Gifted Child Quarterly
XX(X) 1–15

DOI: 10.1177/0016986210383155
http://gcq.sagepub.com
SAGE

The Effects of Acceleration on High-Ability Learners: A Meta-Analysis

Saiying Steenbergen-Hu[1] and Sidney M. Moon[2]

Abstract

Current empirical research about the effects of acceleration on high-ability learners' academic achievement and social–emotional development were synthesized using meta-analytic techniques. A total of 38 primary studies conducted between 1984 and 2008 were included. The results were broken down by developmental level (P-12 and postsecondary) and comparison group (whether the accelerants were compared with same-age, older, or mixed-age peers). The findings are consistent with the conclusions from previous meta-analytic studies, suggesting that acceleration had a positive impact on high-ability learners' academic achievement (g = 0.180, 95% CI = –.072, .431, under a random-effects model). In addition, the social–emotional development effects appeared to be slightly positive (g = 0.076, 95% CI = –.025, .176, under a random-effects model), although not as strong as for academic achievement. No strong evidence regarding the moderators of the effects was found.

Putting the Research to Use

The findings of this meta-analysis suggest that acceleration influences high-ability learners in positive ways, especially on academic achievement. An important message for educators, parents and students is that high-ability learners can benefit from acceleration both in the short-term and in the long run. Specifically, accelerated students tend to outperform students who are not accelerated in their performance on standardized achievement tests, college grades, degrees obtained, status of universities or colleges attended, and career status. Accelerants equal or surpass non-accelerants in self-concept, self-esteem, self-confidence, social relationships, participation in extracurricular activities, and life satisfaction. It is informative for policy-makers that acceleration programs, especially university-based early college entrance programs, have been frequently assessed and appear to be the most effective. In summary, acceleration can be effective both in K-12 education and in college. Parents are encouraged to consider acceleration for their academically talented children and educators are encouraged to make acceleration options available.

Keywords

acceleration, meta-analysis, high-ability learners, academic achievement, social emotional development

Acceleration is defined as a type of educational intervention based on progress through educational programs either at rates faster than or at ages younger than one's peers (Pressey, 1949). Acceleration is commonly referred to as academic acceleration (Southern & Jones, 1991). The effects of acceleration, particularly on student achievement and social–emotional development, have been extensively studied. A growing body of research reviews has focused on integrating the results of acceleration. *A Nation Deceived: How Schools Hold Back America's Brightest Students* (i.e., The Templeton National Report on Acceleration; Colangelo, Assouline, & Gross, 2004) presented a recent comprehensive review of acceleration.

However, the number of extant meta-analytic studies on the effects of acceleration on high-ability learners is limited. A review of the literature only identified four meta-analyses or best-evidence syntheses. They are the studies of Kulik and Kulik (1984), Rogers (1991), Kent (1992), and Kulik (2004). These meta-analytic studies have some limitations. For instance, all the four meta-analytic studies are heavily based on early studies of acceleration in the previous century. The earliest study included was published in 1928 (see Rogers, 1991, for reference), and the most recent study included was published in 1991 (see Kent, 1992, for reference). Kulik's (2004) meta-analysis was a reanalysis of the studies included in the three previous meta-analytic studies. This means that

[1]Duke University, Durham, NC, USA
[2]Purdue University, West Lafayette, IN, USA

Corresponding Author:
Saiying Steenbergen-Hu, Department of Psychology & Neuroscience, Duke University, 417 Chapel Drive, Box 90086, Durham, NC 27708, USA
Email: ss346@duke.edu

the most recent studies in these meta-analyses were conducted more than 15 years ago.

The contextual and methodological discrepancies of these previous meta-analytic studies led to the need for a new comprehensive meta-analysis. First, previous meta-analytic studies have different emphases. For instance, Kulik and Kulik (1984) investigated the effects of acceleration on elementary and secondary school students; Rogers's (1991) best-evidence synthesis examined the effects of educational acceleration on gifted students; Kent (1992) focused on the effects of acceleration on the social and emotional development of gifted elementary students; and Kulik (2004) studied the academic, social, and emotional effects of acceleration. In addition, even when they had very similar study foci (e.g., all the four meta-analytic studies examined social–emotional development), they included different sets of studies and reached different conclusions. A new meta-analysis would refine and extend these previous meta-analytic studies.

In the field of gifted education, researchers and educators have also provided a sound rationale to encourage the application of meta-analysis. For example, Asher (1986) noted that meta-analysis enabled educators and researchers to overcome interpretation obstacles such as imprecise measurement and the small sample sizes of gifted education research. Lauer and Asher (1988) defined meta-analysis as "a systematic, replicable, and relatively unbiased method of summarizing the overall results of a particular body of experimental research literature" (p. 284). Vaughn, Feldhusen, and Asher (1991) stated that meta-analysis was a more comprehensive method to conduct program evaluations in gifted education. Again, Asher (2003) encouraged meta-analysis in gifted education. He asserted that the generalization of the findings of a meta-analysis would be greater than any individual study; thus, meta-analysis would provide the best evidence available to build bases for theory development.

This meta-analysis addresses the following research questions:

Research Question 1: How does acceleration affect high-ability learners' academic achievement?
Research Question 2: How does acceleration affect high-ability learners' social–emotional development?
Research Question 3: What differences exist between content-based acceleration and grade-based acceleration (National Work Group on Acceleration, 2009) in terms of their effects on high-ability learners?
Research Question 4: What moderators are significantly associated with the effects of acceleration on high-ability learners?

Conceptual and Operational Definitions

Academic achievement. Academic achievement is one of the most important outcome variables in acceleration research. In this meta-analysis, the definition of academic achievement relied on Rogers's (2004) summarization of the indicators of academic effects and the predictors of probable student success, which included (a) process and achievement well above age peers, (b) mastery well above grade/age level in a specific subject area or topic, (c) being two or more grade levels ahead, and (d) ability/achievement in the upper 3%.

Based on the existing research literature, the outcomes in this study were categorized into P-12 level and postsecondary level academic achievement. For the P-12 level, the outcomes mainly included test results and the status of higher-education institutions to which accelerated high-ability learners were admitted. For the postsecondary level, there were four subgroup outcomes (a) educational background (e.g., degrees obtained, the status of higher-education institutions attended), (b) college grade point average (GPA), (c) ages when certain degrees were obtained or any career achievement was attained, and (d) career status.

Social–emotional development. The term *social–emotional development* is used in this meta-analysis because this term has been most commonly used in the existing education and psychology literature (Kent, 1992; Kulik, 2004; Neihart, 2007; Robinson, 2004). As Neihart (2007) stated, social effects were typically measured via "social maturity scores, teacher ratings of social skills, participation in extracurricular activities, and leadership positions held" (pp. 331-332), whereas emotional effects were usually examined by "measures of self-concept or teacher or parent ratings of risk taking, independence, and creativity" (p. 332). Self-reported questionnaires are often used to measure social–emotional development. Some relatively widely used inventories include the Minnesota Multiphasic Personality Inventory (Tellegen et al., 2003) and the California Psychological Inventory (Aiken, 2004).

High-ability learners. The subjects of the primary studies are high-ability learners. The term *high-ability learners* primarily refers to academically gifted/talented students (cf., United Sates Department of Education, 1993). This definition is preferred because it stresses the significance of potential and emphasizes comparisons between gifted students and others of equivalent age, experience, or environmental background. It also implies that giftedness exists in all cultural groups and across all socioeconomic levels. This definition is widely accepted in the field of gifted education. Furthermore, the term *high-ability learners* is used to avoid biased perceptions or opinions, which may subconsciously consider gifted students as those with an impressive IQ or extremely extraordinary achievement records.

