American Educational Research Journal

The Big-Fish-Little-Pond Effect and a National Policy of Within-School Ability Streaming : Alternative Frames of Reference

Gregory Arief D. Liem, Herbert W. Marsh, Andrew J. Martin, Dennis M. McInerney and Alexander S. Yeung

The online version of this article can be found at:

Published on behalf of

American Educational Research Association

and

SAGE

http://www.sagepublications.com

Additional services and information for *American Educational Research Journal* can be found at:

Email Alerts: http://aerj.aera.net/alerts

Subscriptions: http://aerj.aera.net/subscriptions

Reprints: http://www.aera.net/reprints

Permissions: http://www.aera.net/permissions

>> Version of Record - Mar 22, 2013

OnlineFirst Version of Record - Nov 15, 2012

What is This?

The Big-Fish-Little-Pond Effect and a National Policy of Within-School Ability Streaming: Alternative Frames of Reference

Gregory Arief D. Liem
Nanyang Technological University
University of Sydney
Herbert W. Marsh
University of Western Sydney
University of Oxford
King Saud University
Andrew J. Martin
University of Sydney
Dennis M. McInerney
Hong Kong Institute of Education
Alexander S. Yeung
University of Western Sydney

The big-fish-little-pond effect (BFLPE) was evaluated with 4,461 seventh to ninth graders in Singapore where a national policy of ability streaming is implemented. Consistent with the BFLPE, when prior achievement was controlled, students in the high-ability stream had lower English and mathematics self-concepts (ESCs and MSCs) and those in the lower-ability stream had higher ESCs and MSCs. Consistent with the local-dominance effect, the effect of stream-average achievement on ESCs and MSCs was more negative than—and completely subsumed—the negative effect of school-average achievement. However, stream-average achievement was stronger than, or as strong as, the more local class-average achievement. Taken together, findings highlight the potential interplay of a local dominance effect with variability and/or salience of target comparisons in academic self-concept formations.

KEYWORDS: academic self-concept, big-fish-little-pond effect, Singapore, ability stream, social comparison

Positive academic self-concept, or students' favorable perceptions of their academic achievement, has been seen as a desirable quality in its own right and a critical factor that facilitates growth of other valued educational outcomes (for reviews, see Branden, 1994; Marsh, 2007; Marsh & Craven,

2006). As synthesized in a meta-analysis by Valentine, DuBois, and Cooper (2004), the preponderance of evidence drawn from longitudinal studies has shown that academic self-concept and achievement—both teacher-assigned grades and standardized test scores—have reciprocal effects, with enhancements in prior academic self-concept and achievement leading to improvements in subsequent achievement and academic self-concept. These effects have been consistently found at different developmental stages (e.g., childhood, adolescence), in various performance domains (e.g., academic, sports), and in cross-cultural research—resulting in the reciprocal effect model (REM; for a review, see Marsh & Craven, 2006; see also e.g., Guay, Marsh, & Boivin, 2003; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005, for more specific studies). Marsh and O'Mara (2008), for example, conducted a longitudinal analysis of the Youth in Transition (YIT) database that comprises a large and nationally representative sample of 10th-grade boys in U.S. public school. The YIT survey was conducted on five occasions over 8 years, from Year-10 to 5 years after high school graduation (see Bachman, 2002, for a more detailed description of this database). Their analysis showed that students' academic self-concepts in Year-10 were a better predictor of their educational attainments 5 years after high school graduation than their school grades, standardized achievement test scores, intelligence, and socioeconomic status and that students' school grades also predicted their academic self-concepts after high school—attesting to the reciprocal, long-term, and causal effects between academic self-concept and achievement.

GREGORY ARIEF D. LIEM is an assistant professor at Psychological Studies Academic Group, National Institute of Education, Nanyang Technological University, Singapore. His research interests center on the role of self-concept, motivation, engagement, and character strengths in human development.

HERBERT W. MARSH holds a joint appointment at the Centre for Positive Psychology and Education at the University of Western Sydney and at Oxford University; e-mail: *H.Marsh@uws.edu.au*. His major research/scholarly interests include self-concept and motivational constructs; evaluations of teaching effectiveness; developmental psychology, quantitative analysis; substantive-methodological synergy, value-added and contextual models; sports psychology; the peer review process; gender differences; peer support and anti-bullying.

ANDREW J. MARTIN is a professorial research fellow in the Faculty of Education and Social Work at the University of Sydney. He specializes in motivation, engagement, and quantitative methods.

DENNIS M. MCINERNEY is a chair professor of educational psychology at the Hong Kong Institute of Education. His research interests include cross-cultural studies of learning and motivation, and self-processes such as self-regulation and self-concept.

