when character education practices are a regular part of the curriculum and does not exist as separate from class happenings; (3) Principal and/or leadership support happens when school administration is involved in the implementation of character education; (4) Primary prevention is established when a program has goals of primary prevention for risk behaviors such as drug use; (5) Staff training requires that staff are trained to teach or administer character education in some way, or they are given professional development opportunities relating to the implementation of character education; (6) Practicing and building character-related skills occurs when students are provided with the opportunity to develop their character; (7) Parent and/or community involvement happens when parents and/or community partnerships are created with the school; and (8) Character trait development requires at least one character trait is targeted and promoted. Programs (9) were also coded for length of implementation which was either less than a year or greater than or equal to a year and (10) whether they were developed by the user (homegrown) or purchased by an outside party.

Two elements of effective programs to note include length of implementation and whether they were developed by the user (homegrown) or purchased by an outside party. The rationale guiding this approach was that it could provide additional information concerning factors related to implementation such as fidelity, dosage, and length of implementation. Although not direct conversions, this information can be found in most of studies and can be related to influential implementation factors.

Berkowitz and Bier (2007) provided recommendations that future research explore the differences in programs. Additionally, Corrigan et al., 2007) assert that character
education is just as effective when homegrown and can be effectively used by rural communities or schools with lower budgets. By coding programs as purchased or homegrown, this meta-analysis also attempts to shed light on these concerns. If character education is effective in homegrown programs, it is possible for schools to practice it without having to purchase expensive curriculums. Likewise, if purchased character education curriculums are equally effective, it will be additional knowledge for schools making decisions concerning character education. For example, if both were equal, instead of purchasing a program, schools could use their resources to pay staff to develop a character education program that fits their school’s particular needs.

Character Education Outcomes

Research can only inform practice when provided with more and better evaluation of what already exists (Leming, 1993). The present study operates on that principle.

Due to these past findings and work within the domain of character education, the current study will investigate the particular effects of character education on outcomes for middle school students. These will span academic, social and internal perceptions, and behavioral domains.

Method

The purpose of this meta-analytic review was to investigate the relationship between character education programs and academic, behavioral, and social and personality outcomes for early adolescents in middle school. Descriptive categories that related to developmental opportunity and elements of character elements of program effectiveness were also examined.

Procedure

  Inclusion and exclusion criteria. Based on Berkowitz and Bier’s (2007) approach, the following criteria were set to determine which programs or interventions were used in the analysis: (1) The program or intervention definition aligns with the constructed definition of character education. (2) The program or intervention targets middle school (grades 6-8) or early adolescents (ages 11-14). (3) A pre-test and post-test design is used to assess the effect that character education has on the desired outcome(s). (4) The academic (e.g., GPA), behavioral (e.g., referrals), or social and internal perception outcomes (e.g., prosocial attitudes) are quantitatively measured. (5) The program or intervention in the study aimed to teach, develop, or strengthen at least one character trait. (6) Studies of interest were published since Berkowitz and Bier’s research (i.e., published in 2006 or later) because Berkowitz & Bier’s findings could have influenced character education program implementation. In this way, findings would reflect the current state of character education.

Any character education study that did not fit all six criteria was excluded from this meta-analytic review. Studies were not restricted to the US; however, for reasons concerning interpretation, studies not in English were excluded from this review. Additionally, programs implemented in non-public school settings were excluded from this study because character education in the context public schools was a primary interest.

Materials

  Articles+, PsycInfo, ERIC, Google Scholar, and the Journal of Research in Character Education were used to conduct this review. These databases were used due to the range of peer-reviewed articles that they contain. Multiple search terms were used to more closely approximate the recall of the true sample of middle school character education studies conducted. Examples include character education, social-emotional learning, and prevention programs.

The initial screening process entailed examining the abstracts and bodies of the articles to find compliance with the first, second, and sixth criterions.

Independent and Dependent Variables

Broadly stated, the current study examined the effect of the independent variable of character education on a number of student outcomes in academic, behavioral, and social and internal perception domains (see Appendix A). Specific indicators of the dependent variables were not rigidly based on a priori criteria. Measures included academics (GPA, state standardized test scores), behavior (referrals, suspensions), and social and internal perception (scores on a variety of attitude and perception scales) outcomes that are used across the population.
Calculating Effect Sizes

Due to the variety of public middle school settings, this meta-analysis assumed a random effects model. Cohen’s (1988) d value was used as the measure of effect size to calculate the difference between pre-test to post-test scores for the variables across domains. As a result, the magnitude of statistically significant change in scores can be determined.

Review Manager version 5.3 was used to conduct this analysis.