Method

Study Search

Studies conducted during the years from 1984 to 2008 were the focus of this meta-analysis. The year 1984 was chosen as a cutoff time because the first meta-analysis of the effects of acceleration was published in 1984 (i.e., Kulik & Kulik, 1984). We also checked to see whether there were studies appearing in the period between 1983 and 1984 while Kulik and Kulik's (1984) meta-analysis was in preparation or press. No study published during those 2 years that met all our inclusion criteria was identified. The following procedures were used to search studies for inclusion: (a) a search of electronic databases including Education Resources Information Center (ERIC), Web of Science, Psychological Abstracts Index (PsycINFO), and ProQuest Dissertation and Theses; (b) web searches using the Google and Google Scholar search engines; (c) a hand search through three primary empirical journals in gifted education: *Gifted Child Quarterly, Roeper Review*, and *Journal for the Education of the Gifted*; (d) manual scrutiny of reference/bibliography lists of the relevant previous systematic reviews to locate additional studies; and (e) personal contact with researchers in the gifted education field to solicit ongoing research.

Study Inclusion and Exclusion Criteria

Study inclusion criteria were formulated in light of the research questions. The criteria covered the issues of study design, populations, interventions, comparisons, and outcome measures. Under the guidance of these criteria, decisions about whether to include or exclude studies were based on reading the full text of the study. Study inclusion and exclusion decisions were made by the first author and discussed with the second author.

1. *Scholarship:* Studies had to be empirical in nature. Publication venues included both peer-reviewed journal articles and other unpublished research documents such as dissertations, conference presentations, or research technical reports.
2. *Relevance:* The target subjects of acceleration described in each study had to be high-ability learners and the identification procedures for the high-ability participants had to be specified. Studies outside the field of gifted education were included so long as they dealt with acceleration of high-ability learners and met the other inclusion criteria.
3. *Appropriate comparison groups:* The comparison groups and the accelerated groups had to be matched in terms of subjects' major aptitudes. Appropriate comparisons included, for example, accelerated high-ability learners versus nonaccelerated high-ability learners or accelerated high-ability learners versus nonaccelerated ones (usually older peers) in the equivalent grade level. Studies using inappropriate comparison groups, such as comparisons between talented and average youngsters, were excluded, except where appropriate statistical control was applied in the comparison.
4. *Evident study design:* Studies had to have a recognizable study design, such as an experimental, a quasi-experimental, or a causal–comparative design.
5. *Measured outcomes:* At least one of the two major effects of acceleration on high-ability learners had to be reported in the studies: academic achievement and/or social–emotional development, both for the accelerated group and for the comparison group.
6. *Enough quantitative information:* Studies had to report credible, quantitative information, either in descriptive or in inferential form (e.g., means, standard deviation, t values, F values, ANOVA tables, or p values), so that effect sizes could be calculated or estimated.
7. *Completeness and nonredundancies*: To avoid the situation of a single study exerting a disproportionate influence on the overall results, if the same study was reported in several articles, only the most complete version was included for analysis. However, when an individual article reported the results from several studies, only those results that were relevant and met all the inclusion criteria were included.
8. *Language*: Studies had to be available in English. Studies from countries outside the United States but presented in English were also included.

Study Coding

Study coding was conducted using a coding protocol, which was created with a series of successive efforts. First, a draft of the coding sheet was created. Then, pilot coding practices were performed with six studies, and this led to some revisions. Furthermore, a review and discussion of the revised coding sheet was conducted with an expert on meta-analysis. The coding protocol contains information about the major characteristics of the primary studies. These characteristics included (a) the general features of the studies (e.g., the forms of acceleration, the study focus, the construction of comparison groups), (b) the methodological features of the studies (e.g., research designs, sample sizes, statistical analysis methods, outcome measures), and (c) the effect size information (e.g., the magnitude of the effect sizes, effect size extraction methods). The coding protocol containing full descriptions of all the major characteristics can be obtained from the first author.

Study coding was conducted in two steps. For each study, the information presented was first recorded in text in a Microsoft Excel spreadsheet as raw data. The raw data were then coded into numerical data, following the coding protocol. For example, for the Chilton (2001) study, the study focus variable was first recorded as "both academic achievement and social–emotional development" in the raw data sheet. Then, the number "3" was given to this variable. Later, the raw data were rechecked to examine the coding accuracy after all the coding was finished. No major error was found. Generally, coding decisions were based on the original claims in the primary studies, and some efforts were made to verify the original claims. The time it took to code one study ranged from 1 to 8 hours. The total time cost for coding all the 38 studies was 103.7 hours. The average time for coding one study was 2.73 hours. Independent double-coding of two (5.3%) randomly selected studies was conducted. Out of all possible codes, the two coders had an agreement rate of 75%.

Effect Size Calculation

Hedges's (1981) *g* was chosen as the effect size index for continuous outcome measures. Hedges's *g* is an appropriate effect size index for this meta-analysis because there are no universal metrics with which to measure the diverse study variables, such as achievement scores or social–emotional development in existing acceleration research. The preference for Hedges's *g* over other alternative indices, such as Cohen's *d* and Glass's Δ, was based on the fact that Hedges's *g* can be corrected to reduce the bias that may arise when the sample size is very small (i.e., $n < 40$; Glass, McGaw, & Smith, 1981). The What Works Clearinghouse (2007) adopted Hedges's *g* as the default effect size measure for continuous outcomes in its review. The choice of Hedges' *g* was an important advantage for this study because the samples in many studies of high-ability learners are quite small.

Log odds ratio was used as the effect size measure for dichotomous outcomes. Although both Hedges's *g* and log odds ratio appeared in the process of effect size extraction, Hedges's *g* was chosen as the primary effect size index. There were two reasons for doing so: (a) the majority of the effect sizes extracted from the included studies were Hedges's *g* and (b) the software (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2006) provided various effect size formats (including both Hedges's *g* and log odds ratio) for every effect size, so choosing one of them to present another was feasible.

Extracting and Combining Effect Sizes Within Studies

In the preliminary stage, effect sizes or subeffect sizes were extracted for all the outcome variables in the primary studies. With a few exceptions (e.g., Hsu, 2003; Callahan & Smith, 1990), multiple or at least two subeffect sizes were extracted from each of the primary studies. The effect sizes extracted at this stage are called preliminary effect sizes. As recommended by Hedges, Shymansky, and Woodworth (1989), the preliminary effect sizes were averaged within each primary study. This procedure is called effect size combination and was performed after all the preliminary effect sizes were extracted.

Grouping Effect Sizes Across Studies

The preliminary effect sizes were then grouped by each major outcome category (i.e., academic achievement and social–emotional development). When a study provided results of both outcome categories, two overall effect sizes were extracted. For example, the Ambruster (1995) study produced two effect sizes, one for academic achievement and the other for social–emotional development effects.

Investigations of academic achievement outcomes were subdivided into two groups by the developmental level studied: P-12 and Postsecondary Achievement. The outcome variables investigated at these two developmental levels varied slightly. Investigations of acceleration at the P-12 level tended to focus on student SAT scores, cumulative GPA, number of awards earned, composite score on achievement tests, or subject test scores. Investigations of acceleration at the postsecondary level used outcomes such as university credits earned, college GPA, mean age of graduation from college, graduating with honors, or pursuing graduate degrees.

In addition, studies of both types of outcomes (i.e., academic achievement and social–emotional development) were subgrouped by the type of comparisons used: comparisons of accelerated high-ability learners with their same-age peers, older peers, or "mixed-age" peers. The subgroup of "mixed-age peers" was used to describe studies that reported little or no information about the age of the comparison groups.

Computing Combined Effect Sizes

Combined (average) effect sizes were obtained through the CMA software (Borenstein et al., 2006), using studies as the unit of analysis, under random-effects models. A random-effects model was chosen because the existing literature shows that there may be other significant additional sources of variance in effect sizes resulting from factors other than the sampling errors. These factors might include the types of educational settings in which the studies were conducted, the length of the acceleration program, and the sample sizes. Therefore, the true effect of acceleration on high-ability learners could be a distribution of effect values. In addition, the coding phase of this meta-analysis revealed that there was a great deal of variation among the studies of the effects of acceleration in terms of study features, providing further support for the choice of a random-effects model.

