When the Big Fish Turns Small: Effects of Participating in Gifted Summer Programs on Academic Self-Concepts

David Yun Dai[1], Anne N. Rinn[2], and Xiaoyuan Tan[3]

Abstract

The purposes of this study were to (a) examine the presence and prevalence of the big-fish-little-pond effect (BFLPE) in summer programs for the gifted, (b) identify group and individual difference variables that help predict those who are more susceptible to the BFLPE, and (c) put the possible BFLPE on academic self-concept in a larger context of self-concept stability and change during adolescence. Longitudinal data were gathered from adolescents participating in a summer program for the gifted over a 3-year period. The results indicate no prevalent patterns of declines in academic self-concepts after participating in summer programs, though suspected cases of BFLPE can be identified, and there was evidence pointing to the moderation of the BFLPE by gender and self-esteem. Longitudinal patterns of self-concept stability and change also show no consistent pattern of long-term effects on self-concept. Implications of these findings are discussed that highlight developmental, social, and individual conditions under which the BFLPE may exert itself and conditions under which it may be mitigated.

Keywords

self-concept, gifted programs, big-fish-little-pond effect

In a study titled "The Effects of Gifted and Talented Programs on Academic Self-Concept: The Big Fish Strikes Again," Marsh, Chessor, Craven, and Roche (1995)

[1]University at Albany, State University of New York, USA
[2]University of North Texas, Denton, USA
[3]Department of Education, State of Arizona, Phoenix, USA

Corresponding Author:
David Yun Dai, ED233 Educational Psychology and Methodology, University at Albany, State University of New York, Albany, NY 12222, USA.
Email: ydai@albany.edu

launched what, to our knowledge, was the first study of the big-fish-little-pond effect (BFLPE) as applied to gifted education. Simply put, the BFLPE predicts that equally able students have lower academic self-concepts when attending schools or participating in programs wherein the average ability levels of peers are high, and higher academic self-concepts when the average ability levels of peers are low (Marsh et al., 2008). The underlying assumption is that individuals typically use their local peer group (either school or class) as a frame of reference when gauging their own academic competence in general or in specific school subjects. Although Marsh and colleagues (2008), as well as some other researchers, did find the BFLPE from attending gifted programs (e.g., Zeidner & Schleyer, 1998), the findings in general are mixed when BFLPE is applied to gifted education settings (see Dai & Rinn, 2008, for a review). Although the BFLPE model is not specific to gifted programs, facets of the BFLPE have been examined with gifted and high-ability students ranging in grade from the early elementary years (Tymms, 2001) to the college years (Rinn, 2007), and the practical implications are obvious and have already produced repercussions in the gifted education community (e.g., Plucker et al., 2004). If attending selective schools or participating in gifted programs indeed has negative effects on one's academic self-concept, then it behooves us to closely examine this putative effect for educational or intervention purposes, as academic self-concept plays a role in the development of one's academic achievement (Marsh & Yeung, 1997) and aspirations (Rinn, 2007), among other important academic indicators and outcomes.

The purposes of this study are to (a) examine the longitudinal presence and prevalence of the BFLPE in summer programs for the gifted, (b) identify variables that help predict those who are more susceptible to the BFLPE, and (c) put possible program effects on academic self-concept in a larger context of self-concept stability and change during adolescence.

Reframing the BFLPE

Dai (2004; see also Dai & Rinn, 2008) argued that in order for the BFLPE model to have a real practical impact, it has to specify "Where" (under what condition), "For Whom" (persons with specific characteristics), and "When" (at what juncture of one's life or development) the BFLPE is likely to occur. The research design adopted by Marsh and colleagues (2008) uses the effect of general group-average ability (at the school or class level) on academic self-concept as an estimation of the BFLPE, without clear indication of how frequently it occurs or how many people experience such a decline. In the context of gifted programs, we suggest a simpler but more direct approach for estimation of the BFLPE, similar to the one used by Marsh et al. (1995); that is, establish a baseline academic self-concept at the beginning of a program and then assess positive and negative changes at the end of such a program. The rationale is the following: According to Marsh et al., (2008) social comparison and ensuing changes in self-concept are likely when individuals are exposed to new peer groups and become uncertain about their abilities, or when the new realities compel them to

take a second look, so to speak, at their own competencies. Participation in a gifted program might result in such a closer examination. A more recent study (Makel, Lee, Olszewski-Kulibius, & Putallaz, 2012) also used this within-subject rather than between-subject design.

The Question of "Where"

A gifted summer program, in particular, is such an event that potentially has an impact on how academically advanced adolescents see themselves. We intended the current design to address the "where" question. By following adolescents outside of their normal school environment and in the context of a summer program for gifted students of similar high ability, we hoped to spur the exposure of a new peer group of equally able peers and the feelings and changes in self-concept that might result.

Although numerous researchers have provided support for the BFLPE in various samples and across different settings, some researchers have shown an increase in students' self-concepts after participating in programs for the gifted, and others have shown no changes in self-concept upon participating in programs for the gifted. Kolloff and Moore (1989) found an increase in general self-concept among gifted students who participated in residential, summer programs for the gifted. Using a sample of 140 gifted adolescents, Cunningham and Rinn (2007) found adolescents experienced an increase in general self-concept (i.e., self-esteem) over the course of the summer program, but did not experience a change in academic self-concept, which is inconsistent with the BFLPE. More recently, using a large sample, Makel et al. (2012) found no decline of academic self-concept of those identified through Talent Searches after participating in a summer program and enhanced nonacademic self-concepts. They argued that supplementary programs like summer programs may operate in a different way than replacement programs; the 2- or 3-week experience may produce sufficient positive effects without detrimental effects on academic self-concept. It is also possible that in such short-term programs, intellectual stimulation is more salient than social-evaluative pressure to the effect of overriding any possible BFLPE. Availability and accessibility of peer comparison information may also determine whether forced social comparison, which is one of the conditions for BFLPE, is present (Huguet et al., 2009).

The Question of "Who"

To address the "who" question, a research design needs to consider a host of possible moderators and mediators of the BFLPE. Motivation for identifying individuals who are more vulnerable to the BFLPE is based on an assumption that individuals respond to socially competitive academic environments differently, and characteristics of those who are more vulnerable to the BFLPE can be identified. Seaton, Marsh, and Craven (2010), in direct response to this criticism (see Dai & Rinn, 2008), examined moderators in the BFLPE and found that the BFLPE was

moderated by a variety of individual characteristics. In particular, the BFLPE was more pronounced for those students who were more intelligent, those who were highly anxious, those who use memorization as a learning strategy (rather than elaboration, or deep thinking and concept integration, strategies), and those who sanction a cooperative orientation (rather than a competitive orientation). Furthermore, previous research suggests a possible gender moderation whereby females are more likely to lower their self-concept under social-evaluative pressure (Dweck, 1999; Eccles, Adler, & Meece, 1984).

There is also evidence that self-esteem moderates how social comparison is made; that is, individuals with low self-esteem tend to engage in upward social comparison, a strategy that likely depresses their own self-concept (Taylor, Wayment, & Carrillo, 1996). The social comparison literature also suggests individual differences in one's disposition to engage in social comparison. Using a social comparison measure, Gibbons and Buunk (1999) were able to identify individuals who were more prone to make social comparisons than others. There are, of course, other contextual and individual factors that potentially mitigate the BFLPE. In a recent study, Huguet et al. (2009) attempted to separate two social comparison processes and their respective effects, one forced by the immediate social environment (i.e., salient social comparison information) and the other deliberately engaged or chosen by some individuals, which can involve either upward or downward social comparison. The results showed the coexistence of both social comparison processes, thus reconciling the different findings by the BFLPE research and the classic social comparison research that defines social comparison as chosen by individuals (Festinger, 1954). Taken together, the BFLPE research findings suggest that, to have a full understanding of the BFLPE, contextual, individual, and developmental factors need to be considered in an integrated manner.