ALEXANDER S. YEUNG is an associate professor in the Centre for Positive Psychology and Education, University of Western Sydney. His specializations include self-concept, cognition and instruction, motivation, and teachers and teaching.
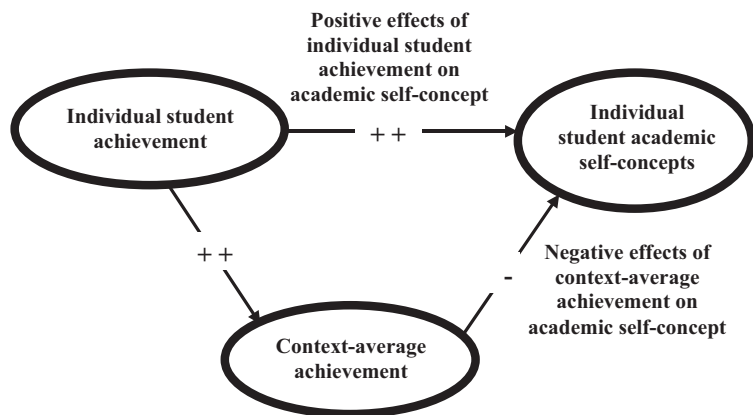
Figure 1. **The big-fish-little-pond effect (BFLPE) (adapted from Marsh, 2007).**

Beyond its achievement yields, academic self-concept in high school has been found to be more salient than actual academic achievement in predicting learning effort, educational and occupational aspirations, and subsequent university course selection and attendance (Guay, Larose, & Boivin, 2004; Marsh, 1991; Marsh & O'Mara 2008; Marsh & Yeung, 1997). Guay and his colleagues (2004), for example, found that students' positive academic self-concepts were associated with better educational outcomes 10 years later—the findings that these researchers concluded as providing "good support for the long-lasting effects of academic self-concept" (p. 64). Collectively, these reviews suggest that promoting positive academic self-concepts is a crucial approach to optimizing achievement and other educational accomplishments and that if students' academic self-concepts are inadvertently undermined, then these lowered self-beliefs are likely to undermine subsequent educational outcomes, including academic and occupational aspirations.

The literature has established that students' academic self-concepts are partly developed through a social comparison process in which students use the achievement of their peers as a frame of reference to judge their own achievement (Marsh, 1987, 2007; Marsh et al., 2008; Skaalvik & Skaalvik, 2002). This process is encapsulated in the big-fish-little-pond effect (BFLPE) model (Marsh, 1987; see Figure 1) positing that when the positive effect of individual student achievement on academic self-concept is taken into account (the brighter I am, the better my academic self-concept), class-average and school-average achievement has a negative effect on

academic self-concept (the brighter my classmates, the lower my academic self-concept) (see also Marsh, 2007; Marsh et al., 2008).

The negative effect of school-average achievement (i.e., the BFLPE) has long-term implications. For example, in a more recent analysis of the YIT database, Marsh and O'Mara (2010; see also Marsh, Kong, & Hau, 2000) demonstrated that school-average achievement had long-term negative effects not only on academic self-concepts a year after graduation but also on educational and occupational attainments and aspirations 5 years after high school. In almost all cases, the negative long-term effects of school-average achievement were substantially mediated by academic self-concept—further attesting to the importance of academic self-beliefs in facilitating the growth of key educational and occupational outcomes and also to the need to attenuate the negative BFLPE on academic self-concepts of students in schools with high average achievement. In the present study we extended prior research by examining the effects of within-school ability grouping on students' academic self-concepts—and compared these effects with the effects of school-average and class-average achievements as well as with those of students of the same gender or ethnicity who were in the same class or stream.

The BFLPE and Ability Grouping

Early evidence of the BFLPE in ability group settings. The BFLPE framework has shed light on early research into ability grouping or tracking effects. In a meta-analytic study, Kulik and Kulik (1982) reported that ability-grouped and non–ability-grouped students did not differ systematically in their academic self-concepts. Consistent with BFLPE predictions, however, Marsh (1984) pointed out that the Kulik and Kulik meta-analysis confounded negative effects of placement in high-ability groups with positive effects of placement in low-ability groups. It was not surprising, therefore, that ability-grouping effects were small and nonsignificant when tracking effects were averaged across high-ability and low-ability groups. In a subsequent reanalysis of their results, Kulik (1985) confirmed Marsh's predictions based on the BFLPE when the effects of high-ability and low-ability grouping were considered separately (see Hattie, 2002, for a more recent meta-analysis of ability grouping research, supporting the BFLPE model).

BFLPEs in tracked-school settings. More recent evidence has shown the operation of BFLPEs across education systems in which ability grouping is implemented. Based on a large sample of 14,341 German students from upper-, middle-, and lower-ability schools, Trautwein, Lüdtke, Marsh, Köller, and Baumert (2006, Study 1) found that math self-concepts were positively related to individual student achievement but negatively related to school-average achievement. This shows that when students' prior achievement was controlled, students in the high-track schools had relatively lower

academic self-concepts than those in middle- and low-track schools. The same pattern was also reported by Marsh et al. (2000) in a longitudinal study of 7,997 Hong Kong students assigned to secondary schools differentiated by ability bands. In this study, they found that school-average achievement had negative effects on students' general academic self-concepts and these effects were consistent over time and in size, with regression coefficients (βs) ranging between –.20 and –.30. Given the focus on school effects, these studies appropriately took into account the potentially shared attributes of students from the same school and, hence, based their analyses on two-level models (students at Level 1 and school at Level 2).