Heterogeneity Analysis

The purpose of heterogeneity analysis is to examine whether sampling error alone might be responsible for the variance among the effect sizes (Cooper 2007; Hedges et al., 1989). A set of heterogeneity statistics were computed through the CMA software. These included the Q statistic, p values, and I-squared (I^2) values. A Q statistic represents the variance associated with a group or subgroup of effect sizes that are used to calculate the combined effect size of the group or subgroups. A statistically significant heterogeneity test suggests that there is a high degree of heterogeneity among the group of effect sizes tested. I-squared values quantitatively indicate the degree of heterogeneity, with a range from 0 to 100. If a high degree of heterogeneity exists, further analyses are conducted to search for possible moderators.

Testing for Moderators

Our goal in terms of moderators was to identify variables relating to the general and methodological features of the primary studies that might moderate the effect of acceleration on high-ability students' academic and social–emotional development. An analysis of variance (ANOVA) analog was used to test for moderators in this meta-analysis. The relevant statistics include Q_b and p values. The Q_b statistic denotes the variance that is caused by the variation between/among all the subgroups of the effect sizes when these effect sizes are grouped by the coding categories for a certain variable. The p value denotes the results of a significance test for the difference between or among the subgroups.

Assessing Publication Bias

Two approaches were used to detect and correct for publication bias. First, a funnel plot (Borenstein et al., 2006; Cooper, Hedges, & Valentine, 2009) was produced. The diagnostic function of a funnel plot is based on the key idea that if the studies distribute symmetrically around the mean effect size, it can be concluded that there is no publication bias; if more studies (usually small studies) asymmetrically cluster in the bottom of the plot, there might be publication bias. In addition, Duvall and Tweedie's (2000) trim-and-fill procedure was conducted to further assess and adjust for publication bias.

Meta-Analytic Interpretation Strategy

A "meta-analytic thinking" approach was adopted for interpreting the results in this meta-analysis (cf., Cumming & Finch, 2001; Thompson, 2002). As Thompson (2006) recommended, two questions should be asked when interpreting effect sizes: (a) Was the effect size noteworthy? (b) Were the effects consistent with the related prior literature? In educational research, even small effect sizes may be considered noteworthy, because in many cases, multiple factors contribute to an observed outcome and a single or even a few interventions may be found to have only a limited impact on the outcome (Ahadi & Diener, 1989; Berliner, 2002; Strube, 1991). Specifically, in this meta-analysis, the results were reported and interpreted in terms of the direction and magnitude of the combined (average) effects and the heterogeneity of effects, rather than solely in terms of their statistical significance. In addition, as Borman and Grigg (2009) proposed, narrative descriptions were also used for the interpretation of effect sizes as needed.

Results

Study Overview

A total of 38 studies were identified that met our inclusion criteria. For academic achievement, the most common outcome variables in the primary studies were standardized achievement test results, college GPA, educational background (e.g., the degrees obtained, the status of higher-education institutions attended), career status, ages when certain degrees were obtained or when some career goal was reached. For social–emotional development, the most widely studied variables were self-concept, self-acceptance, self-reliance, self-esteem, self-confidence, social relationship, participation in extracurricular activities, locus of control, life satisfaction, and educational or vocational plans.

In terms of study focus, among the 38 studies included, 15 (39.5%) investigated academic achievement effects, 11 (28.9%) examined the effects of social–emotional development, and 12 (31.6%) provided results for both effects. The most common forms of acceleration included early kindergarten/school/college entrance (14 studies), subject matter acceleration (10 studies), multiple forms (6 studies), advanced placement (AP)/dual credits/international baccalaureate (2 studies), curriculum compacting (2 studies), mentoring (2 studies), grade skipping (1 study), and early graduation (1 study). These forms were determined based on three information resources: (a) the descriptions of the 18 forms of acceleration provided by Colangelo et al. (2004) in *A Nation Deceived: How Schools Hold Back America's Brightest Students*, (b) the terms used by the authors in the primary studies, and (c) the actual content of the acceleration interventions that was described in the primary studies.

Descriptive Results

A total of 274 preliminary effect sizes (i.e., effect sizes before combination) were extracted. On average, there were 7.2 effect sizes extracted from each study. The multiple effect sizes from each study usually indicate that there were various outcome variables used in the primary study. For example, three preliminary effect sizes were extracted from the Janos and Robinson (1985) study because this study reported three outcome variables (i.e., college GPA, university credits earned, and Concept Mastery Test score). For academic achievement, the 141 effect sizes ranged from –4.145 to 3.843, including 108 positive, 4 zero, and 29 negative values. For social–emotional development, the 133 effect sizes ranged from –0.746 to 1.281, including 81 positive and 52 negative values.

Academic Achievement

As discussed in the Method section, we reported and interpreted the effect sizes using current guidelines for the interpretation of the results of a meta-analysis, that is, primarily in terms of their direction, their magnitude, the heterogeneity of effects, and the practical or policy importance of their magnitude, rather than solely in terms of statistical significance (Borman & Grigg, 2009; Thompson, 2006).

Table 1. Summary of Combined Effect Sizes for Academic Achievement

Outcome Group	Number of Studies	Combined Effect Size	Standard Error	95% Confidence Intervals Lower Limit	95% Confidence Intervals Upper Limit	Test of Null (Two-Tailed) *p* Value
Achievement						
Overall	28	0.180	.128	–.072	.431	.162
Achievement						
P-12 achievement	21	0.147	.164	–.174	.467	.370
Postsecondary achievement	7	0.313	.294	–.262	.889	.286
Achievement						
With same-age peers	13	0.396	.187	.029	.762	.034*
With older peers	9	0.224	.222	–.212	.660	.315
With mixed-age peers	6	–0.323	.265	–.842	.197	.223
Postsecondary achievement						
With same-age peers	4	0.498	.277	–.045	1.042	.072
With older peers	4	0.255	.266	–.266	.776	.337
With mixed-age peers	1	0.294	.490	–.667	1.255	.549
P-12 achievement						
With same-age peers	9	0.347	.261	–.165	.858	.184
With older peers	6	0.334	.319	–.290	.959	.294
With mixed-age peers	6	–0.320	.314	–.936	.295	.308

*$p < .05$.

Combined effects. Table 1 presents the overall combined effect sizes for academic achievement. The overall combined (average) effect size for academic achievement was positive (g = 0.180, 95% CI = –.072, .431, 28 studies), although it was not statistically significant (df = 27, $p > .05$). It is relevant to note that among the 28 studies that provided results for academic achievement, three studies yielded obviously bigger effect sizes than did the remaining 25 studies. Specifically, the Hsu (2003) study yielded the lowest negative effect size (g = –2.493), whereas the Janos (1987) study yielded the highest positive effect size (g = 1.809), followed by Callahan and Smith (1990; g = 1.671). A close examination of these three studies revealed that in each of them, only a very limited number of outcome variables were investigated. This resulted in a limited number of preliminary effect sizes being extracted. It is possible that when there are a limited number of preliminary effect sizes involved in the combination of the effect size for a study, extremely big (either positively or negatively) values are yielded because of lack of balance between or among the preliminary effect sizes. The exceptionally low effect size from the Hsu (2003) study seems to be especially counterintuitive. We found that for the Hsu (2003) study, a potential positive preliminary effect size was not calculable because of lack of information. As a result, only a single negative preliminary effect size (g = –2.493) was extracted, and this was the only effect size that could be used to represent this study. When the Hsu (2003) study was deleted from the analysis, the combined effect size was increased (g = 0.240, 95% CI = .128, .358, 27 studies), and it was statistically significant (df = 26, $p < .001$), which is more consistent with the results of prior meta-analyses. However, because we used random-effects models in this meta-analysis, assuming that the effects may vary across contexts, and also because all meta-analyses are based on calculable quantitative information from primary studies, we decided to include all three studies in subsequent analyses.