The Question of "When"

To address the "when" question, a longitudinal design with special attention to developmental timing, phases, and stages of a population in question is necessary. The program effects on academic self-concepts highlight the situational influences, whereas academic self-concept always has a distinct development component; that is, self-concepts tend to be more differentiated as children are cognitively more mature and socially exposed to more social comparative information (Byrne & Shavelson, 1996; Nicholls, 1984; Ruble, Grosovsky, Frey, & Cohen, 1992). Marsh (1987) noted that the BFLPE might be smaller for older students, as they

> typically have some basis for the assessment of their own academic skills that is independent of the performances of their classmates, and they often know how the average ability level of their classmates compares with some broader frame of reference. (p. 282)

Once self-concepts are crystallized or established, they have continuity and stability, and can only be disturbed by compelling information to the contrary. Whether temporary

reevaluation in one's own competencies leads to enduring changes and negative effects is an empirical issue. For example, although Marsh, Trautwein, Lüdtke, Baumert, and Köller's (2007) research shows the BFLPE to be enduring in nature, there is research evidence to suggest that depressed self-concept during a gifted program can bounce back to baseline after the program (Gibbons, Benbow, & Gerrard, 1994) or even increase after the program. Moon, Feldhusen, and Dillon (1994) noted that the short-term effects of a gifted program on the self-concepts of gifted students may be negative or nonexistent, but the long-term effects may actually be positive. In their study of the long-term effects of an enrichment program for gifted elementary school students, Moon et al. found that the long-term benefits of participating in a gifted program are mostly positive, including an increased self-concept, increased motivation, an increase in basic thinking skills, and an increase in autonomous learning. However, prior research examining the short-term effects of the same program on the same sample of students indicated the program did not have a significant effect on self-concept. Other researchers have found similar results in their longitudinal investigations of the effects of programs for the gifted. For example, a 15-year longitudinal study concerning the impact of a program for the gifted indicated a largely positive impact on the lives and attitudes of the students involved (Humes & Campbell, 1980). Similarly, a 4-year longitudinal study showed participants experienced an increase in self-esteem after participating in a summer program for the gifted (Thomas, 1989). The positive long-term effects seen in these studies seemingly indicate the BFLPE may, at most, be temporary.

The above review suggests that, to fully understand the nature and impact of BFLPE, contextual (where), individual (who), and developmental (when) factors should be considered in an integrated manner. This study represents a preliminary attempt in this direction. In this study, we addressed three research questions:

> *Research Question 1*: What is the extent to which academic self-concepts can be per-
> turbed by participating in a gifted summer program (the question of "where")?
> *Research Question 2*: Do individual characteristics such as age, gender, self-
> esteem, and the inclination to make social comparison have effects on
> changes in academic self-concepts (including general academic, verbal, and
> math self-concepts; the question of "who")?
> *Research Question 3*: How stable are academic self-concepts over time during
> adolescence (11-16 years of age), and can perturbations be observed in the
> context of alternating between supplementary summer programs and mixed
> regular schools (the question of "when")?

Method

Participants

Participants were recruited from one summer gifted program that has been in operation for more than 20 years at a comprehensive university in the southern United States. This summer program is a 3-week residential program for gifted students entering the

8th, 9th, 10th, or 11th grades the following school year. To qualify for participation in this summer program, students must have been eligible to attend talent search summer programs (e.g., through the Duke Talent Identification Program)[1] within the past 4 years. It involves 6 hr of class and 1 hr of study hall per day, 5 days a week, for 3 weeks. The students have a variety of courses from which to choose (e.g., acting, ecology, geography, science), and they enroll in only one course. The students also engage in various social activities after class each day and on weekends.

Data were gathered at the beginning (pre) and end (post) of the summer program for three consecutive years (2005, 2006, and 2007) from approximately 291 younger adolescents between 11 and 16 years of age. A total of 152 students who had never participated in any gifted programs before participated in this program for at least 1 year. Among those 152 students, 53 participated in this program for at least two consecutive years (2005-2006 or 2006-2007), and only 19 students participated in this program for the entire 3 years from 2005 to 2007.

Instruments and Materials

Participants were given a demographic questionnaire to identify gender and age, among other information. Other data were gathered from participants' applications for the summer programs, including ethnic background and grade level.

Self-Concept. The Self-Description Questionnaire–II (SDQ-II) was designed to measure the self-concepts of young adolescents and is theoretically based on the notion that self-concept is multidimensional and hierarchically structured (Marsh, 1990; Shavelson, Hubner, & Stanton, 1976). The SDQ-II measures self-concept in the following areas: mathematics, verbal, physical abilities, physical appearance, same-sex peer relations, opposite-sex peer relations, parent relations, emotional stability, honesty-trustworthiness, general academic, and general self (i.e., self-esteem). A 6-point Likert-type scale is used to measure self-concept in these areas (1 = *false* to 6 = *true*). Extensive support for the reliability and validity of the SDQ-II has been reported in other research (see Gilman, Laughlin, & Huebner, 1999; Plucker, Taylor, Callahan, & Tomchin, 1997).

For the purposes of this study, only the Mathematics Self-Concept, Verbal Self-Concept, General Academic Self-Concept, and General Self-Esteem subscales were used. The Mathematics subscale measures ability, enjoyment, and interest in mathematics and reasoning (e.g., "I do badly in tests of mathematics"; Marsh, 1990, p. 5). From the normative sample, internal consistency was reported as 0.90 and factor loadings range from .72 to .80. The Verbal subscale measures ability, enjoyment, and interest in English and reading (e.g., "I learn things quickly in English classes"; Marsh, 1990, p. 5). Internal consistency was reported as .86 and factor loadings range from .53 to .75. The General Academic Self-Concept subscale measures one's interests and abilities in schoolwork (e.g., "I learn things quickly in most school subjects"; Marsh, 1990, p. 6). From the normative sample, internal consistency was reported as .87 and

factor loadings range from .48 to .64. The General Self-Esteem subscale measures one's feeling of self-worth, self-confidence, and self-satisfaction (e.g., "If I really try I can do almost anything I want to"; Marsh, 1990, p. 6). From the normative sample, internal consistency was reported as .88 and factor loadings range from .49 to .64.

Social Comparison. Social comparison was assessed using the Iowa–Netherlands Comparison Orientation Measure (INCOM; Gibbons & Buunk, 1999). The INCOM is a self-report measure that was developed to gauge an individual's level of social comparison. Participants are asked to respond to 11 statements on a 5-point Likert-type scale (A = *I disagree strongly* to E = *I agree strongly*). Thus, higher scores indicate higher levels of social comparison and lower scores indicate lower levels of social comparison. Sample items include "I always pay a lot of attention to how I do things compared to how others do things," and "I am not the type of person who compares often with others." From the normative sample, internal consistency was reported as .83 (Gibbons & Buunk, 1999). Validity has been assessed by significant correlations between scores on the INCOM and various other social comparison information measures, including the Attention to Social Comparison Information Scale (Lennox & Wolfe, 1984; $r = .47$ to $.66$, $p < .001$).

The INCOM was administered both pre and post the summer program to gauge fluctuations in students' tendencies to make comparisons. The factor structure of this scale was tested using the responses collected from 152 students who were the first-time participants for the summer program in either Year 1, Year 2, or Year 3.