BFLPEs and within-school ability grouping. Of particular relevance to the present study, Trautwein et al. (2006, Study 2) examined the effect of within-school ability grouping with a sample of 3,243 German students from schools that provide education for students at all levels of achievement. Based on their prior achievement, students in these schools were assigned to two (or three) domain-specific ability groups (in math, foreign languages, or German)—thus, the ability grouping implemented was idiosyncratic to each school. Consistent with BFLPE predictions, their three-level analysis, in which students were nested within streams that were nested within schools, showed that stream-average achievement was a negative predictor of students' math self-concepts. Their study, however, did not consider and juxtapose the effects of school- and class-average achievement, two other contextual predictors that have been found in prior studies as key frames of reference leading to the BFLPE.

Ireson and colleagues (Ireson & Hallam, 2009; Ireson, Hallam, & Plewis, 2001) examined academic self-concepts of students from secondary schools in England. In these studies, schools varied in the extent to which they implemented ability grouping: Some schools tracked students for all academic subjects, some tracked students only on a few subjects, and some others did not track at all. Ireson and colleagues found significant but little effects of the differential degree of tracking practices such that students from schools implementing more ability grouping had relatively lower general academic self-concepts. The main effects of tracking, however, were not found on academic self-concepts in English, math, or science. Although Ireson and colleagues based their analyses on two-level models in which students were nested within schools, they did not consider the track level a student was in or class or track as a unit of analysis. It is therefore not surprising that the reported effects of ability grouping were small and inconsistent particularly given that—as noted earlier—the reverse effects of placement in high and low tracks would lead to confounded negligible effects when analyses do not separate the effects of the different track levels (Hattie, 2002; Kulik, 1985; Marsh, 1984).

In a recent study, Nagengast and Marsh (2011) analyzed the 2006 Programme for International Student Assessment (PISA) database and compared the sizes of BFLPEs across the four United Kingdom countries differing

330

in the degree of selectivity of their secondary school systems. Of the four UK countries, Northern Ireland implemented the most selective school system. In England some schools were selective and some were comprehensive. In Wales and Scotland all schools were comprehensive. Their analysis indicated that the negative effects of school-average achievement on science self-concept were largest in Northern Ireland, nearly as large in England, but relatively small in Wales and Scotland. This study is substantively important by showing that BFLPEs are more negative in the contexts with larger between-school differences in school-average achievement. As noted by these researchers, however, because PISA 2006 did not include information on classrooms and streams, their analysis was not able to disentangle and juxtapose the potentially competing frames of references associated with the achievement of students in the class, the stream, and the school.

Filling the Gaps in Prior Research

Collectively, existing studies on the BFLPE in ability group settings have shed light on the effects of the different track levels on students' evaluations of their academic achievement. More specifically, it has been shown that the effects of high tracks are negative and low tracks are positive. That is, context-average achievement—the average achievement of students in a given context—has reversed effects on students' academic self-concepts, and the size of these effects corresponds with the variability of achievement of students in the context used as a frame of reference of social comparisons.

Unfortunately, most prior research into the BFLPE in ability group settings are substantively limited and methodologically flawed in several ways. First, prior research was conducted in the educational contexts in which the implementation of within-school ability grouping practices was not uniform or common across schools (see Oakes, 1985). Hence, generalizations about the effects of streaming across different schools are called into question. Second, prior studies have confounded the positive effects of low tracks and the negative effects of high tracks on academic self-concepts. Hence, it is not a surprise that the effects of ability grouping were negligible and nonsignificant when averaged across the different track levels (see Hattie, 2002). Third, prior investigations were restricted to the use of either class-average achievement, stream-average achievement, or school-average achievement as a contextual predictor of academic self-concept. The absence of these three contextual predictors in the one study does not attest to their relative salience as frames of references leading to BFLPEs (Nagengast & Marsh, 2011). Lastly, most existing studies have based their analyses on only two-level models with student at the first level and class or stream or school at the second level. The failure of considering the hierarchical nature of the data, especially in a study that examines contextual

effects, may lead to dubious findings attributed to aggregation biases (Raudenbush & Bryk, 2002).