As noted earlier, the effect sizes for academic achievement were also analyzed by developmental level (i.e., P-12 Achievement and Postsecondary Achievement). For P-12 Achievement, the combined effect size was 0.147 (95% CI = –.174, .467, 21 studies). For Postsecondary Achievement, the combined effect size was 0.313 (95% CI = –.262, .889, 7 studies). Neither combined effect size was statistically significant at the $p < .05$ level. In addition, no statistically significant difference in the combined effect sizes was detected between the P-12 Achievement and the Postsecondary Achievement levels (df = 1, $p > .05$).

The effect sizes for academic achievement were also analyzed by comparison group (whether the accelerants were compared with same-age, older, or mixed-age peers). The combined effect size for the subgroup of Achievement With Same-Age Peers was the highest (g = 0.396, 95% CI = .029, .762, $p < .05$, 13 studies), and this result was statistically significant at the .05 level. Positive results were found for studies comparing accelerants with older peers (g = 0.224, 95% CI = –.212, .660, $p > .05$, 9 studies). However, for the subgroup of Achievement With Mixed-Age Peers, the combined effect size was negative (g = –0.323, 95% CI = –.842, .197, $p > .05$, 6 studies). These results may suggest that the positive academic effects of acceleration are more discernable when high-ability learners are compared with same-age peers and that negative effects of acceleration on achievement are more likely in studies with mixed-age group comparison.

Table 2. Summary of Heterogeneity Analysis for Academic Achievement

Outcome Group/Subgroup	Number of Studies	Effect Size	*Q* Value	*df*	Heterogeneity *p* Value	I^2
Achievement						
Overall	28	0.180	640.700	27	.000***	95.786
Achievement						
P-12 achievement	21	0.147	620.092	20	.000***	96.775
Postsecondary achievement	7	0.313	18.711	6	.005**	67.934
Achievement						
With same-age peers	13	0.396	90.218	12	.000***	86.699
With older peers	9	0.224	15.310	8	.053	47.746
With mixed-age peers	6	−0.323	411.967	5	.000***	98.786
P-12 Achievement						
With same-age peers	9	0.347	69.897	8	.000***	88.555
With older peers	6	0.334	8.822	5	.116	43.322
With mixed-age peers	6	−0.320	411.967	5	.000***	98.786
Postsecondary achievement						
With same-age peers	4	0.498	13.812	3	.003**	78.279
With older peers	4	0.255	22.697	3	.000***	86.782
With mixed-age peers	1	0.294	0.000	0	1.000	0.000

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

Again, when the results for P-12 Achievement were grouped by comparison groups, three subgroups were created (see Table 1). None of the combined effects of these subgroups were found to be statistically significant at the significance level of $p < .05$. No statistically significant difference in the combined effect sizes was detected among the above three subgroups. The results for Postsecondary Achievement were also grouped by comparison groups. None of the combined effects of the subgroups was found to be statistically significant at the significance level of $p < .05$. No statistically significant difference in the combined effect sizes was detected among the above three subgroups.

Heterogeneity analysis. Table 2 presents a summary of the results of heterogeneity analysis for academic achievement. As Table 2 shows, the overall achievement outcomes, outcomes for both developmental levels, and outcomes for two of the three types of comparison exhibited high degrees of heterogeneity. The only relatively homogeneous achievement outcomes were those where high-ability peers were compared with older peers.

Cumulative analysis. Cumulative analysis was conducted to detect the emerging change of the combined effects over time as each newer study was included in the analysis. An obvious pattern was that the accumulated effect sizes for academic achievement were consistently higher through the 1990s than in the 2000s. Specifically, from 1990 (i.e., Brody, Assouline, & Stanley, 1990) to 1999 (i.e., Washington, 1999), a total of 12 studies appeared successively. As each newer study was added, the accumulated effect was found to be positive and statistically significant at the significance levels of $p < .05$, $p < .01$, or $p < .001$. Also, the confidence interval lines associated with these 12 accumulated effect sizes lay on the right side of the zero line, indicating that the effects of acceleration favored the treatment group during this period. From 2001 (i.e., Chilton, 2001) to 2008 (i.e., Wells, Lohman, & Marron, 2008), 10 studies appeared successively. With the exception of two cases (i.e., Chilton, 2001; Moon & Callahan, 2001), all the other accumulated effects were not statistically significant.

Assessment of publication bias. The assessment of publication bias for the meta-analysis of academic achievement was performed on the 28 studies included in this group. Among the 28 studies, 18 (64.3%) were published as journal articles, 8 (28.6%) were doctoral dissertations, and 2 (7.1%) were current studies obtained from researchers. First, a funnel plot was produced with Hedges's *g* plotted against its precision (i.e., 1/standard error). The majority of the studies clustered asymmetrically in the bottom part of the funnel. This might suggest the presence of publication bias. It could also indicate that studies with small numbers of subjects were a major contributor in this meta-analysis. This finding confirmed an existing perception that acceleration research has often been conducted with small sample sizes in the field of gifted/talented education. Furthermore, Duvall and Tweedie's trim-and-fill procedure was conducted. A funnel plot produced after Duvall and Tweedie's trim-and-fill procedure still suggested the presence of publication bias. Statistics showed that under the random-effects model, the point estimate and the 95% confidence interval for the combined effects was 0.180 (−.072, .431), whereas after the trim-and-fill procedure, the combined effects became −0.059 (−.276, .159).This implied that the original combined effect size may be overestimated and that it may be difficult to publish studies that show negative effects on achievement. At the very least, it can be concluded that small sample size effects do not account for publication bias alone in this case.

Summary and conclusions related to achievement. The findings from this meta-analysis suggest that, overall, acceleration

Table 3. Summary of Combined Effect Sizes for Social-Emotional Development

Outcome Group	Number of Studies	Combined Effect Size (Hedges's *g*)	Standard Error	95% Confidence Intervals Lower Limit	95% Confidence Intervals Upper Limit	Test of Null (Two-Tailed) *p* Value
Social–emotional						
Overall	22	0.076	.051	–.025	.176	.139
Social–emotional						
With same-age peers	9	0.141	.079	–.013	.295	.072
With older peers	10	0.052	.083	–.111	.215	.533
With mixed-age peers	3	–0.036	.124	–.280	.208	.773

does improve high-ability learners' academic achievement. The findings reveal that acceleration had a positive influence on high-ability learners' academic achievement (g = 0.180, 95% CI = –.072, .431). The 28 effect sizes that contributed to the combined effect ranged from –2.493 to 1.809, with 0.294 as the median effect size. Five effect sizes were negative, and the remaining 23 were positive. Effect sizes most frequently occurred in the 0.301 to 0.353 range, with 7 effect sizes in this category. Moreover, positive effects were found for two developmental stages, the P-12 (g = 0.147, 95% CI = –.174, .467, 21 studies) and the Postsecondary (g = 0.313, 95% CI = –.262, .889, 7 studies). The effects of acceleration appeared to be more discernable when accelerated high-ability learners were compared with their nonaccelerated same-age peers. An examination of the methodological features of the included studies indicated that studies were conducted with more rigorous methods since 2000's. Overall, the findings from this study are consistent with those from previous meta-analytic studies, such as Kulik and Kulik (1984), Rogers (1991), and Kulik (2004), which found that acceleration had a significant positive impact on high-ability learners.

This finding is also partially inconsistent with those from some previous meta-analytic studies. For example, in terms of the subgroup of "with older peers," there is a discrepancy between the findings from this study and those from prior meta-analytic studies. Specifically, when Academic Achievement Overall, P-12 Achievement, and Postsecondary Achievement were sorted by comparison group, the subgroups of "with older peers" exhibited positive effects in all three analyses (i.e., g = 0.224, g = 0.334, and g = 0.255, respectively). However, in two previous meta-analytic studies (i.e., Kulik, 2004; Kulik & Kulik, 1984), an effect size of –0.04 was yielded when older-age comparison groups were used.