Descriptive Categories

Each study’s character education program or intervention was analyzed and coded for particular characteristics across developmental opportunity, Elements of Program Effectiveness, Character Trait, and Implementation Length. A codebook was constructed to maintain definitions and procedures. When the preexisting instructions did not account for issues in the coding process, the codebook was updated with new rationale for the decision reached. As a result, future situations were handled in similar ways.

The character traits used to describe instances of character were gathered from Peterson and Seligman (2004), which should be consulted for additional insight.

Everything, from the Abstract to Conclusion, was used to make decisions concerning the presence or lack of a descriptive category. Typically, information was found in the Methods and Results. Codes functioned as dichotomous variables. Studies were examined and assigned descriptive codes that reflect the presence, at any magnitude of the particular factor.

The principle investigator coded the studies. As advised by Card (2011) and Wilson (2009), each study was coded twice with a one month gap between the coding sessions to account for the potential introduction of measurement error in the process. In order to prevent bias, these studies were unmarked copies that did not contain notes from the previous coding (Card). Agreement rate was used to establish a reliability estimate (Orwin & Vevea, 2009). Agreement rate was calculated by dividing the number of agreements by the total number of studies coded. Intra-rater reliability was 94%.

Results

A total of 11 studies were included in this analysis (k=11), and a total of 112 studies were excluded after initial screenings. This proportion of excluded studies was similar to Berkowitz and Bier’s (2007) proportion of excluded materials, although their sample size was larger given their inclusion of K-5.

The sample of this meta-analysis consisted of students between 11 and 14 years old in middle school. Middle school settings varied in location and demographics; however, all were public school settings. Sample sizes of the studies ranged from 28 to 634 students (see Table 1). Of these 11 studies, 5 of them were included in academic outcome analysis, 9 in behavioral, and 8 in social and internal perception outcomes. Length of implementation ranged from 3 weeks to 3 years.

Tests for homogeneity (Tau²) indicated significant differences between measures. As a result, a random effects model was adopted for analysis. Effect sizes were adjusted and weighted based on sample size.

Effect sizes were interpreted as, Weak is 0.2 – 0.5, Moderate is 0.5 – 0.8, and Strong is 0.8 and higher (Card, 2011; Cohen, 1988). And these effect sizes can be interpreted as the effect (or benefit) in standard deviation units that a character education program had on that particular outcome. For example, among the character education programs included in this analysis, an effect size of -0.2 means there is a one-fifth of a standard deviation unit decrease in negative behaviors from pre-test to post-test. Cohen (1988) recommends that each area of inquiry within a research discipline develop its own standards; however, explanatory character education research has yet to establish these standards. In light of this, we adopted the general recommendations of Card (2011) and Cohen (1988).

As can be observed from Table 2, analysis yielded an effect size between character education and academic outcomes of $d = 0.15$. These results suggest that, although there is a statistically significant effect of character education on GPA, it is not meaningful.

Concerning behavioral outcomes, there is a weak negative effect between character education
programs and the measured behavior outcomes ($d = -0.20$). From pre-test to post-test, there was a decrease in negative behavior in middle school student populations receiving character education. The confidence interval of the true effect size falls between -0.29 and -0.10.

There is a weak effect between character education programs and the measured social and internal perception outcomes ($d = 0.26$). From pre-test to post-test, there was an increase in positive social and internal perception outcomes in middle school student populations receiving character education. The confidence interval of the true effect size falls between 0.09 and 0.66. Finally, the strongest effect was seen in internal factors ($d = 0.45$). With the implementation of character education programs, there was a moderate increase in student life satisfaction, social efficacy, and hope.

**Descriptive Categories**

Following coding, descriptive categories were analyzed for levels of prevalence. This was done by dividing the observed instances by the total studies. The proportion is represented in Table 3. Relevant proportions are listed here as percentages rounded to the nearest whole number: Curriculum integration (45%), parent and/or community involvement (9%), Principal and/or leadership support (18%), Differing peer values (55%), Moral distinction (45%), Moral fortitude (18%), Nuance of choices (27%), Perspective-taking (36%), Personal value reflection (36%), and Real-world application of issues (73%). The most regularly occurring targeted character traits across studies were hope and teamwork (see Table 4). Coding was based on the observable presence of these elements. This does not mean that if a study was not coded as possessing a particular descriptive element, it was not actually present for the study’s implementation or that it did not have an effect on outcomes. It only means that it was not observed during the coding process.

Overall academic outcomes for character education programs that were observed to be integrated into the curriculum ($d = 0.46$) had higher effect sizes than those not a part of the curriculum ($d = 0.12$).