Singapore educational context. In the present investigation we considered an apparently unique situation in the Singapore education system characterized by the implementation of national educational policy and curricula that guide instructions in all schools and national examinations. Based on academic performance obtained in Primary-4, Primary-5, and Primary-6 pupils are tracked into three main streams: EM1 (higher), EM2 (standard), and EM3 (foundation) streams in which pupils differ in their levels of proficiency in English, math, science, and mother tongue. While EM1 and EM2 pupils take English, math, and science at the standard level, the former take mother tongue at the higher level and the latter take this subject at the standard level. Unlike EM1 and EM2 pupils, EM3 pupils take all the four subjects in the foundation level. In the final year of their primary education (Primary-6), all pupils are required to take a set of high-stakes standardized achievement tests in the four academic subjects (i.e., Primary School Leaving Examination or PSLE). It is individual students' PSLE aggregate score across these four subjects that constitutes the basis to assign them into one of the three core ability streams in the secondary school (i.e., Express, Normal Academic, and Normal Technical streams—hereafter called high-ability, middle-ability, and low-ability streams). When a student shows an outstanding performance, however, he or she is allowed to move to a "higher" ability stream (e.g., from the low-ability stream to the middle-ability stream or from the middle-ability stream to the high-ability stream).[1]

To our knowledge, this is apparently the first BFLPE study conducted in an education setting that implements a *nationwide within-school ability grouping* practice. This setting allows us to test with a large sample the predicted positive effect of high-ability streams and the predicted negative effect of low-ability streams as students are placed to different ability groups with comparable cut-off ranges of PSLE aggregate scores across schools. Furthermore, while we know of no BFLPE studies that have examined the relative salience of the different levels of contexts, in the present investigation we juxtaposed BFLPEs in relation to three different frames of reference—the school, the stream within the school, and the class within the stream—in the one analytic study. In doing so, we based our analyses on four-level models, taking into account student, class, stream, and school as different units of analysis.

Alternative Frames of Reference in the BFLPE: Theoretical Perspectives

This research is particularly relevant to addressing two distinct but related theoretical perspectives in social comparison theory that have not been considered simultaneously in prior BFLPE research: the level of locality

and the characteristic of the target frames of reference. The local dominance effect (Alicke, Zell, & Bloom, 2010; Zell & Alicke, 2009, 2010) posits that individuals tend to use the most "local" (or proximal) frame of reference to inform their self-evaluations of competence, even when they know that their local group is less representative and when they are aware of other comparative information that is more broadly representative. With respect to referent characteristics, social comparison perspectives (Festinger, 1954; Marsh et al., 2008; Suls & Wheeler, 2008) predict that students are most likely to choose other students who share salient attributes or characteristics as referents in evaluating their own achievement. These two perspectives and their empirical evidence are reviewed in more detail in the following.

Local Dominance Effect

Empirical evidence for the BFLPE also come from laboratory studies with random assignment to conditions (e.g., Cleveland, Blascovich, Gangi, & Finez, 2011; Seta & Seta, 1996; Zell & Alicke, 2009, 2010). Of particular relevance to the present study, Zell and Alicke (2009) pitted "local" against more "general" frames of reference or comparison standards and tested the hypothesis that the effects of local comparison information on individuals' self-evaluations *supersede* or *dominate* those of more general comparison information—the local dominance effect. In their experiments, Zell and Alicke (2009) asked participants to complete a verbal reasoning task and gave them different levels of comparative feedback. Three feedback sources were manipulated in different combinations, ranging from most local to most general; the experimenters provided information on how well participants performed in relation to a small group of five (most local), to almost 1,500 other test-takers (intermediate), and to other schools (most general). Some participants received all levels of feedback, some received two, and some received one. Consistent with their hypothesis, participants in each condition used the most local comparison information available to them. Interestingly, the three feedback sources had comparable effects when given alone.

Zell and Alicke (2009) provided compelling evidence that individuals based their self-evaluations on the most local frame of reference when more than one comparison standard was available. The local dominance effect experiments provided clear evidence that even when participants had access to frames of reference with a stronger diagnostic value in providing information about one's relative standing to broader populations, the most local one dominated the more general ones. Predictions based on the local dominance effect offer a potentially important extension of existing BFLPE research. However, support for the local dominance effect is based on laboratory, experimental settings and not on large-scale naturalistic applied contexts such as schools. Studies of the local dominance effect

333

have strong internal validity based on the use of random assignment but are not as strong in terms of the external validity of the experimental manipulation to establish alternative frames of reference like those that students actually use in school settings.

To date, there has been no BFLPE study that has specifically juxtaposed the effects of the different levels of frame of reference on academic self-concept. One study close to this juxtaposition was conducted by Rogers, Smith, and Coleman (1978) who ranked a small group of underachieving elementary school pupils (N = 159) from 17 classrooms in seven schools relative to other students in their own classroom and relative to those of the total sample. These researchers found that consistent with the local dominance effect, pupils' self-perceptions in a diverse set of domains (e.g., behavior, intellectual and school status, popularity, anxiety) were more highly associated with their within-classroom rankings than their rankings relative to the total sample. This study, however, did not specifically examine domain-specific academic self-concepts although the participants were rank-ordered according to their reading and math achievement, did not perform a multilevel analysis that takes into account similar attributes of pupils within the same class or school, and, as noted by the researchers, generalizability of the finding was limited due to the use of a small sample comprising only low-achieving primary school pupils.