In the first step, confirmatory factor analyses were conducted using SAS procedure PROC CALIS. Results indicated that the two-factor solution did not fit the data very well (goodness of fit indices [GFI] = 0.69; adjusted GFI = 0.82; standardized root mean square residual [SRMSR] = 0.079; and root mean square error of approximation [RMSEA] = 0.089), and the two-factor solution did not fit the data significantly better than the one-factor solution. The chi-square value dropped only slightly from 95.4 in the single-factor solution to 94.1 in the two-factor solution, which was not significant at $p < .05$ with one degree of freedom.

In the second step, exploratory factor analyses were conducted to display the factor structure of INCOM measures in the 152 student sample. Image factor analysis with varimax rotation was used. The scree plot for image factor analysis in Figure 1 suggested the existence of two notable factors. However, the rotated factor loading matrix (see Table 1) did not support the same two-factor structure revealed in Gibbons and Buunk (1999). Results showed that Items 1, 3, 4, 6, 7, 10, and 11 had higher loadings on Factor 1, Items 2, 5, 8, and 9 had higher loading on Factor 2, and the two factors were highly correlated ($r = .63$). The inconsistent factor structure of the INCOM was probably caused by different characteristics of the two samples. The current sample consisted of younger adolescents (aged between 11 and 16), whereas the samples from Gibbons and Buunk consisted of older adolescents (*M* age = 17) and college students. Given the fact that the two factors were highly correlated and the two-factor solution was not significantly better than the one-factor solution, the 11 INCOM items were averaged to generate one single social-comparison composite score.
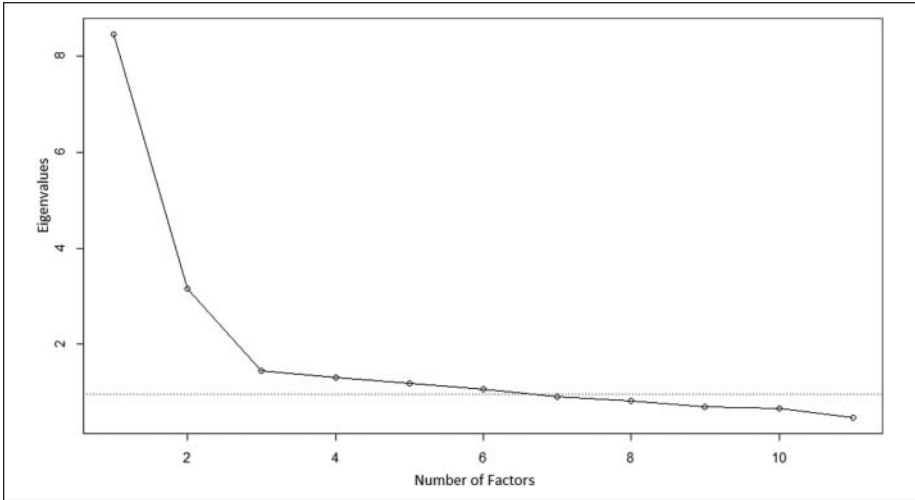
Figure 1. Scree plot for image factor analysis (*N* = 152).

Table 1. Image Factor Analysis After Varimax Rotation Constrained to Two Factors (*N* = 152).

Items	Factor 1	Factor 2	h2	Cpx
1	0.62	0.21	0.43	1.20
2	0.31	0.49	0.33	1.70
3	0.53	0.23	0.33	1.40
4	0.51	0.22	0.31	1.30
5	0.25	0.57	0.38	1.40
6	0.68	0.07	0.47	1.00
7	0.58	0.16	0.36	1.20
8	0.10	0.66	0.45	1.00
9	0.07	0.60	0.37	1.00
10	0.31	0.29	0.18	2.00
11	0.37	0.06	0.14	1.00
Cssm	0.19	0.15	0.34	1.29

Note: Reports the factor loadings, the communality estimates (h2), the complexity estimates (cpx), and the amount of variance explained (cssm).

Procedure

Parental consent was obtained prior to the start of the summer program. Adolescents who gave assent and whose parents gave consent were invited to take part in the study, but they were given the option to decline participation (none declined). Data were gathered in a single session during the 1st week of each summer program (A05, A06, and A07), and then again 2 days before the end of the program (B05, B06, and B07).

Table 2. Results of Dependent-Sample *t* test for General Academic Self-Concept, Math Self-Concept, and Verbal Self-Concept.

	n	Pre-mean	Post-mean	Mean difference [95% CI]	*t*(*p*)	Standardized effect size
General academic self-concept	147	5.47	5.46	−0.01 [−0.08, 0.07]	−0.21(<.83)	−0.02
Math self-concept	148	4.98	4.95	−0.03 [−0.10, 0.05]	−0.7(<.49)	−0.06
Verbal self-concept	148	5.01	5.04	0.03 [−0.06, 0.11]	0.55(<.58)	0.05

Note: CI = confidence interval.

Statistical Analysis

Analyses of the data were carried out in two stages. The first stage examined the short-term effects of the BFLPE on academic self-concepts just over a 3-week period in the summer program; while the second stage addressed the longitudinal presence and prevalence of the BFLPE over a 2-year period.

Results

In the first stage, the students who participated in the program for at least one summer in these 3 years and who filled out the self-concept and other questionnaires twice prior to and 2 days before the end of the program were included in the analyses. This approach is justified in that what is of interest for detecting the BFLPE is not in what year the participants attended the program, but what was the first time they participated. The resultant total sample size was 152, with 74% being White and 58% being male. The data of these 152 students from their participation in the program for the first time in any of the 3 years were compiled together, and a dependent-sample *t* test was used to assess changes in their self-concepts over the 3 weeks in this program.

Self-Concept Changes Prior to and at the End of the Program

As most of the BFLPE research was conducted at the level of general academic self-concept, we decided to start out at the general level and then move to more domain-specific self-concepts. Table 2 displays the average self-concepts before the program and after the program, the differences between the two averages, dependent-sample *t* statistics and their corresponding *p* values, and standardized effect sizes. The results show that, overall, there were almost no measurable changes in self-concept. Across the three types of self-concepts (academic, math, verbal), the pre and post mean differences were very close to zero, and their corresponding *t* statistics were nonsignificant with negligible effect sizes.