It should be emphasized that inferences derived from the subgroup analyses should be viewed with caution and need to be replicated by future research. In addition, when only a limited number of studies were involved in the subgroup analysis, studies with extreme effect sizes, such as the Hsu (2003) study, may have had a strong influence on the results. This is particularly troubling in the case of Hsu (2003), for the reasons previously elaborated. Because Hsu (2003) investigated P-12 academic achievement with mixed-age peers, the results related to these two categories need to be interpreted with caution.

Social–Emotional Development

Combined effects. A separate meta-analysis was conducted on the social–emotional development effects. Table 3 summarizes the major statistical information on the combined effects. As Table 3 shows, the overall combined effect size for social–emotional development was slightly positive (g = .076, (95% CI = –.025, .176, 22 studies). The effect was not statistically significant (df = 21, p > .05). When the effect sizes for social–emotional development were analyzed by comparison group, for the subgroup of same-age peers, the average effect size was 0.141 (95% CI = –.013, .295, 9 studies); for the subgroup of older peers, the average effect size was 0.052 (95% CI = –.111, .215, 10 studies); and for the subgroup of mixed-age peers, the average effect size was –0.036 (95% CI = –.280, .208, 3 studies). None of the results for the three subgroups was statistically significant (p > .05).

Heterogeneity analysis. There were two subgroups of effect sizes exhibiting high degrees of heterogeneity. These two subgroups were Social–Emotional Overall (g = 0.076, Q = 43.515, df = 21, I^2 = 51.741, p < .01) and Social–Emotional With Same-Age Peers (g = –0.036, Q = 26.928, df = 8, I^2 = 70.292, p < .01).

Cumulative analysis. The cumulative analysis of the combined effect sizes for social–emotional development sorted by year of study revealed that none of the accumulated effects was statistically significant as the 22 studies appeared successively, although the first two accumulated effect sizes (as the first two studies, Janos & Robinson, 1985; Robinson & Janos, 1986, were involved successively) appeared to be larger than the rest. There was no other obvious change pattern observed in terms of the magnitudes of the accumulated effect sizes as the newer studies appeared over the time period.

However, starting from the early 1990s through the 2000s, the standard errors associated with the accumulated effect sizes showed a decreasing pattern. Specifically, starting from the Swiatek and Benbow (1991) study to the Stamps (2004) study, 15 studies appeared. During this period, the standard errors consecutively decreased from .128 to .036, with the exception of two cases. This suggests that research on the effects of acceleration on social–emotional development might have been conducted with relatively higher precision in recent years, possibly because the researchers used larger sample sizes and/or more reliable measurement tools.

Assessment of publication bias. The assessment of publication bias for the meta-analysis of social–emotional development effects was conducted on 22 studies. Among these studies, 15 (68.2%) were published journal articles, 6 (27.3%) were doctoral dissertations, and 1 (4.5%) was a current study obtained from researchers. A funnel plot was produced for a visual inspection of publication bias. The majority of the studies appeared symmetrically around the mean effect line of the funnel plot. This suggested that there was no publication bias. However, the relatively low precision shown in the funnel plot indicated that, just like in the meta-analysis of academic achievement effects, small studies once again played a major role in the meta-analysis. This leads to further confirmation of the perception of small-sample-size issues in gifted/talented education research.

A funnel plot was created after Duvall and Tweedie's trim-and-fill procedure to further assess publication bias. This funnel plot looked almost identical to the one produced before the trim-and-fill procedure. The statistical information yielded with the trim-and-fill procedure also showed that the point estimate and the 95% confidence interval for the observed and adjusted effects were the same: 0.093 (–.025, .176). Therefore, it can be concluded that publication bias was not of concern in this case.

Summary and conclusions related to the social–emotional effects of acceleration. The results from this meta-analysis suggest that the effects of acceleration on high-ability learners' social–emotional development were slightly positive, although not as large as the effects on academic achievement. For example, the overall combined effect size for social–emotional development was 0.076 (95% CI = –.025, .176, 22 studies). The effect sizes associated with this combined effect exhibited a high degree of heterogeneity. However, compared with academic achievement, the heterogeneity of effect sizes in the social–emotional groups appeared to be a little lower. No extreme effect sizes existed. The 22 effect sizes that contributed to the combined effects ranged from –0.528 to 0.664, with 0.058 as the median effect size. Seven effect sizes were negative, and the remaining 15 were positive. Effect sizes most frequently occurred in the 0.024 to 0.083 range (6 effect sizes) and the 0.105 to 0.174 range (6 effect sizes). The assessment of publication bias suggested that publication bias was not of concern for the social–emotional findings. Furthermore, compared with prior meta-analytic studies, a more positive impression of the effects of acceleration on social–emotional development was found in this meta-analysis, perhaps because the related cumulative analysis revealed that research on the social–emotional development effects of acceleration might have been conducted with higher precision (indicated by smaller standard errors) since the early 1990s than before.

When combined effects were assessed in terms of comparison groups, the effects appeared mixed. Specifically, comparisons with same-age peers yielded a positive effect (g = 0.141, 95% CI = –.013, .295, 9 studies), followed by those with older peers (g = 0.052, 95% CI = –.111, .215, 10 studies), whereas comparisons with mixed-age peers exhibited a slightly negative effect (g = –0.036, 95% CI = –.280, .208, 3 studies). However, the latter group was investigated in very few studies, so the results may be different in future studies. Overall, the results derived from this meta-analysis were slightly more positive than in previous meta-analytic studies, in which claims were made that acceleration had mixed effects on social–emotional development (e.g., Kulik, 2004) or did "no harm" (e.g., Kent, 1992). At the very least, the results of this meta-analysis support the notion that acceleration is not harmful to social–emotional development.

Testing for moderators

On the basis of the heterogeneity analysis results, testing for moderators was done only on the groups with high degrees of heterogeneity and consisting of more than 10 studies (cf., Higgins & Green, 2006). Three groups of studies met these criteria: P-12 Achievement (21 studies), Achievement With Same-Age Peers (13 studies), and Social–Emotional Development (22 studies). To ensure meaningful testing, we regrouped the original coding by combining or deleting some subcategories of certain variables in the original coding protocol. Furthermore, for each group, we selected the variables that had at least two subcategories where each subcategory had at least 3 studies in it. As a result, there were 14 variables qualified for moderator testing for the P-12 Achievement group, 9 variables for the Achievement With Same-Age Peers group, and 10 variables for the Social–Emotional Development group. Except in a few cases, most of the variables tested in the three groups were overlapping. Table 4 shows all the variables tested and the testing results.

Unfortunately, some variables we would like to have tested for moderators were not reliably reported in the studies and so could not be included. For example, among the 38 primary studies included in this meta-analysis, in more than two thirds of the studies (71.1%), the researchers presented no specific information about the ethnicity of the samples; in close to two thirds of the studies (68.4%), the researchers provided no specific information about the socioeconomic status (SES) of the study subjects; in more than half of the studies (63.2%), the researchers reported no specific information about school type; and in about one third of the studies (28.9%), the researchers provided no information about the gender ratio of the subjects.