The present study, then, aimed to test the local dominance effect in an applied setting and extended prior work through its multilevel examination of BFLPEs by juxtaposing the effects of school-, stream-, and class-average achievement as frames of reference varying in their degrees of locality. More specifically, both the BFLPE and the local dominance effect predict that school-, stream-, and class-average achievement should each have a negative effect when considered separately. However, according to the local dominance effect hypothesis, the negative effects of class-average achievement (the most local frame of reference) should dominate the negative effects of stream-average achievement (the middle-level frame of reference), which in turn should dominate the negative effects of school-average achievement (the most general frame of reference).

Specific Target Referents

Skaalvik (1997; Skaalvik & Skaalvik, 2002) maintained that the use of contextual frames of reference leading to BFLPEs does not always mean that students evaluate their achievement against the aggregate achievement of *all* other students in a given achievement context. Skaalvik and Skaalvik (2002) suggested that in addition to the *generalized other*, students are likely to use a *specific group of students* in the achievement context as a target reference. According to a *related attribute* hypothesis (Dijkstra, Kuyper, van der Werf, Buunk, & van der Zee, 2008), students are more likely to select a target

group whose members share proximities with them in salient characteristics. This is consistent with Festinger (1954) who stated that an individual's need for accurate self-evaluations leads the person to select similar others as a target comparison.

Past studies have shown the salience of *gender* and *ethnicity* as two attributes that comparers sought in their referents for social comparisons (e.g., Blanton, Buunk, Gibbons, & Kuyper, 1999; Huguet et al., 2009; see also Dijkstra et al., 2008, for a review). Preckel and Brüll (2008), for example, found that students reported preferences to compare their test scores with other students of the same gender irrespective of the achievement of the comparison target. Huguet et al. (2009) demonstrated that when students were asked to nominate a specific comparison target, they chose other students who were of the same gender and slightly higher performers than themselves. In support of these findings, a recent study by Liem and Martin (2011) showed that it was students' perceptions of their relationships with same-sex peers, and not with opposite-sex peers, that directly affected their academic performance, suggesting the more salient role of same-sex peers than opposite-sex peers in a student academic trajectory.

Prior research has also provided evidence that adolescents typically view peers of the same ethnic background as more favorably than those from a different ethnic group (Aboud, 2003; Tarrant, 2002). Furthermore, there is some evidence that most of adolescents' friends are ethnically similar to them (Howes & Wu, 1990). Taken together, there are reasons to believe that students spend their time with same-ethnicity peers and, hence, have more access to and use information about these peers than those from different ethnic groups. In a study with White and Black American students, for example, Meisel and Blumberg (1990) found a significant pattern of students' preferences to compare their achievement with other students who were of the same ethnicity status. Similarly, in an early study of information-seeking strategies among students with different ethnicity status, Aboud (1976) found that both White and Chicano American students tended to select ethnically similar peers to evaluate their own achievement.

With a few exceptions (e.g., Huguet et al., 2009), there is surprisingly little research on BFLPEs that has juxtaposed the generalized others and the more specific groups in the achievement context that students may use as comparison targets. In the present study, we sought to investigate the BFLPE specific to the contexts of gender and ethnicity. In terms of ethnicity, the Chinese constitutes 75% of the Singapore population and is therefore regarded as the ethnic majority whereas the Malay, Indian, Eurasian, and people from other ethnic groups constitute 25% of the population and are considered as the ethnic minority. Here, we evaluated the effects of average achievement of students who were of the same gender or the same ethnicity status (majority or minority) who were in the same class (i.e., gender-class-average achievement or ethnicity-class-average achievement, respectively) or in the same

335

stream, grade, and school (i.e., gender-stream-average achievement or ethnicity-stream-average achievement, respectively) and juxtaposed these effects with the effects of class-average achievement and stream-average achievement. This juxtaposition seems to be a logical extension of the local dominance effect in that a specific group of students who share similar attributes (gender or ethnicity) within the class or stream can be viewed as a more local frame of reference than all students within the class or stream as a whole.

The Present Study

The nationally mandated ability streaming implemented in Singapore secondary education is a particularly suitable setting to extend BFLPE research. This is the case because it encompasses in the one naturalistic setting the three core ability streams (i.e., high-ability, middle-ability, and low-ability). More specifically, this Singaporean context enables us to address novel questions and unique issues in the overall program of BFLPE research in the following ways. First, it enabled us to examine the BFLPE across schools with a common ability grouping practice because the assignment of students to one of the three core streams is based on performances on the same achievement test and on comparable ranges of cut-off values across schools. Hence, this setting provided an avenue to test BFLPE predictions by separating the predicted positive effect of the low-ability stream from the predicted negative effect of the high-ability stream.