Each of the dependent-sample analyses was followed by a dependent-sample assessment plot (Pruzek & Helmreich, 2009). In Figures 2, 3, and 4, the preprogram
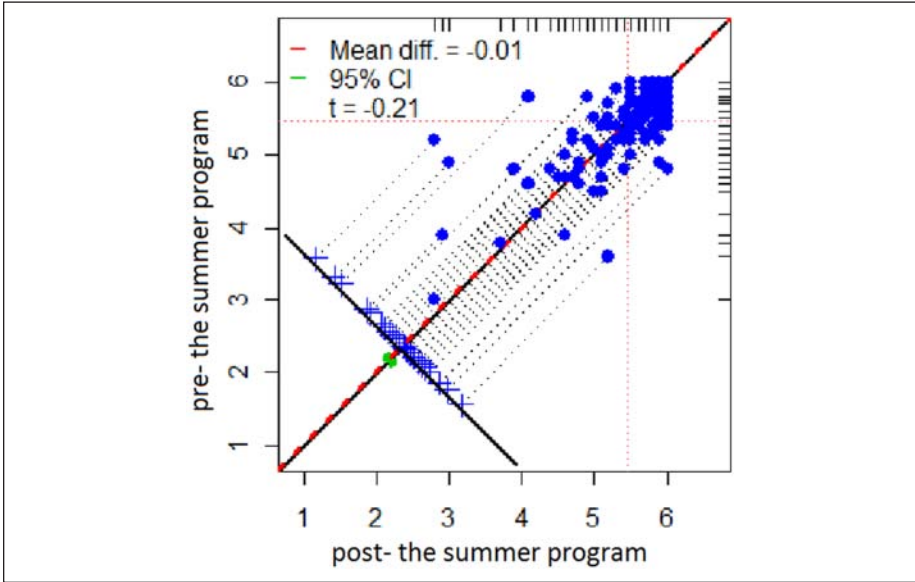
Figure 2. Effect of BFLPE on general academic self-concept.
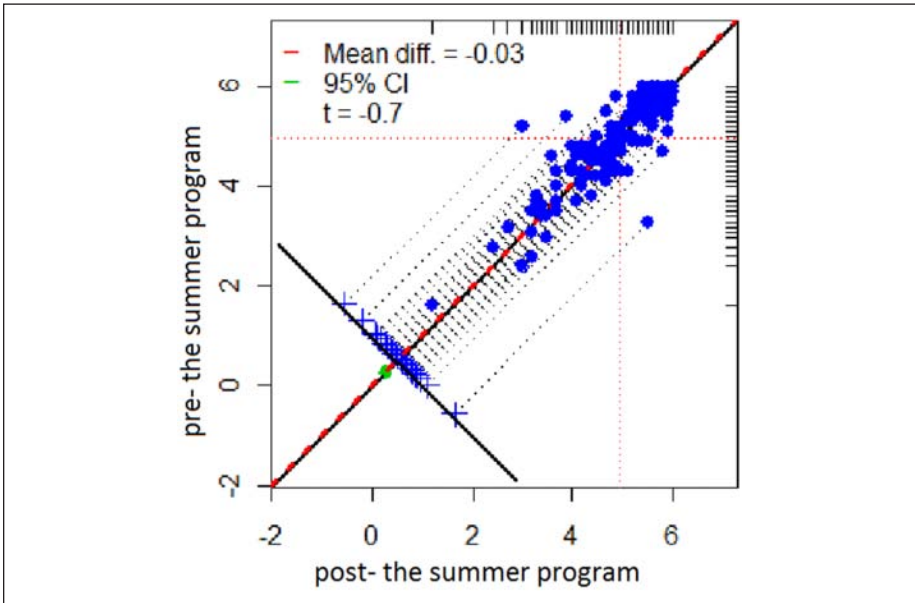Note: BFLPE = big-fish-little-pond effect; CI = confidence interval.

Figure 3. Effect of BFLPE on math self-concept.
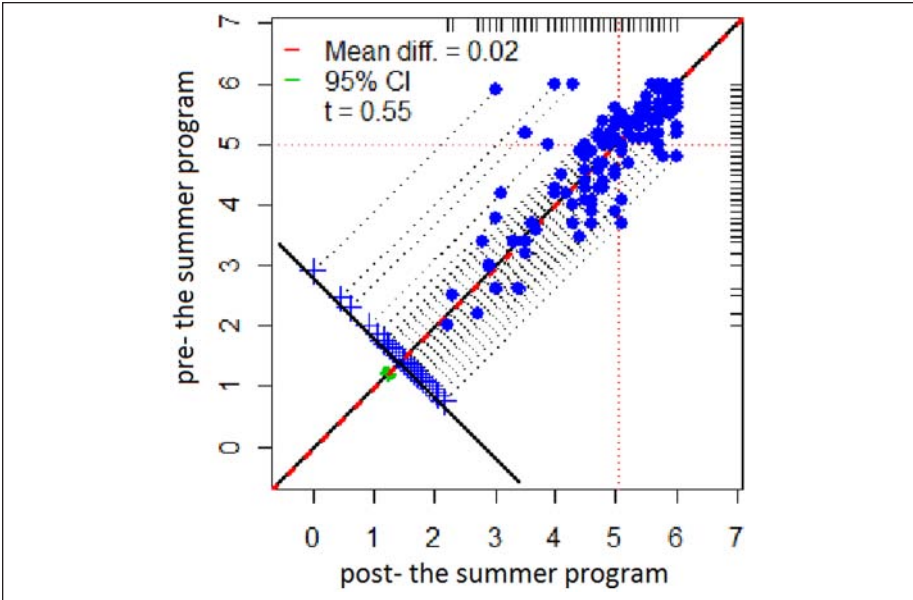Note: BFLPE = big-fish-little-pond effect; CI = confidence interval.

Figure 4. Effect of BFLPE on verbal self-concept.
Note: BFLPE = big-fish-little-pond effect; CI = confidence interval.

self-concept scores were displayed on the *y*-axis, while the postprogram scores were displayed on the *x*-axis. The heavy dashed line, parallel to the solid diagonal line, depicts the mean of differences. A heavy dashed line falling exactly on the diagonal line indicates no difference between pre- and post-measures of self-concept. The heavy line perpendicular to the identity, depicts a 95% confidence interval for the mean difference. A heavy line across the diagonal line indicates no significant differ-ence at *p* < .05 level. These graphs yield more detailed individual differences than numerical summaries. Diagnostically, although there were suspected cases (those dots below and far from the diagonal line on the right side) possibly implicating the BFLPE, more evident in their general academic self-concept and verbal self-concept, most cases show high stability. The data plots also suggest that initially low scores on aca-demic self-concept were more volatile and prone to fluctuating than higher scores (wider dispersion at the lower end), suggesting changes in self-concept are more likely for those who initially were uncertain about their academic abilities (score falling in the average or below average range on the 6-point Likert-type scale of the SDQ-II).

Moderation of Stability and Changes by Age, Gender, Type of Classes Attended, Self-Esteem, and Social Comparison

To address the second question regarding possible moderation of the BFLPE by indi-vidual differences (the question of "for whom"), contrast scores were generated for

Table 3. Results of Multiple Regression Analyses With Self-Concept Contrast Scores as Dependent Variables.

	b_0	b_{age}	b_{gender}	b_{class}	$b_{self-esteem}$	$b_{social\ comparison}$	R^2
General academic self-concept	0.75	−0.06	0.04	0.05	−0.03	0.02	0.02
Math self-concept	0.34	−0.05	−0.02	−0.10	0.05	0.04	0.04
Verbal self-concept	1.12	−0.07	−0.21*	0.01	0.04	−0.03	0.08

*$p < .05$.

each academic self-concept measure using postprogram scores minus preprogram scores, such that zero scores indicate no change, positive scores indicate increases, and negative scores indicate decreases, and the contrast scores were regressed on age, gender, class, students' initial self-esteem, and social comparison tendency before their participation in the program. The variable "Class" was formed as a dummy variable; those who predominantly take science, technology, engineering, and mathematics (STEM)–related classes were coded "1" and those who predominantly took language arts and social studies classes were coded "0."

Results from these multiple regression analyses, as displayed in Table 3, show little moderation of academic self-concepts by the five variables, probably because the contrast scores did not provide a sufficient range of variations (fluctuations) in the first place. Only the effect of gender on verbal self-concept were significant at $p < .05$ level, suggesting that females were more likely to experience a decrease in their verbal self-concept after participating in the gifted program. Although statistically not significant, those who took STEM-related classes instead of language arts/social studies classes appeared to show a decrease in their math self-concept. Because the pre–post difference or contrast did not produce sufficient variations, the moderation analyses lack the power to detect related effects.