Table 4. Results of Testing for Moderators

Variable	P-12 Achievement			Achievement With Same-Age Peers			Social–Emotional Development		
	g	Number of Studies	*p*	*g*	Number of Studies	*p*	*g*	Number of Studies	*p*
Acceleration category	0.231	18	.578				0.087	21	.070
Grade-based acceleration	0.243	7					0.146	15	
Content-based acceleration	0.060	11					−0.052	6	
Acceleration duration	0.286	17	.491	0.518	9	.820	0.100	21	.244
Less than 28 weeks	−0.180	5		0.397	3		0.228	3	
More than 28 weeks	0.292	12		0.522	6		0.070	18	
Grade level at acceleration[a]	0.165	16	.722	0.281	11	.698	0.078	19	**.046***
Elementary school and earlier	0.181	9		0.195	5		0.107	5	
Middle and high school	0.008	7		0.300	6		−0.116	6	
Postsecondary[b]							0.206	8	
Study time	0.197	19	.736	0.350	12	.468	0.093	19	.153
Study conducted within 1 year after acceleration	0.041	8		0.289	8		−0.024	8	
Longitudinal studies	0.199	11		0.479	4		0.150	11	
Gender[c]	0.314	9	**.019***				0.158	11	.102
Male predominant (more than 60%)	0.421	5					0.336	5	
Female predominant (more than 60%)	0.198	4					0.011	6	
Study form[d]	0.157	21	.945	0.319	13	**.008****	0.096	22	.221
Published	0.142	13		0.635	6		0.125	15	
Unpublished	0.162	8		0.118	7		−0.030	7	
Research design	0.168	21	.367						
Experimental	−0.086	3							
Quasi-experimental	0.555	3							
Causal–comparative research	0.110	15							
Sample size	0.217	21	.502						
≤500	0.229	15							
>500	−0.214	6							
Statistical analysis	0.322	18	.636	0.330	12	.059			
Student-level *t* test, ANOVA, or MANOVA	0.299	12		0.625	6				
Student-level ANCOVA, MANCOVA, logistic regression analysis	0.354	6		0.264	6				
Data source	0.352	11	.368	0.404	8	.468			
Standardized tests	0.371	5		0.433	4				
Nonstandardized tests	0.158	6		0.217	4				
Measurement reliability status	0.183	17	.793	0.354	13	.301	0.095	19	.866
Low or moderate reliability	0.190	11		0.227	8		0.097	10	
High reliability	0.026	6		0.470	5		0.076	9	
Study validity status[e]	0.201	21	.741	0.339	13	.484	0.082	22	.808
Low or moderate validity	0.090	10		0.405	6		0.102	14	
High validity	0.214	11		0.243	7		0.078	8	
ES reporting[d]	0.165	21	.951	0.334	13	**.032***	0.073	22	.721
Yes	0.167	9		0.124	6		0.087	18	
No	0.146	12		0.568	7		0.053	4	

a. This moderator appeared in the Social–Emotional Development group.
b. Only the Social–Emotional Development group has this subcategory.
c. This moderator appeared in the P-12 Achievement group.
d. These two moderators appeared in the Achievement With Same-Age Peers group.
e. Study validity status denotes five kinds of validity, as proposed by Kelley (2007). They are internal validity, construct validity, statistical validity, external validity, and descriptive validity. Study validity status was measured through a scale adopted from Kelley (2007). An average score was calculated for each study to indicate its study validity status. A value between 0 and 2 represents low validity, a value between 2.1 and 3 represents moderate validity, and a value between 3.1 and 4 represents high validity. Among the 38 studies included, 2 (5.3%) had low validity, 19 (50%) had moderate validity, and 17 (44.7%) had high validity.
$^*p < .05$. $^{**}p < .01$.

The answer to the question about the possible differences between content-based acceleration and grade-based acceleration was drawn from the results of the moderator analyses (see Table 4). The results suggested no significant difference between the effects of grade-based and content-based acceleration in terms of academic achievement and social–emotional development. As no previous meta-analytic studies had explored this issue, no comparison could be made. However, we would like to offer a few possible explanations for this finding. First, in practices, grade-based acceleration is usually mixed with content-based acceleration. It is often the case that high-ability learners with experience of content-based acceleration also undergo grade-based acceleration along their educational journey, or there are students whose experience of grade-based acceleration is the result of some previously accumulated content-based acceleration. As a result, it is probably hard to disentangle the effects of the two. Second, grade-based acceleration and content-based acceleration were not rigorously defined and differentiated in many primary studies, partly for the reasons discussed above. The coding of these two categories of acceleration in this meta-analysis was mainly based on the insufficient descriptions from the primary studies. It is possible that the ambiguous coding led to a failure to detect the difference. Finally, there were not enough studies included in this meta-analysis to enable detection of the difference. In the three analyses testing the difference between content-based and grade-based acceleration, there were only 21, 18, and 9 studies involved, respectively. It is possible that future analyses based on an adequate number of studies would to able to detect the difference between the two types of acceleration.

Summary and conclusions related to moderators. The results of the moderator testing suggest that there is no strong evidence that any factor significantly moderates the effect of acceleration on high-ability learners' academic achievement or social–emotional development. Although, as Table 4 indicates, four variables (i.e., gender, study form, effect size reporting, and grade level at acceleration) were identified as moderators, this finding ought to be interpreted holistically. Of the 14 variables tested in the group of P-12 Achievement, only gender appeared to be a moderator; of the 9 variables tested in the group of Achievement With Same-Age Peers, only study form and effect size reporting appeared to be moderators; and of the 10 variables tested in the group of Social–Emotional Development, only grade level appeared to be a moderator. In other words, at a significance level of .05, only four variables appeared as moderators among the 33 variables tested. Taking into account the matter of chance, the presence of moderating influence can be considered as minimal. Furthermore, the moderating influence of none of the four moderators identified was replicated in any other groups tested. Taken together, there is no strong moderating influence on the effect of acceleration as identified in this meta-analysis.

However, two findings deserve mention. First, published studies were found to have higher average effect sizes than did unpublished studies. This finding provides further evidence for the presence of publication bias relating to the studies of academic achievement as previously discussed. It also confirms some similar existing findings in the literature. Second, studies that reported effect sizes in the original reports were found to have lower average effect sizes than those with no effect sizes reported directly.

Discussion

Study Limitations

Interpretation and application of the findings from this meta-analysis need to be undertaken with consideration of the study limitations. Apart from the general limitations associated with the methods of meta-analysis (see, e.g., Glass, McGaw, & Smith, 1981), this study is subject to the following four additional limitations. First, no strong evidence of coding reliability was provided in this study. Although the coder agreement rate obtained is close to satisfactory (75%), the confidence level of this information is quite restricted, because only a very limited number of studies (2 out of 38) were double-coded. However, a few additional procedures were carried out to increase the coding reliability. For example, pilot coding was performed on a set of studies, and further revision was done based on the pilot experiences. Moreover, during the coding process, a second check was performed when the coded text information was further coded into numerical data. No minor errors were found. These efforts increased the study coding reliability.

Second, a sensitivity analysis was not undertaken in this meta-analysis. As described in the Method section, it is believed that a random-effects model would fit better than a fixed-effect model for this meta-analysis. As a result, no analysis was conducted with the fixed-effect or mixed-effects models, and no comparison was made thereafter. It should be noted, however, that Duvall and Tweedie's trim-and-fill procedures, which were used in this meta-analysis, can be considered an alternative sensitivity analysis.

Third, this meta-analysis is subject to problematic definitions of acceleration practices. This limitation is closely related to the existing situation in the acceleration research literature. One example would be that researchers usually provided no clear definitions of the acceleration interventions they investigated, nor did they give a detailed description of the acceleration practices in their study. This posed some difficulties for study coding. Furthermore, researchers usually claimed that their study effects were due to the acceleration intervention that the subjects experienced when the studies were conducted, ignoring what had happened previously. For instance, researchers who chose early college entrants as their study subjects would typically claim that their study was designed to investigate the effects of early college entrance on high-ability learners, regardless of the fact that some of their subjects were able to enter college early because they had had at least one other acceleration experience earlier in their educational career, such as grade skipping, subject matter acceleration, or advanced placement (AP). It seems that the observed effect in most studies of postsecondary students might involve an accumulated effect of multiple acceleration experiences, but this was not investigated in any of the studies. Therefore, the findings from this meta-analysis may only serve as a rough picture of the effects of acceleration. Solutions to this limitation require researchers to provide more precise and complete descriptions of the acceleration interventions.