Second, the setting also allowed us to examine the extent to which stream-average achievement (the focus of this study) has a negative effect on academic self-concept compared with class- and school-average achievement (the focus of most BFLPE studies). Consistent with the BFLPE model (Marsh, 1987; Marsh et al., 2008), we predicted that after controlling for the positive effect of individual student achievement on academic self-concept, the average achievement of students in the class, the stream, and the school would have a negative effect on academic self-concept. Consistent with the local dominance effect hypothesis (Zell & Alicke, 2009, 2010), we predicted that these contextual predictors—varying in their degree of locality/generality—would have a negative effect on academic self-concept when considered separately. The more local frames of reference, however, were expected to have a more salient effect than the more general ones when they were considered together.

Third, the Singapore setting also enabled us to examine BFLPEs specific to the contexts of gender and ethnicity. Integrating BFLPE, social comparison, and local dominance perspectives, we tested the effects of frames of reference established by peers from the same class and stream who were also of the same gender or of the same ethnicity status. The ethnic composition of the Singapore population (approximately 75% Chinese, 25% non-Chinese) provided an ideal setting to test ethnicity-referenced BFLPEs. Substantively, this

336

investigation extended our test of the local dominance effect by juxtaposing the effects of average achievement of same-gender or same-ethnicity students in the same class or stream (the more local frames of reference) with those of all students in the same class or stream (the more general frames of reference). The extent that BFLPEs specific to the contexts of gender and ethnicity were present, the effects of gender- or ethnicity-average achievement should remain a significant predictor of academic self-concepts even when stream- or class-average achievement was simultaneously included as a predictor.

Lastly, the present study aimed to examine BFLPEs on self-evaluations in two specific domains, English and math self-concepts (ESC and MSC, respectively). While many BFLPEs studies have focused on MSC (e.g., Marsh, Trautwein, Lüdtke, Baumert, & Köller, 2007), those that examined verbal self-concepts have focused on perceived competence in reading (e.g., Marsh, 1987) or the samples' first language (e.g., Preckel & Brüll, 2008). Examination of ESC in the present study is particularly interesting because under a bilingualism policy (see Pakir, 1993), Singaporean students are generally proficient in both English (i.e., the medium of instruction at all Singaporean schools) and one other language (i.e., Mandarin for the Chinese, Malay for the Malays, or Tamil for the Indians). However, given that English is the official working language in the country, Singaporeans are taught by their parents and teachers since early childhood that English is important and instrumental to do well both academically and occupationally (Liem, Lau, & Nie, 2008). The importance of English in the Singaporean society may intensify the BFLPE on Singaporean students' ESCs.

Method

Sample

A survey was administered to 4,461 Singaporean Secondary-1 to Secondary-3 (or Grade 7 to Grade 9) school students from 136 classes in nine schools. The number of students drawn from each class ranged between 20 and 42 (M = 32.80, SD = 5.32). Distributions of the participants based on stream and grade are presented in Table 1. The average age of the sample was 13.94 (SD = 1.07; range: 11-19). In total, 2,005 (44.9%) of the participants were girls and 2,422 (54.3%) were boys (34 [0.8%] missing values). In terms of ethnicity, 3,056 (68.5%) of the participants were Chinese, 897 (20.1%) were Malay, 179 (4%) were Indian, and the remaining 329 (7.4%) were categorized as "Others" (e.g., Eurasian, Filipino, Japanese). This distribution of students across streams and across ethnic groups represented the recent proportions of secondary school students and general population in Singapore, respectively (Singapore Department of Statistics, 2010; Singapore Ministry of Education, 2009).

Table 1
Sample Distribution by Ability Stream and Grade

	Stream						Total	
	High Ability		Middle Ability		Low Ability			
	n Student (%)	n Class (%)	n Student (%)	n Class (%)	n Student (%)	n Class (%)	n Student (%)	n Class (%)
Grade 7	768 (17.22)	23 (16.91)	503 (11.28)	15 (11.03)	345 (7.73)	11 (8.09)	1,616 (36.22)	49 (36.03)
Grade 8	732 (16.41)	21 (15.44)	441 (9.89)	13 (9.56)	260 (5.83)	10 (7.35)	1,433 (32.12)	44 (32.35)
Grade 9	697 (15.62)	20 (14.71)	495 (11.10)	15 (11.03)	220 (4.93)	8 (5.88)	1,412 (31.66)	43 (31.62)
Total	2,197 (49.25)	64 (47.06)	1,439 (32.26)	43 (31.62)	825 (18.49)	29 (21.32)	4,461 (100)	136 (100)

As English is the primary medium of instruction at schools and the official working language, Singaporeans in general are proficient in English. Under Singapore's bilingualism policy (see Pakir, 1993), however, Singaporean students are required to study their mother tongue. Hence, 42.4% of our participants reported that they were also proficient in Mandarin Chinese, 14.7% in Malay, and 1.4% in Tamil. Socioeconomic status (SES) background of the students' families was inferred from the information obtained from the participants about their parents. Paternal and maternal highest levels of education spanned from primary to tertiary education, with the majority having completed secondary education. Parents were of a wide and diverse range of occupations (e.g., factory worker, taxi driver, doctor, lawyer). The sampling of schools was carried out in a way that ensured representation of each Singapore's educational jurisdiction (North, South, West, and East). Given the sampling procedure, the sample size, and the range of sample characteristics (i.e., ethnic groups, languages spoken at home, and SES), the sample was broadly representative of Singaporean Secondary-1 to Secondary-3 school students.