Longitudinal Stability and Changes and Possible Perturbations of BFLPE

On the second stage of data analysis, long-term stability and changes in self-concept were examined, and the possible perturbations as a result of BFLPE were explored. Students who participated in the summer program for at least two consecutive years (2005-2006 or 2006-2007) and who provided responses to all four surveys (pre- and post-measures for each year) were included in this stage of analysis, resulting in a sample of 53 students. Among the 53 students, 94% were White, 64% were male, and 19 of them participated in the program for three consecutive years from 2005 to 2007.

Table 4 presents the means and confidence intervals of academic self-concepts, self-esteem, and social comparison at each of the six waves of data collections over the

Table 4. Counts of Students, Their Means, and 95% Confidence Intervals at the Six Waves of Data Collection for the Three Time-Varying Academic Self-Concept Outcomes and the Two Time-Varying Covariates: Self-Esteem and Social Comparison.

		Year 1		Year 2		Year 3	
		Pre	Post	Pre	Post	Pre	Post
General academic	Counts	53	49	51	44	19	17
	M [95% CIs]	5.52 [5.38, 5.66]	5.62 [5.5, 5.74]	5.58 [5.42, 5.74]	5.54 [5.36, 5.72]	5.59 [5.41, 5.77]	5.68 [5.52, 5.84]
Math	Counts	53	48	51	44	19	17
	M [95% CIs]	5.07 [4.82, 5.32]	5.09 [4.84, 5.34]	4.89 [4.64, 5.14]	4.96 [4.71, 5.21]	4.72 [4.31, 5.13]	4.89 [4.48, 5.3]
Verbal	Counts	53	49	51	44	19	17
	M [95% CIs]	5.06 [4.81, 5.31]	5.14 [4.89, 5.39]	5.12 [4.87, 5.37]	5.16 [4.87, 5.45]	5.49 [5.24, 5.74]	5.55 [5.33, 5.77]
Self-esteem	Counts	53	49	51	44	19	17
	M [95% CIs]	5.31 [5.15, 5.47]	5.44 [5.3, 5.58]	5.32 [5.08, 5.56]	5.38 [5.11, 5.65]	5.58 [5.36, 5.8]	5.56 [5.27, 5.85]
Social comparison	Counts	53	49	51	45	19	16
	M [95% CIs]	3.36 [3.2, 3.52]	3.27 [3.09, 3.45]	3.39 [3.19, 3.59]	3.4 [3.18, 3.62]	3.63 [3.32, 3.94]	3.55 [3.16, 3.94]

Note: CI = confidence interval.

3 years for these 53 students. Following the rationale of Gibbons et al. (1994), we would expect self-concepts to be depressed when the students were exposed to equally able peers during the program and self-concepts to bounce back when they returned back to their normal school environment. From Table 4, however, the mean math self-concepts, mean verbal self-concepts, and mean self-esteems of the first four waves displayed an opposite pattern, the postprogram means were slightly higher than the preprogram means in the same year while the preprogram means of the 2nd year were slightly lower than the postprogram means of the 1st year, indicating math self-concept, verbal self-concept, and self-esteem increased a little bit when the students were in the program and decreased a little bit when they went back to their normal schools. The tooth pattern displayed by the means of math self-concept, verbal self-concept, and self-esteem across the first four waves can also be observed when they were plotted in Figure 5. The means of general academic self-concept and social comparison did not show such a tooth pattern. As there were only 19 students who participated in the program for 3 years, the counts of students at Waves 5 and 6 were much less, and data collected at these two waves were limited in their generalizability.

To test long-term stability and changes of academic self-concepts in early adolescence, data for the 2005-2006 participants and for the 2006-2007 participants were merged so that the data represent Year 1 and Year 2 participation, rather than chronological order of participation. In other words, initially six waves of data were sorted and reorganized according to the availability of four waves of data into a 2-year sequence of four data points: Year 1 pre- and post-measures, and Year 2 pre- and post-measures.

The data were modeled using mixed-effects regression (MRM) via SAS procedure PROC MIXED. Besides the three academic self-concept measures, self-esteem and social comparison measured at the beginning and end of the program each year were used as time-varying covariates. Other covariates included in the models were gender, class (language arts–related vs. STEM-related), and a pre- or post-indicator. The moderations of age on self-concepts were ignored as it was highly correlated with time. Descriptions of the three self-concept measures and the five covariates are provided in Table 5.

Missing data problems are inherent in longitudinal studies. Likewise, there were missing data in the three self-concept measures as well as in the two time-varying covariates, self-esteem and social comparison. SAS PROC MIXED deletes cases with missing data, which can give rise to biased and misleading results (Little & Rubin, 1987). To handle missing data problems, we invoked a multivariate linear mixed-effects model (Schafer & Yucel, 2002) that was incorporated in R package "pan." This model can simultaneously produce multiple imputations of missing values in both the longitudinal measures and time-varying covariates. Ten imputed data sets were produced from this multiple imputation procedure.

Mixed-effects regression models were first run on two randomly selected imputed data sets with each self-concept measures as the outcome, one after the other sequentially. The basic model was the random intercept model, which was followed by random intercept and slope model. Both linear and quadratic time trends were tested. The
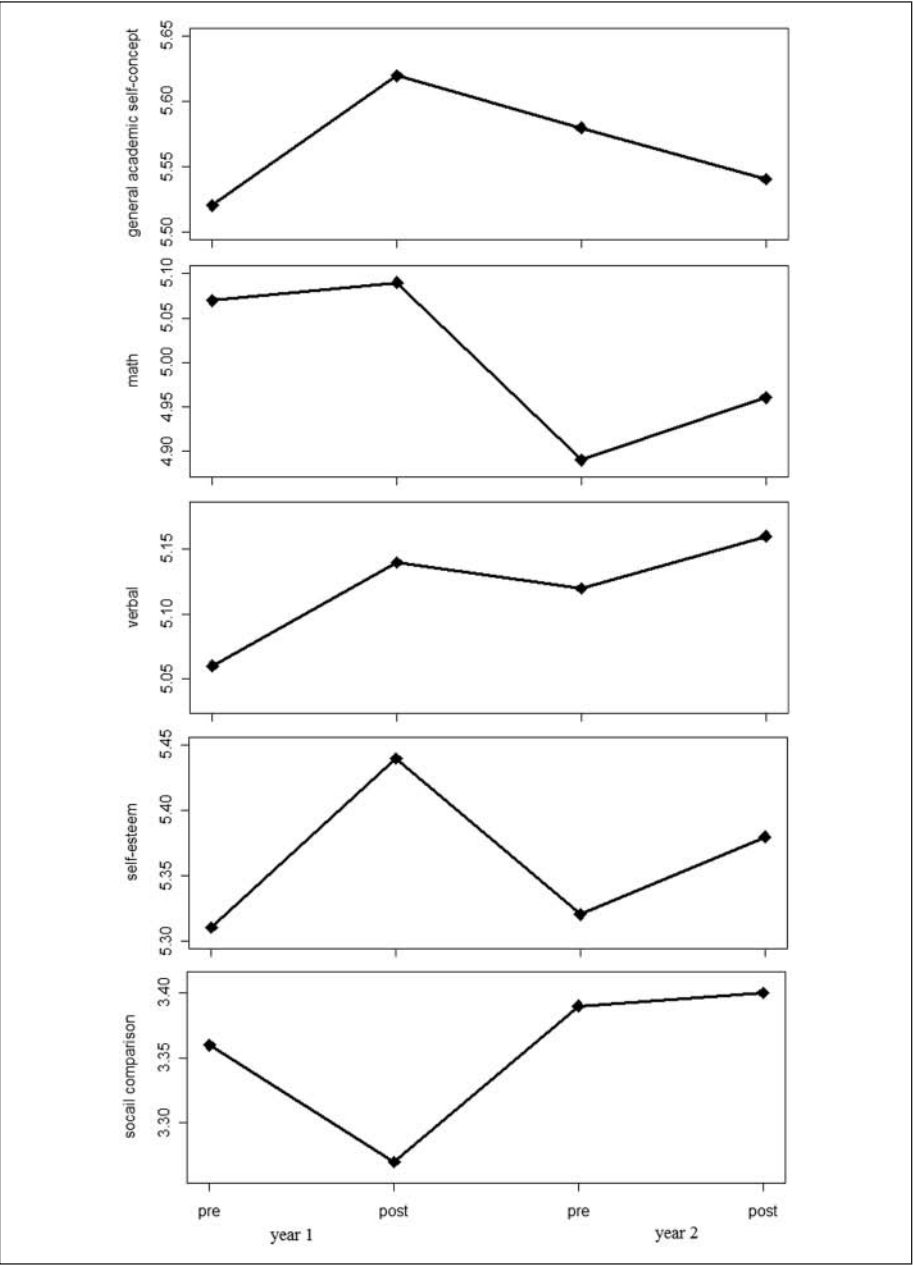
Figure 5. Longitudinal plotting of the means for the three academic self-concepts, self-esteem, and social comparison at the four waves of data collections in the first and the second years.