**Discussion**

Previous research suggests that character education can have a significant impact on academic, behavioral and social and internal perception outcomes for students. Our results begin the process of answering questions about the magnitude of the impact these programs have on middle school student outcomes.

Character education does have a significant effect on academic outcomes. This effect, however, was not meaningful. In other words, character education was and has been shown to have a statistically significant effect on academics, yet that same effect was not meaningful in an applied sense. The magnitude of the effect was less robust than predicted. Approaching these findings, from an applied research perspective, character education’s effect on academics, such as GPA, is still meaningful. Slightly affecting a student’s overall academic performance is an important step towards improvement; however, more targeted, evidence-based academic interventions appear to be the best approach to improving student academic outcomes. Character education is important, but it is not a magic bullet to academic difficulty. Despite this, character education appears to have a weak, multifaceted effect on middle school student behavioral and social outcomes.

Character education had a weak overall effect on reducing negative behavioral outcomes in early adolescents. With the implementation of character education programs, schools reported less instances of negative behavior. These effects appeared to be particularly observable in reducing documented instances of student misconduct such as referrals, tardiness, and suspensions. Additionally, character education programs linked to behavior modification interventions did appear to have more favorable results (Caldarella et al., 2011; Lassen, Steele, & Sailor, 2006). That association, however, may lend itself to the benefits of implementing programs that are based on current research and evidence.

The social and internal perception data also had an overall weak effect on student outcomes. These findings suggest that character education has the strongest effects on how a student self-reports about their own character and social strengths. Following the implementation of character education programs, students scored
highest on measures of perceived commitment to school, prosocial behavior, efficacy concerning social interactions, and measures of self-worth and life satisfaction. In essence, character education programs affect how students perceive their own character. The connection to applied social skills is much less meaningful, however.

In addition to these findings, we classified results based on certain descriptive variables (see Table 3). These classifications resulted from the coding process. The results are exploratory, so they are of a less causal nature and provide potential lines for additional research. Intriguing observations are discussed below.

**Exploratory Observations of Descriptive Elements**

Studies with the observable descriptive element, curriculum integration, were associated with stronger program outcomes. For example, Samuelsson (2008) developed the character trait of teamwork through the context of a math class. The effects of that program, in relation to performance measures on math-based assessments, were higher than the general trend for both all and non-integrated academic outcomes. This exploratory result suggests that character education practices show strong academic results when explicitly tied to academic skills.

In a way, this observation is similar to character education programs that integrated behavior modification interventions. Like any other skill, students must also learn to apply character. Guiding students to understand and observe good character in themselves and others is the first step. This mirrors the early stages of Lockwood’s (2013) theory of developmental character education. Next, educators should be providing students with opportunities to develop, practice, and apply those traits. Whether it is learning to regulate behavior or applying character traits in math class, students must receive instruction on how to develop these skills as well as be provided with time to practice them. This observation raises the question: is the application of character in schools domain-specific? If so, this mirrors the findings of Leming’s (1993) research synthesis of character education that the application of character varies by context. Character traits (like honesty or good teamwork) learned in one setting or class does not guarantee its application in another. In this way, character may work best when learned and practiced in a variety of meaningful settings.

Another interesting observation concerns factors outside of schools. Exploratory observations oriented our attention toward parent and/or community involvement. There was only one study in our meta-analysis that explicitly mentioned (and therefore was coded to contain) elements of parent and/or community involvement. Interestingly enough, this study had one of the highest behavioral effects, which were generally weak in other studies (Domino, 2011). For a variety of reasons, parental involvement often declines during the middle school years and is less prominent than in elementary school (Hill & Tyson, 2009). This observation raises an important question for character educators and researchers: Is parental involvement in middle school character education programs important?

Although autonomy is developmentally appropriate for middle school students, having parents and communities involved in the character education process appears to matter. Evidence has been found in previous studies linking family and community involvement and school partnerships to increased student well-being, health outcomes and prosocial behaviors in addition to reductions in aggression, gang activity, and engaging in risky behaviors (Agans, 2014; Bulotsky-Shearer, Wen, Faria, Hahs-Vaughn, & Korfmacher, 2012; Kerr, Shattuck, Kattar, & Uriburu, 2003; Marsh, Foley, & Maddison, 2013; McMahon, Singh, Garner, & Benhorin, 2004). Our exploratory observations cannot provide conclusive explanations – only stimulate discussion. However, to provide an answer to our own question, we would say that parent involvement is important to middle school character education. Involving parents and the community in the character education of early adolescents was worthwhile in previous cases (Domino, 2011). Support outside the school has the potential to bolster and further solidify value stances favored by the school and community at large, thereby increasing the likelihood that students adhere to particular values. However, parent and community involvement do not appear to be regularly discussed in terms of published character education research in middle school.