Measures

The survey was administered in English, the medium of instruction at all schools in Singapore. The measures, samples of items, and the Cronbach's internal consistency reliability computed with the data of the present study are reported in the following.

Academic self-concept. To measure English self-concept and math self-concept, we used the Self-Description Questionnaire II (SDQ-II; Marsh, 1992), a multidimensional measure of self-concept that is considered to be one of the most robust self-concept instruments (Byrne, 1996). The ESC scale, comprising five items ($\alpha = .93$; I am good at English; I get good marks in English; I have always done well in English; Work in English is easy for me; I learn things easily in English), measures students' self-evaluations of their competence in English. The MSC scale also consisted of five items with wording identical with the ESC items described previously, but students were asked to evaluate their competence in mathematics ($\alpha = .97$; e.g., I am good at mathematics, etc.). The SDQ-II items were rated on a 6-point response scale ranging from 1 (*false*) to 6 (*true*).

Academic achievement. To construct indicators of individual student achievement and context-average achievement, we used students' English and mathematics Primary School Leaving Examination scores obtained by the participants at the end of their Primary-6, that is, prior to their start of secondary schools the following academic year. These PSLE scores were the key determinant of student placements into different ability streams. The possible range of PSLE raw scores is between 0 and 200 for English and between 0 and 100 for math. English and math PSLE scores made available to students

(and used in this study) are T-standardized scores. Thus, the performance of a nationwide cohort of all students taking the examination in a given year is taken into account in the calculation of the reported domain-specific PSLE scores of individual students.

The use of the domain-specific PSLE scores optimizes the degree of comparability of the achievement indicators across individual students, grades, streams, and schools, which is one of the methodological prerequisites in investigating the BFLPE (Marsh et al., 2008).

In the present study, the English and math PSLE scores were self-reported by the participants (i.e., they were asked, "What is your PSLE score for English and math?"). This method generated a substantial amount of responses (N = 4,045 for English and N = 4,040 for math). Three schools also provided a complete set of PSLE scores of their 1,425 students. The correlations between students' self-reported PSLE scores and those provided by the schools was substantial, r = .93 ($p < .001$), suggesting the high reliability of students' self-reported scores. This was not surprising given the fact that the PSLE scores are extremely crucial to streaming and, therefore, students were most likely to accurately remember their scores. Furthermore, a recent study suggests that researchers can assume the validity of students' self-reported grades as they are not subject to systematic bias (Dickhäuser & Plenter, 2005). To deal with missing data, we implemented the Expectation Maximization (EM) algorithm as the most widely recommended approach to imputation for missing data (Graham & Hoffer, 2000).

Stream membership. Two dummy variables representing high-ability (HA) and low-ability (LA) streams were created to examine BFLPEs at both ends of the ability continuum using stream membership. HA and LA are dichotomous variables with students in the group used to name the variable coded 1 and the remaining students coded 0 as a reference group (e.g., LA is a variable with students in the low-ability group coded 1 and students in the other two streams coded 0).

Gender and ethnicity. Students also supplied information about their gender and ethnicity. For analysis purposes, female students were coded 1 and male students were coded 0. In this study, students were asked to indicate their ethnicity, as stated in their identity card, by choosing one of the options provided (i.e., Chinese, Malay, Indian, Others). Students who chose "Others" were asked to specify their ethnicity. As there were minority groups (e.g., Filipino, Vietnamese) that were represented by only one student in a class/stream, the same-ethnic referent may not be available for some students. In alignment with Singapore's ethnic proportion and prior work (e.g., Aboud, 1976; Meisel & Blumberg, 1990) showing the use of peers of the same ethnicity status (majority or minority) as referents in comparing achievement, the dichotomy of "ethnic majority" (comprising Chinese students, coded 1) and "ethnic minority" (comprising non-Chinese students, coded 0) was used to examine the BFLPE in the context of ethnicity.

Procedure

A pilot study was first carried out by administering a paper-and-pencil survey to 308 students from one secondary school in Singapore. The purpose of the pilot study was to ensure that students at all grades (i.e., year groups) and all streams understood the items in the survey, to record the time taken to complete the survey, and to monitor all other possible issues that might be encountered in the main study. Students in general did not find any difficulty in responding to the survey. On the basis of the pilot study a few items were modified for better clarity of expression. However, no item used in this study was modified from its original version.