Table 5. Descriptions of the Three Academic Self-Concept Outcomes and the Covariates Involved in the Longitudinal Data Analyses.

Variable	Description	Category
General academic self-concept	Assessed by the SDQ-II general academic subtest at the beginning and the end of the summer program for 3 years	Time-varying outcome
Math self-concept	Assessed by the SDQ-II math subtest at the beginning and the end of the summer program for 3 years	Time-varying outcome
Verbal self-concept	Assessed by the SDQ-II verbal subtest at the beginning and the end of the summer program for 3 years	Time-varying outcome
Pre–post	Pre = 0, post = 1	Time-fixed covariate
Gender	Male = 1, female = 2	Time-fixed covariate
Class	Art-and-humanity related = 0, STEM related = 1	Time-varying covariate
Self-esteem	Assessed by the SDQ-II general self subtest at the beginning and the end of the summer program for 3 years	Time-varying covariate
Social comparison	Assessed by the INCOM at the beginning and the end of the summer program for 3 years	Time-varying covariate

Note: SDQ-II = Self-Description Questionnaire–II; STEM = science, technology, engineering, and mathematics; INCOM = Iowa–Netherlands Comparison Orientation Measure.

fit difference between any two models is distributed as a chi-square distribution, and a significant chi-square statistic indicates that the model was significantly improved by the additional terms added to the model. Through this model-selection process, the random linear trend model was determined to be the most efficient model that can account for change and stability over time.

We then ran the random linear model on the 10 imputed data sets. In addition to time effects, the full model also included, at the within-subject level, a pre- or post-indicator to account for the short-term effects during the program, and the interaction between pre- or post-indicator and time to investigate the possibility of differing short-term effects over time. The effects of other covariates—gender, class, self-esteem, and social comparison—were tested by adding their within-subject effects into the full model one after another. The results from the 10 imputed data sets were pooled together using SAS PROC MIANALYZE to get one single set of parameter estimates, and a 95% confidence interval was constructed for each estimate to account for imputation uncertainty.

The regression coefficients from these mixed-effects regression models and their 95% confidence intervals are presented in Table 6. Time effects were significant at $p <$ .05 level only on math self-concept. The math self-concept decreased as time went by,

Table 6. Regression Coefficients and Their 95% Confidence Intervals for the Mixed-Effects Regression Models.

Self-concept	Model	b_0	b_{time}	$b_{pre-post}$	$b_{time \times pre-post}$	$b_{covariate}$
General academic self-concept	Model 1	5.52	0.03 [−0.03, 0.09]	0.10 [−0.04, 0.23]	−0.04 [−0.11, 0.03]	
	Model 2	5.73 [5.38, 6.07]	0.03 [−0.03, 0.09]	0.10 [−0.04, 0.23]	−0.04 [−0.11, 0.03]	−0.15 [−0.39, 0.08]
	Model 3	5.55 [5.4, 5.7]	0.03 [−0.04, 0.09]	0.10 [−0.04, 0.24]	−0.04 [−0.11, 0.03]	−0.05 [−0.20, 0.09]
	Model 4	3.59 [3.13, 4.04]	0.03 [−0.02, 0.08]	0.06 [−0.07, 0.18]	−0.04 [−0.11, 0.03]	0.36* [0.28, 0.45]
	Model 5	6.06 [5.70, 6.42]	0.04 [−0.02, 0.10]	0.07 [−0.07, 0.20]	−0.03 [−0.11, 0.04]	−0.16 [−0.26, −0.06]
Math self-concept	Model 1	5.07	−0.09* [−0.17, −0.02]	0.02 [−0.13, 0.18]	0.04 [−0.04, 0.12]	
	Model 2	5.52 [4.84, 6.20]	−0.09* [−0.17, −0.02]	0.02 [−0.13, 0.18]	0.04 [−0.04, 0.12]	−0.33 [−0.80, 0.13]
	Model 3	5.03 [4.75, 5.30]	−0.09* [−0.17, −0.00]	0.01 [−0.15, 0.17]	0.04 [−0.04, 0.12]	0.07 [−0.14, 0.27]
	Model 4	4.54 [3.82, 5.26]	−0.09* [−0.17, −0.02]	0.01 [−0.15, 0.17]	0.04 [−0.04, 0.12]	0.10 [−0.03, 0.23]
	Model 5	5.14 [4.63, 5.66]	−0.09* [−0.17, −0.02]	0.02 [−0.14, 0.17]	0.04 [−0.04, 0.13]	−0.02 [−0.16, 0.11]
Verbal self-concept	Model 1	5.06	0.03 [−0.05, 0.11]	0.05 [−0.10, 0.21]	0.01 [−0.08, 0.09]	
	Model 2	4.97 [4.26, 5.67]	0.03 [−0.05, 0.11]	0.05 [−0.10, 0.21]	0.01 [−0.08, 0.09]	0.07 [−0.41, 0.56]
	Model 3	5.18 [4.91, 5.45]	0.01 [−0.08, 0.09]	0.07 [−0.08, 0.23]	0.01 [−0.08, 0.09]	−0.20 [−0.43, 0.03]
	Model 4	4.09 [3.33, 4.86]	0.03 [−0.05, 0.11]	0.03 [−0.12, 0.19]	0.01 [−0.08, 0.09]	0.18* [0.05, 0.32]
	Model 5	5.24 [4.65, 5.83]	0.03 [−0.05, 0.11]	0.04 [−0.12, 0.20]	0.01 [−0.08, 0.09]	−0.05 [−0.21, 0.11]

Note: Model 1 is random linear trend model; Model 2 is random linear trend model plus the within-subject effect of gender; Model 3 is random linear trend model plus the within-subject effect of class; Model 4 is random linear trend model plus the within-subject effect of self-esteem; Model 5 is random linear trend model plus the within-subject effect of social comparison.
*p < .05.

but it is hard to know if this decrease was caused by the summer gifted programs, as there were a lot of other moderators not accounted for by this model. The short-term effects of the program were not significant on any of the three academic self-concepts, nor was its interactions with time. Only the moderations of self-esteem were positively significant at $p < .05$ level on general academic self-concept and verbal self-concept, but not on math self-concept.