Finally, the findings from this meta-analysis do not represent all the potential information that the included studies could provide. In other words, more findings could be drawn from the coded data with additional effort. As previously described, a total of 274 preliminary effect sizes were extracted from the 38 studies, serving as a rich source of data. In this study, separate meta-analyses were conducted on the effects of academic achievement and social–emotional development. As a result, findings were derived in terms of these two outcome types. Within each of the outcome types, however, information regarding specific outcome variables was available. For example, for academic achievement, further meta-analyses could be conducted to obtain the combined effect in terms of some of the most common outcome variables, such as standardized achievement test results or college GPA. With regard to social–emotional development, there was even more variation in the measurement of outcomes, so meta-analyses could be conducted on some comparatively common subconstructs or variables, such as social relationships, participation in extracurricular activities, mental maturity/adjustment, locus of control, life satisfaction, self-concept, self-acceptance, self-esteem, self-confidence, and educational/vocational plans. These findings, if obtained, would be very informative to future acceleration research and practices.

Implications for Future Research

One implication of this study for future research is that researchers in gifted and talented education need to do a better job of specifying important demographic information so that these variables can be analyzed as potential moderators of acceleration effects. As we reported in the moderator analysis section, in most of the primary studies, specific information about the ethnicity of the treatment samples, the SES of the study subjects, the school type and environment, or the gender ratio of the participants was missing. As a result, questions still remain as to the effects of acceleration on high-ability learners in terms of school context, student SES, ethnicity, and gender distribution.

These gaps in the research may be considered as a possible explanation for one of the conclusions in the Templeton National Report on Acceleration (Colangelo et al., 2004), which stated that throughout the nation, the effects of acceleration remained vague to educational administrators, teachers, and parents, although considerable evidence regarding the benefits of acceleration had been documented in the literature. Furthermore, this also suggests that acceleration research, like educational research overall, needs to be turned in a direction in which it is better linked to the education system practically.

A new research area suggested by this meta-analysis is the investigation of the impact of acceleration on high-ability learners during their transition from high school to college. Among the 38 primary studies from 1984 to 2008 that were included in this meta-analysis, no study investigated the effects of acceleration on high-ability learners in both high school and college. Researchers might want to explore this important transition. One can expect that these research efforts would be in line with the current increasing attention on high school to college transitions in the United States. Accompanying this interest in the high school to college pathway, more options are opening up for high-ability learners. For example, they can join an early-college high school, take AP classes, choose to enter college early, skip high school completely, or take dual-enrollment classes in their high school. As a result, more and more high-ability learners will be experiencing a nontraditional pathway between high school and college. They may have different experiences in high school, during the transition, or even after the transition. New research on the effects of acceleration during these transition periods is needed.

The criterion of requiring appropriate comparison groups played a significant role in determining which studies were included in this meta-analysis. A number of studies that met all the other criteria but failed to meet this one were excluded from this meta-analysis. A comparison of this meta-analysis with another current best-evidence synthesis of research on academic acceleration, conducted by Rogers, Young, and Lonergan (2008), revealed that a set of studies that were included in the latter were not qualified for inclusion in this meta-analysis because there were no appropriate comparison groups used and statistical controls were not applied. Therefore, researchers should make sure that future research is conducted with appropriate comparison groups or with necessary statistical control procedures.

Implications for Practice

Students and parents may find this meta-analysis helpful in their decision making about acceleration. As the end users of acceleration interventions, students' and parents' biggest concerns about acceleration are related to its short-term and long-term influence on students' academic achievement and social–emotional development. Often, through their personal experience, students and parents can understand and appreciate the short-term benefits of acceleration. However, they remain unsure or unconvinced about the long-term impact of their acceleration decisions. It is always the case that the evidence of the long-term effects of any educational intervention would have to rely on comprehensive practices across different contexts, for various subjects, and over a long period of time. In other words, the evidence eventually needs to be based on longitudinal/retrospective research studies. However, these types of studies were scarce when the prior meta-analytic studies were conducted. In the past two decades, more and more longitudinal studies were conducted. As mentioned previously, half of the studies (19 out of 38) included in this meta-analysis were longitudinal/retrospective studies. In other words, the findings of this meta-analysis are supported by a number of longitudinal studies that provided important and valuable information about the long-term effects of acceleration. Armed with this synthesized new information, students and parents will become more confident when they make decisions about acceleration.

A message from this meta-analysis for educators is that acceleration can be a valuable bridge to help them build and maintain optimal cooperation between the P-12 and postsecondary education systems. University-based acceleration programs played a very important role in implementing the acceleration practices in this study. The study results showed that, overall, students viewed their experiences in the university-based acceleration programs, such as early college entrance, quite positively; they believed that they benefited greatly in academic and social–emotional development. These positive outcomes will motivate educators in P-12 systems to devote more efforts to acceleration practices and to work more closely with universities to ensure appropriate acceleration opportunities during the transition to college. Similarly, the results of this study suggest that university-based acceleration programs need the support and cooperation of educators in the P-12 system, to recruit more qualified high-ability learners and achieve further success. One can expect that such a supply-and-demand relationship, based on successful acceleration programs, would benefit both P-12 and higher education greatly.

Conclusions

With the aim of updating previous meta-analytic studies and comprehensively synthesizing the current research findings regarding the effects of acceleration, this meta-analysis contributes to the field of gifted education in the following three ways. First, the findings from this meta-analysis generally confirm the positive influence of acceleration on high-ability learners, in terms of academic achievement and social–emotional development. Second, this meta-analysis suggests that researchers should particularly pay attention to ensuring that future research will be conducted with appropriate comparison groups and that more research is needed to investigate the impact of acceleration on high-ability learners during the transition from high school to college. Third, newly synthesized information based on a number of longitudinal studies about the long-term effects of acceleration was derived from this meta-analysis. In summary, while supporting the generally positive effects of acceleration found in previous research, this meta-analysis resulted in new information about acceleration that can benefit researchers, students, parents, educators, and policy makers.

Acknowledgement

We thank Marcia Gentry and Rebecca Mann for their insightful feedback to early versions of this study. We thank Harris Cooper, Yan Ping Xin, William Asher, Oliver Wendt, and Amy Dent for their invaluable meta-analysis expertise and guidance. We also thank Dake Zhang for her assistance with checking coding reliability. We are especially grateful to John Steenbergen, who provided tremendous help with proofreading and review over the course of this study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the authorship and/or publication of this article.

Funding

This research was partially supported by a graduate fellowship awarded to the first author by the College of Education at Purdue University; it was also partially supported by a postdoctoral fellowship to the first author by the Department of Psychology & Neurosciences at Duke University.

References

*Studies marked with an asterisk were included in the meta-analysis. The complete list of the 38 studies included in the meta-analysis can be obtained from the first author on request.

Ahadi, S., & Diener, E. (1989). Multiple determinants and effect size. *Journal of Personality and Social Psychology*, *56*, 398-406.

Aiken, L. R. (2004). *Psychological testing and assessment*. New York, NY: Allyn & Bacon.

*Ambruster, R. L. (1995). *The effectiveness of an early entrance screening instrument as an accurate predictor of success in elementary grades* (Unpublished doctoral dissertation). Southern Illinois University at Edwardsville.

Asher, J. W. (1986). Conducting research with meta-analysis: A new direction for gifted education. *Gifted Child Quarterly*, *30*, 7-10.

Asher, J. W. (2003). Meta-analysis and gifted education. *Journal for the Education of the Gifted*, *27*, 7-19.

Berliner, D. C. (2002). Educational research: The hardest science of all. *Educational Researcher*, *31*, 18-20.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2006). Comprehensive Meta-Analysis (Version 2.2.027) [Computer software]. Englewood, NJ: Biostat.

Borman, G. D., & Grigg, J. A. (2009). Visual and narrative interpretation. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 497-519.). New York, NY: Russell Sage Foundation.

*Brody, L., Assouline, S. G., & Stanley, J. C. (1990). Five years of early entrants: Predicting successful achievement in college. *Gifted Child Quarterly*, *34*, 138-142.

*Callahan, C., & Smith, R. M. (1990). Keller's personalized system of instruction in a junior high gifted program. *Roeper Review*, *XIII*, 39-44.