Another aspect notably absent descriptive element in the studies was principal and/or leadership support. Having the support of school
leadership can positively influence program outcomes as seen in Caldarella et al. (2011) and Lassen et al. (2006). Integrating character education into a regular schedule or into the curriculum is important, and instructional leadership plays an instrumental role in that integration by making character development a school-level effort (Auciello, 2008; Berkowitz & Bier, 2004). In this way, our observations suggest that character education operates at its best when it is part of school culture. These observations create an anecdotal endorsement for the full integration of character into meaningful school processes.

Finally, the appropriateness of character education programs, as it relates to Lockwood’s (2013) theory of developmental character education, is an important concern. In essence, do character education programs address the emerging developmental capacities and opportunities of early adolescents during the middle school years? Our findings are mixed. While many programs were observed to distinguish between moral decisions and also applied character dilemmas to real life, other factors were less represented. Only a small portion of character education programs were observed to integrate the exploration of nuances in value decisions, even less for moral fortitude. The majority of character education programs included in this study appear to altogether neglect the reality that alternative choices exist and that students’ peers can and will make different decisions.

One interpretation is that when programs are implemented at the middle school level, they do not deviate much from models at the elementary school level. This could be related to feasibility on the school level, such as time constraints and individual differences in classrooms.

Another response to this is that, given some of the overarching program themes such as interactive teaching and instruction and integration in the curriculum, there may be core elements linked to teaching these skills well. How these practices are expressed at the elementary and middle school levels may change, and it’s possible that programs did adapt these practices, but those changes were not evident in the published research study. In conjunction with that interpretation, one takeaway is clear. As character education researchers and practitioners, we must think about our methodology, especially as it relates to developmental appropriateness. In order to understand current practices and what works best, we must know about the implementation of character education programs. It is a nuanced field and merits equally sophisticated attention to its processes. And that comes with sound, descriptive methodology. With time, research will be able to refine theories of developmentally appropriate character education.

Limitations and Future Research

The majority of the studies reviewed did not meet criteria for analysis. This is a similar exclusion proportion to Berkowitz and Bier’s results (90%). Additionally, Hippel’s (2015) findings suggest that meta-analyses including seven or fewer studies is fairly common. What this suggests is that meta-analyses with small sample sizes are both common and have merit; however, researchers must be wary of extrapolation when interpreting their results (Hippel; Zoogman, Goldberg, Hoyt, & Miller, 2014).

Given our small sample size, two elements are clear. First, our sample serves as an indicator to the limited scientifically rigorous research conducted on character education in middle schools. Outcome data therefore are quite narrow, with some subcategories composed of one or two studies. Second, given the modest number of studies that fit criteria and that were included in this analysis, conclusions should be made cautiously.

Future studies examining the effects of character education should take a baseline measure before implementation and a measure following implementation for results. This was a common missing element that resulted in excluding a number of studies. As schools increasingly turn to character and other social and emotional developmental programs, a need for evidence will remain. Schools and researchers should consider adopting scientifically-sound experimental designs when implementing and monitoring programs. That way, future researchers pursuing meta-analytic reviews of character education programs can generalize findings.

Likewise, variability in the measures and data collection tools is characteristic in this line of education research. We handled this by creating overarching groups of academic, behavioral, and
social and internal perception outcomes, each with their own subfields of factors, and coding studies for meaningful categorical elements. Regardless of the methodological approach, future researchers should be prepared for various challenges and plan to problem-solve accordingly.

Additional research should investigate, in more detail, how level of curriculum integration affects character education. Provided that curriculum integration was a key descriptive category, exploring the factors that influence it could lead to compelling research. More specifically, research might examine teacher perceptions towards character education, character education’s prevalence throughout the school day, and teacher implementation fidelity. Future findings could yield compelling data that may assist in the implementation and improvement of current and future character education programs in middle schools.