Discussion and Conclusion

In general, this study does not find a prevalent pattern of declines in academic self-concepts as a result of participating in a summer gifted program, known as the BFLPE. Only a few cases of increases and decreases in academic self-concepts can be identified in our plots. Longitudinal patterns also do not support a BFLPE perturbation explanation. It is important to note up front that the null finding does not falsify the BFLPE theory per se; rather it raises questions about conditions under which the BFLPE is more likely to be observed. Although Marsh et al. (2008) considered gifted programs less optimal for testing the BFLPE theory, the relevance and applicability of the BFLPE to gifted programs is obvious. Indeed, in a qualitative study of an early college entrance program conducted by the first author (under review), BFLPE was found to be a distinct experience for quite a few students. Then, why is it in summer programs like the one we investigate here and many others (e.g., Cunningham & Rinn, 2007; Makel et al., 2012) the BFLPE does not seem to be prevalent, whereas in replacement programs (Marsh et al., 1995) the BFLPE appears to be a more distinct issue? The distinction Makel et al. (2012) made between replacement and supplementary programs provides an important clue. Although not included in our measurement, we suspect that a 3-week summer program like the one we investigate has several features that mitigate against the BFLPE. First, these programs are short in duration, and students take various classes in such a program. Therefore, there is not a sufficient basis for forming a relatively stable peer group as a frame of reference for social comparison. We also suspect that in such short summer programs, social-evaluative pressure is not as salient, and peer performance information is less accessible than self-contained replacement programs. Without a clear sense of how well the peer group performs, the BFLPE is less likely to occur.

Alternative explanations are also plausible. It could be that the positive effects of participating in gifted programs (i.e., assimilation effects, such as identification, inspiration, and reflected glory) on self-concept offset the negative ones (i.e., contrast effects, such as the BFLPE), resulting in no distinct changes either way. For example, gifted programs may provide better curriculum and instruction as well as more inspiration and stimulation than regular programs; this positive impact can, in turn, "confound" what otherwise would be deemed "pure" measures of effects of social comparison (Marsh et al., 2008). Unfortunately, the present study did not include a direct measure of assimilation effects. However, the longitudinal patterns of the three academic self-concepts (the first three diagrams in Figure 5) show a distinct contrast

between positive changes during the 3-week program, and a decline during the internal between the Year 1 and Year 2 summer programs. The natures of the positive and negative changes warrant further research. A recent study (Huguet et al., 2009) showed it is possible to tease apart both assimilation and contrast effects, hence clarifying circumstances for both the positive and negative impact on self-concept. In our follow-up study, we plan to use other variables in the data set to examine this issue, such as changes in participants' educational and career aspirations and possible selves.

A third explanation is that students used different frames of reference, beyond the program or school boundaries. Therefore, participating in a summer program might produce a negative situational effect on academic self-concept, but there are multiple standards of comparison students will use that are powerful enough to wash off a distinct pattern of declines in academic self-concept in the face of a more competent peer group than they typically have in regular classrooms. We suggest this was more likely the case, particularly when the exposure to the new peer group in such a summer program was not a prolonged and exclusive one (see Dai & Rinn, 2008).

The fourth explanation is that academic self-concept has its own developmental continuity. Particularly during adolescence, individuals' academic self-concepts are more differentiated and consolidated and may be less susceptible to situational influences. For that matter, gifted students' academic self-concepts may be crystallized earlier than other peers due to their more distinct performance and achievement in their domains of strengths. To the extent that academic self-concepts become stabilized, the situational BFLPE may be of less leverage in self-evaluation of academic competencies. Such an explanation is to some extent supported by the fact that most students' self-concepts remained virtually unchanged or unperturbed before and at the end of the summer program. Only those with relatively low self-concepts show some fluctuation and volatility.

Although the results of this study do not suggest the high prevalence of the BFLPE in summer programs for the gifted, they do indicate the presence of the BFLPE under certain tractable conditions. The finding of this study that lower academic self-concept was more volatile than higher academic self-concept is consistent with the theoretical prediction of social comparison theory that when individuals are less certain about their abilities, they are more likely to seek social comparison information for self-evaluation purposes (Festinger, 1954). This gives us further clues as to where the BFLPE might occur and where interventions should be aimed. It is also worth noting that gender moderates a possible BFLPE in such programs. Namely, female students are more likely to experience declines in verbal self-concept, a finding consistent with the general literature that females are more sensitive to social comparison information when evaluating themselves (see Dai, 2002). Although statistically not significant due to small sample size and limited range of variations, the moderation of math self-concept by attending STEM-related class is also worth mentioning. It suggests the specificity of a possible BFLPE in a more advanced academic setting. It is possible that performance standards in STEM-related classes are more objective and evaluation information more consistent than language arts and social studies classes, leading to a slight decline in math self-concept.

There are strategic and methodological ramifications of testing the BFLPE in terms of "where," "who," and "when." A unique strength of the present study is its longitudinal follow-up of two (and to a more limited scope, three) consecutive participations. The longitudinal patterns of self-concepts give us a better sense of how much impact participating in a challenging summer program might have on academic self-concept in the context of long-term stability and change in academic self-concepts. It allows us to put the possible program effects in broader developmental and social contexts. This study also incorporated several moderators that potentially accentuate or mitigate against the BFLPE. However, given the consistent findings of the lack of a distinct BFLPE in summer gifted programs, a more productive research strategy would be to systematically compare supplementary and replacement programs, and explore positive as well as negative impacts of these programs on academic self-concept and other psychosocial and motivational outcomes. More refined instruments can be incorporated to tease apart the positive (assimilation) and negative (contrast) effects of gifted programs on academic and social self-concepts. But more importantly, contextual information needs to be incorporated into research design about these programs, such as the nature of classes and activities offered, and the nature of social dynamics and information exchanges among participants.

As Dai and Rinn (2008) pointed out, a distinct weakness of the BFLPE paradigm had been a lack of direct measurements of social comparison. This situation is changing. Huguet et al. (2009) attempted to derive data-based social comparison indexes. They separated two social comparison processes and their respective effects, one forced by the salient social comparison information and the other deliberately engaged or chosen by some individuals. In this study, the social comparison disposition measure was incorporated to estimate the latter effect. This disposition measure did not do well in predicting changes (increases or declines) in academic self-concept. A simple explanation is the lack of a BFLPE in the first place. But a more technical explanation is that the measurement is too general and not sensitive enough to capture how individuals make social comparison in academic settings, in terms of frequency, direction (upward or downward), and valence (positive and negative). Future research should adapt social comparison measures to specific circumstances to improve their predictive efficacy.

Another issue that has been drawing increasing attention is the way the BFLPE is estimated. Marsh and colleagues used a paradigm (Marsh & Hau, 2003) that estimates the BFLPE through a between-subject and between-group design. Makel et al. (2012) pointed out that a within-subject estimate (i.e., pre–post change or contrast scores, used initially by Marsh et al., 1995) may yield a more direct measure of the BFLPE. The present study also used the within-subject estimate. More research is needed to understand how contrast scores can be formed while being sensitive to the distribution properties. For example, the present study found more dispersion at the lower end of the self-concept spectrum, suggesting that change estimates and errors are not invariant across the distribution. Traditional ANOVA or regression analyses treat residual variances as "errors" or "noises." The data plotting technique we adopted in the present study (Figures 2-4) allows us to spot those cases of ups and downs in self-concepts for diagnostic purposes. This is particularly important when the question of "who" is concerned.