*Chilton, C. M. (2001). *The math achievement, experiences, and attitudes of gifted and promising math students* (Unpublished doctoral dissertation). Arizona State University, Tempe.

Colangelo, N., Assouline, S. G., & Gross, M. U. M. (2004). *A nation deceived*: *How schools hold back America's brightest students* (Vol. II, pp. 13-22). Iowa City, IA: Connie Belin and Jacqueline N. Blank International Center for Gifted Education and Talent Development.

Cooper, H. (2007). *Evaluating and interpreting research syntheses in adult learning and literacy* (Harvard Graduate School of Education, National Center for the Study of Adult Learning and Literacy [NCSALL] Occasional Paper). Retrieved from http://www.ncsall.net/fileadmin/resources/research/op_research_syntheses.pdf

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage Foundation.

Cumming, G., & Finch, S. (2001). A primer on the understanding, use and calculation of confidence intervals that are based on central and noncentral distributions. *Educational and Psychological Measurement*, *61*, 532-575.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Duvall, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455-463.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*, 107-128.

Hedges, L., Shymansky, J., & Woodworth, G. (1989). *A practical guide to modern methods of meta-analysis*. Washington, DC: National Science Teachers Association.

Higgins, J. P. T., & Green, S. (Eds.). (2006). *Cochrane handbook for systematic reviews of interventions*. Chichester, England: Wiley.

*Hsu, L. (2003). Measuring the effectiveness of summer intensive physics courses for gifted students: A pilot study and agenda for research. *Gifted Child Quarterly*, *47*, 212-218.

*Janos, P. M. (1987). A fifty-year follow-up of Terman's youngest college students and IQ matched agemates. *Gifted Child Quarterly*, *31*, 55-58.

*Janos, P. M., & Robinson, N. M. (1985). The performance of students in a program of radical acceleration at the university level. *Gifted Child Quarterly*, *29*, 175-179.

Kent, S. D. (1992). The effects of acceleration on the social and emotional development of gifted elementary students: A meta-analysis. *Dissertation Abstracts International*, *54*(2-A).

Kulik, J. A. (2004). Meta-analytic studies of acceleration. In N. Colangelo, S. G. Assouline, & M. U. M. Gross (Eds.), *A nation deceived*: *How schools hold back America's brightest students* (Vol. II, pp. 13-22). Iowa City, IA: Connie Belin and Jacqueline N. Blank International Center for Gifted Education and Talent Development.

Kulik, J. A., & Kulik, C. C. (1984). Effects of accelerated instruction on students. *Review of Educational Research*, *54*, 409-425.

Lauer, J., & Asher, J. W. (1988). *Composition research: Empirical designs*. New York, NY: Oxford University Press.

*Moon, T. R., & Callahan, C. M. (2001). Curricular modifications, family outreach, and a mentoring program: Impacts on achievement and gifted identification in high-risk primary students. *Journal for the Education of the Gifted*, *24*, 305-321.

National Work Group on Acceleration. (2009). *Guidelines for developing an academic acceleration policy*. Iowa City: Institute for Research and Policy on Acceleration, at the University of Iowa's Belin-Blank Center for Gifted Education and Talent Development.

Neihart, M. (2007). The socioaffective impact of acceleration and ability grouping: Recommendations for best practice. *Gifted Child Quarterly*, *51*, 330-341.

Pressey, S. L. (1949). *Educational acceleration: Appraisals and basic problems*. Columbus: Ohio State University.

Robinson, N. M. (2004). Effects of academic acceleration on the social-emotional status of gifted students. In N. Colangelo, S. G. Assouline, & M. U. M. Gross (Eds.), *A nation deceived*: *How schools hold back America's brightest students* (Vol. II, pp.59-67). Iowa City, IA: Connie Belin and Jacqueline N. Blank International Center for Gifted Education and Talent Development.

*Robinson, N. M., & Janos, P. M. (1986). Psychological adjustment in a college-level program of marked academic acceleration. *Journal of Youth and Adolescence*, *15*, 51-60.

Rogers, K. B. (1991). A best-evidence synthesis of research on types of accelerative programs for gifted students. *Dissertation Abstracts International*, *52*, 796-A. (UMI No. AAT91-22206)

Rogers, K. B. (2004). The academic effects of acceleration. In N. Colangelo, S. G. Assouline, & M. U. M. Gross (Eds.), *A nation deceived*: *How schools hold back America's brightest students* (Vol. II, pp. 47-57). Iowa City, IA: Connie Belin and Jacqueline N. Blank International Center for Gifted Education and Talent Development.

Rogers, K. B., Young, M., &Lonergan, R. (2008). *A best-evidence synthesis of the research on academic acceleration 1990-present*. Paper presented at the Ninth Biennial Henry B. & Jocelyn Wallace National Research Symposium on Talent Development, Iowa City, IA.

Southern, W. T., & Jones, E. D. (Eds.). (1991). *The academic acceleration of gifted children*. New York, NY: Teachers College Press.

*Stamps, L. (2004). The effectiveness of curriculum compacting in first grade classrooms. *Roeper Review*, *27*, 31-41.

Strube, M. J. (1991). Multiple determinants and effect size: A more general method of discourse. *Journal of Personality and Social Psychology*, *61*, 1024-1027.

*Swiatek, M. A., & Benbow, C. P. (1991). Ten-year longitudinal follow-up of ability-matched accelerated and unaccelerated gifted students. *Journal of Educational Psychology*, *83*, 528-538.

Tellegen, A., Ben-Porath, Y. S., McNulty, J. L., Arbisi, P. A., Graham, J. R., & Kaemmer, B. (2003). *The MMPI-2 Restructured Clinical Scales: Development, validation, and interpretation*. Minneapolis: University of Minnesota Press.

Thompson, B. (2002). What future quantitative social science research could look like: Confidence intervals for effect sizes. *Educational Researcher*, *31*, 24-31.

Thompson, B. (2006). Research synthesis: Effect sizes. In J. Green, G. Camilli, & P. B. Elmore (Eds.), *Complementary methods for research in education* (pp.583-603). Washington, DC: American Educational Research Association.

U.S. Department of Education. (1993). *National excellence: A case for developing America's talent*. Washington, DC: Government Printing Office.

Vaughn, V. L., Feldhusen, J., & Asher, J. W. (1991). Meta-analyses and review of research on pull-out programs in gifted education. *Gifted Child Quarterly*, *35*, 92-98.

*Washington, J. (1999). *Achievement and social-emotional outcomes of a multigraded, accelerated mathematics program* (Unpublished doctoral dissertation). The Johns Hopkins University. (AAT 9964225)

What Works Clearinghouse. (2007, April 16). *Technical details of WWC-conducted computations*. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/conducted_computations.pdf

*Wells, R., Lohman, D., &Marron, M. (2008). *What factors are associated with grade acceleration? An analysis and comparison of two U. S. databases*. (Note: This study was obtained from the authors in 2008 when it was in the process of preparation for publication.)

Bios

Saiying Steenbergen-Hu is a postdoctoral fellow in the Department of Psychology and Neuroscience at Duke University. She earned a PhD degree in educational psychology from Purdue University. Her research interests include meta-analysis, research evidence synthesis, research methods, research design, use of research for policy and practices, accelerated learning, motivation, teacher education, and classroom communication.

Sidney M. Moon is Professor of Gifted, Creative, and Talented Studies and Associate Dean for Learning and Engagement in the College of Education at Purdue University. She has been involved in the field of gifted, creative, and talented studies for almost 30 years. In that time, she has contributed more than 75 books, articles, and chapters to the field. Her most recent book is *The Handbook of Secondary Gifted Education*. She is active in the National Association for Gifted Children (NAGC), where she has served as Chair of the Research and Evaluation Division, a member of the Board of Directors, and Chair of the Bylaws Committee. Currently, she is association editor for NAGC. Her research interests include talent development in the STEM disciplines (science, technology, engineering, and mathematics), gifted program evaluation, and social/emotional development.