**Conclusions**

Data from this meta-analysis mirrors Berkowitz and Bier’s (2007) findings that character education does impact a variety of outcomes. In addition, it shows the strength of the effect that character education has on academic, behavioral, and social and internal perception outcomes. It appears that character education in middle school is a reasonable means to reduce problem behaviors, increase prosocial behaviors and social cognitions, but does not provide a meaningful effect to academic outcomes. Schools should consider their specific outcome goals before adopting or creating a character education program. Additionally, they should consider the variance in character education implementation such as the degree to which they can make character education a part of their school culture.
References


promote achievement. Development Psychology, 45(3), 740-763. doi: 10.1037/a0015362


Table 1
*Characteristics of Included Middle School Character Education Research Studies*

<table>
<thead>
<tr>
<th>Citation</th>
<th>Treatment n</th>
<th>Mean Age</th>
<th>Homegrown</th>
<th>Academic</th>
<th>Behavioral</th>
<th>Social and Internal Perceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caldarella et al. (2011)</td>
<td>153</td>
<td>N/A - MS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Domino (2013)</td>
<td>160</td>
<td>N/A - 7th</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Gueldner &amp; Merrell (2011)</td>
<td>40</td>
<td>11.5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Kuperminc et al. (2011)</td>
<td>86</td>
<td>11.13</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Lassen et al. (2006)</td>
<td>634</td>
<td>N/A - MS</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Marques et al. (2009)</td>
<td>31</td>
<td>10.96</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Proctor et al. (2011)</td>
<td>218</td>
<td>12.98</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Samuelsson (2008)</td>
<td>119</td>
<td>13</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Savage (2011)</td>
<td>28</td>
<td>11.35</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Turner-Musa et al. (2008)</td>
<td>42</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Willer (2009)</td>
<td>40</td>
<td>12.18</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

*When age is not presented, indicated with N/A followed by grade level or middle school (MS)*

*0 indicates that outcome measure was not present, 1 that it was.*
Table 2
*Effect Sizes and 95% Confidence Intervals for Middle School Character Education in Academic, Behavioral, and Social and Internal Perceptions*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Effect size (d)</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic</td>
<td>0.15*</td>
<td>0.07</td>
<td>0.23</td>
</tr>
<tr>
<td>GPA</td>
<td>0.24</td>
<td>-0.09</td>
<td>0.58</td>
</tr>
<tr>
<td>Math Performance</td>
<td>0.41*</td>
<td>0.31</td>
<td>0.52</td>
</tr>
<tr>
<td>Behavioral</td>
<td>-0.20*</td>
<td>-0.29</td>
<td>-0.10</td>
</tr>
<tr>
<td>Referrals, Tardiness, Suspensions</td>
<td>-0.24*</td>
<td>-0.35</td>
<td>-0.13</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>-0.15</td>
<td>-0.34</td>
<td>0.04</td>
</tr>
<tr>
<td>Social and Internal Perceptions</td>
<td>.26*</td>
<td>0.09</td>
<td>0.66</td>
</tr>
<tr>
<td>Internal</td>
<td>0.45*</td>
<td>0.19</td>
<td>0.71</td>
</tr>
<tr>
<td>Social Behavior</td>
<td>0.17</td>
<td>-0.38</td>
<td>0.72</td>
</tr>
<tr>
<td>Self-worth</td>
<td>0.46</td>
<td>-0.36</td>
<td>1.28</td>
</tr>
<tr>
<td>Social Efficacy</td>
<td>0.42</td>
<td>-0.29</td>
<td>1.13</td>
</tr>
<tr>
<td>Positive Affect</td>
<td>0.02</td>
<td>-0.13</td>
<td>0.18</td>
</tr>
</tbody>
</table>

*Confidence interval does not contain zero
Note: Notable subcategories are italicized and indented under the associated outcome measures
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Curriculum Integration</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5/11</td>
</tr>
<tr>
<td>Parent / Community Involvement</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1/11</td>
</tr>
<tr>
<td>Principal / Leadership Support</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2/11</td>
</tr>
<tr>
<td>Differing Peer Values</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>6/11</td>
</tr>
<tr>
<td>Moral Distinction</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5/11</td>
</tr>
<tr>
<td>Moral Fortitude</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2/11</td>
</tr>
<tr>
<td>Nuance of choices</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3/11</td>
</tr>
<tr>
<td>Perspective-taking</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4/11</td>
</tr>
<tr>
<td>Personal Value Reflection</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4/11</td>
</tr>
<tr>
<td>Real-world Application of Issues</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>8/11</td>
</tr>
</tbody>
</table>

Note: Not all descriptive categories are included in this table, only ones relevant to analysis and discussion. 0 indicates that the descriptive element was not observed in the study, 1 that it was.
Table 4

*Observed Coded Character Traits for Middle School Character Education Programs*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Appreciation of Beauty and Excellence</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Bravery</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Creativity</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Curiosity</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Fairness</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Forgiveness</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Gratitude</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Honesty</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hope</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Kindness</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Love of Learning</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Perspective</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Prudence</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Self-regulation</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Social</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intelligence</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Teamwork</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total Traits</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>24</td>
</tr>
</tbody>
</table>

Note: 0 indicates that outcome measure was not present, 1 that it was.