In summary, we set out in this study to explore when and where the BFLPE is particularly acute, and for whom it has the most distinct effect. The findings of longitudinal stability and changes in academic self-concept while participating in a summer gifted program for a 3-year period put the BFLPE in perspective, while suggesting a more complex reality of positive and negative effects and consequently a more refined methodological strategy. The apparent lack of the BFLPE in our findings does not necessarily refute the theory itself but can be attributed to many mitigating conditions. We suggest that future research use a similar longitudinal design but with special care for teasing out both positive and negative effects of participating in gifted programs (supplementary vs. replacement) on academic self-concepts, situating these effects in rich individual, developmental, and social contexts, and elucidating underlying mechanisms.

Authors' Note

Declaration of Conflicting Interests

Funding

Note

1. The Duke Talent Identification program, for example, is the largest of its kind and identifies academically talented seventh graders based on standardized test scores achieved while attending elementary or middle school. Candidates are invited to take the ACT or the SAT college entrance exam as seventh graders, as above-level testing at least 2 years above a student's current grade placement offers gifted students and their families a far better understanding of how the student compares with his or her gifted peers and what level of educational challenge is appropriate.

References

Byrne, B. M., & Shavelson, R. J. (1996). On the structure of social self-concept for pre-, early, and late adolescents: A test of the Shavelson, Hubner, and Stanton (1976) model. *Journal of Personality of Social Psychology*, *70*, 599-613.

Cunningham, L. G., & Rinn, A. N. (2007). The role of gender and previous participation in a summer program on gifted adolescents' self-concepts over time. *Journal for the Education of the Gifted*, *30*, 326-352.

Dai, D. Y. (2002). Are gifted girls motivationally disadvantaged? Review, reflection, and redirection. *Journal for the Education of the Gifted*, *25*, 315-358.

Dai, D. Y. (2004). How universal is the big-fish-little-pond effect? *American Psychologist*, *59*, 267-268.

Dai, D. Y., & Rinn, A. N. (2008). The big-fish-little-pond effect: What do we know and where do we go from here? *Educational Psychology Review*, *20*, 283-317.

Dai, D. Y., Steenbergen-Hu, S., & Zhou, Y. (under review). Cope and grow: Lived experiences of students in an early college entrance science program. *Journal of Educational Psychology*.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and child development*. Philadelphia, PA: Psychology Press.

Eccles (Parsons), J., Adler, T., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, *46*, 26-43.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117-140.

Gibbons, F. X., Benbow, C. P., & Gerrard, M. (1994). From top dog to bottom half: Social comparison strategies in response to poor performance. *Journal of Personality and Social Psychology*, *67*, 638-652.

Gibbons, F. X. M., & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, *76*, 129-142.

Gilman, R., Laughlin, J. E., & Huebner, E. S. (1999). Validation of the Self-Description Questionnaire–II with an American sample. *School Psychology International*, *20*, 300-307.

Huguet, P., Dumas, F., Marsh, H., Régner, I., Wheeler, L., Suls, J., & Nezlek, J. (2009). Clarifying the role of social comparison in the big-fish-little-pond-effect (BFLPE): An integrative study. *Journal of Personality and Social Psychology*, *97*, 156-170.

Humes, C. W., II, & Campbell, R. D. (1980). Gifted students: A 15-year longitudinal study. *Gifted Child Quarterly*, *24*, 129-131.

Kolloff, P. B., & Moore, A. D. (1989). Effects of summer programs on the self-concepts of gifted children. *Journal for the Education of the Gifted*, *12*, 268-276.

Lennox, R. D., & Wolfe, R. N. (1984). Revision of the Self-Monitoring Scale. *Journal of Personality and Social Psychology*, *46*, 79-122.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York, NY: John Wiley.

Makel, M. C., Lee, S.-Y., Olszewski-Kubilius, P., & Putallaz, M. (2012). Changing the pond, not the fish: Following high-ability students across different educational environments. *Journal of Educational Psychology*. Advance online publication. doi:10.1037/a0027558

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, *79*, 280-295.

Marsh, H. W (1990). *Self-Description Questionnaire (SDQ) II: Manual*. New South Wales, Australia: University of Western Sydney.

Marsh, H. W., Chessor, D., Craven, R., & Roche, L. (1995). The effects of gifted and talented programs on academic self-concept: The big fish strikes again. *American Educational Research Journal*, *32*, 285-319.

Marsh, H. W., & Hau, K.-T. (2003). Big-fish-little-pond-effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academic selective schools. *American Psychologist*, *58*, 364-376.

Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K., O'Mara, A., & Craven, R. (2008). The big-fish-little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, *20*, 319-350.

Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The big-fish-little-pond-effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal*, *44*, 631-669.

Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: Relations to academic self-concept and achievement. *American Educational Research Journal*, *34*, 691-720.

Moon, S. M., Feldhusen, J. F., & Dillon, D. R. (1994). Long-term effects of a pullout enrichment program based on the Purdue Three-Stage Model. *Gifted Child Quarterly*, *38*, 38-48.

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experiences, task choice, and performance. *Psychological Review*, *91*, 328-346.

Plucker, J. A., Robinson, N. M., Greenspon, T. S., Feldhusen, J. F., McCoach, D. B., & Subotnik, R. F. (2004). It's not how the pond makes you feel, but rather how high you can jump. *American Psychologist*, *59*, 268-269.

Plucker, J. A., Taylor, J. W. V., Callahan, C. M., & Tomchin, E. M. (1997). Mirror, mirror, on the wall: Reliability and validity evidence for the Self-Description Questionnaire II with gifted students. *Educational and Psychological Measurement*, *57*, 704-713.

Pruzek, R. M., & Helmreich, J. E. (2009). Enhancing dependent sample analyses with graphics. *Journal of Statistics Education*, *17*(1). Retrieved from www.amstat.org/publications/jse/v17n1/helmreich.html

Rinn, A. N. (2007). Effects of programmatic selectivity on the academic achievement, academic self-concepts, and aspirations of gifted college students. *Gifted Child Quarterly*, *51*, 232-245.

Ruble, D. N., Grosovsky, E. H., Frey, K. S., & Cohen, R. (1992). Developmental changes in competence assessment. In A. K. Boggiano & T. S. Pittman (Eds.), *Achievement and motivation: A social-developmental perspective* (pp. 138-164). New York, NY: Cambridge University Press.

Schafer, J. L., & Yucel, R. M. (2002). Computational strategies for multivariate linear mixed-effects models with missing values. *Journal of Computational and Graphical Statistics*, *11*, 437-457.

Seaton, M., Marsh, H. W., & Craven, R. G. (2010). Big-fish-little-pond effect: Generalizability and moderation—Two sides of the same coin. *American Educational Research Journal*, *47*, 390-433.

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

Taylor, S. E., Wayment, H. A., & Carrillo, M. (1996). Social comparison, self-regulation, and motivation. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition, Vol. 3: The Interpersonal context* (pp. 3-27). New York, NY: Guilford.

Thomas, T. A. (1989). *Acceleration for the academically talented: A follow-up of the academic talent search class of 1984*. California. Retrieved from ERIC database. (ERIC Document Reproduction Service No. ED307303)

Tymms, P. (2001). A test of the big fish in a little pond hypothesis: An investigation into the feelings of seven-year-old pupils in school. *School Effectiveness and School Improvement*, *12*, 161-181.

Zeidner, M., & Schleyer, E. J. (1998). The big-fish-little-pond effect for academic self-concept, test anxiety, and school grade in gifted children. *Contemporary Educational Psychology*, *24*, 305-329.