In the main study, the survey was administered in intact groups by the teachers or by trained research assistants, as deemed appropriate by the school principals. The survey was conducted in the second half of the academic year so that students were aware of their academic standing relative to other students. Participants were first briefed that the purpose of the survey was to understand their school motivation and learning. To ascertain participants' genuine answers, it was emphasized that their responses were confidential, would not affect their school grades, and would be analyzed collectively and not individually. Participants were also told that there were no right or wrong answers to any of the questions and that honest responding was of great importance in the study. It took around 45 minutes for the participants to complete the survey. Human Ethics Research Clearance was obtained.

Statistical Analysis

In research that involves students from a large number of clusters, it is inappropriate to pool responses of individual students without regard to the groups or contexts (e.g., classroom, stream, school) to which students belong unless it can be demonstrated that each of the groups does not differ systematically from each other (see Goldstein, 1995; O'Connell & McCoach, 2008; Raudenbush & Bryk, 2002). Because of the hierarchical nature of the data, conducting a single-level analysis, particularly one that juxtaposes the effects of class, stream, and school, may violate assumptions of independence and lead to associated problems such as aggregation bias, ecological fallacy, heterogeneity of regression, and spurious significant results (e.g., Raudenbush & Bryk, 2002). Therefore, consistent with methodological recommendations by Marsh et al. (2008) in testing the BFLPE and with other recent BFLPE studies (e.g., Huguet et al., 2009; Marsh et al., 2007), multilevel modeling was performed using MLwiN 2.21 (Rasbash, Steele, Browne, & Goldstein, 2009). A detailed presentation of multilevel modeling is beyond the scope of the present investigation and is available elsewhere (e.g., Raudenbush & Bryk, 2002).

In this investigation, we considered four-level analyses in which students (Level 1) are nested within classes (Level 2), which are nested within streams (Level 3), which are nested within schools (Level 4). We conducted two sets of multilevel regression analyses that aimed at addressing the main purposes of the study: (1) to examine the BFLPE as a function of contextual predictors varying in their degree of locality/generality (class-average achievement, stream-average achievement, and school-average achievement) and (2) to examine the BFLPE in the contexts of gender and ethnicity. To render multi-level regression coefficients in each of the models comparable and to enhance the interpretability of findings, we standardized (z-scored) all the continuous variables (e.g., English self-concept, English achievement scores) to have $M =$ 0 and $SD = 1$ across the entire sample (see Aiken & West, 1991; Raudenbush & Bryk, 2002) and centered the dichotomous dummy variables (i.e., gender and ethnicity) using the corresponding mean of such variables (i.e., in this study, the means were .48 for gender and .69 for ethnicity; see Hox, 2010, p. 61, for this procedure).

BFLPE. The first set of analyses was a test of the BFLPE as a result of class, stream, and school contextual effects. Separate analyses were conducted for English and math domains. These domain-specific academic self-concepts (ESC or MSC) were the outcome variable, and the corresponding domain-specific measures of individual achievement (both linear and quadratic), school-average achievement, stream-average achievement, and class-average achievement were the predictor variables, which were successively included in a series of a priori nested models to address our substantive research goal (see Models 1–8, Table 3). The inclusion of the nonlinear (quadratic) component of student ability in the models allows us to examine the extent to which relations between achievement and academic self-concept are monotonic and similar for students of lower ability, middle ability, and higher ability (Marsh & Rowe, 1996). The quadratic achievement variable was created by squaring the standardized linear ability variable. To ensure that variables were kept in the same metric, neither the quadratic component of achievement nor the context-average achievements were re-standardized. The three contextual variables, school-average achievement, stream-average achievement, and class-average achievement, were calculated by averaging the student English or math PSLE scores separately for each school, for each stream within each grade and each school, and for each class, respectively. There was a total of 68 combinations of stream, school, and grade and a total of 136 classes (i.e., slightly more than two classes in each of the 68 combinations of stream, grade, and school; see Table 1).

For purposes of the present analyses, the effects of predictors and a constant on academic self-concept were estimated as fixed effects, and variation in the student (Level 1), class (Level 2), stream (Level 3), and school (Level 4) intercepts were estimated as random effects. The random effects demonstrated the degree of the variation that existed in ESC or MSC intercepts

342

that vary from student to student (Level 1), from class to class (Level 2), from stream to stream (Level 3), and from school to school (Level 4). To test if stream membership (i.e., being in a particular stream) is predictive of academic self-concept in the directions that are supportive of the presence of the BFLPE (i.e., students in the high-ability stream have lower academic self-concepts and students in the low-ability stream have higher academic self-concepts), we entered the dummy variables HA and LA in separate models (see Models 9 and 10, Table 3).

Gender-based and ethnicity-based frames of reference. The second set of analyses tested the BFLPE in relation to gender- and ethnicity-based frames of reference. For purposes of these analyses, domain-specific measures of gender-stream-average achievement and ethnicity-stream-average achievement indicators were calculated by averaging English or math PSLE scores of students who are of the same gender (or the same ethnicity status) who are within the same school, grade, and stream. Hence, gender- or ethnicity-stream-average achievement is a more local contextual predictor than stream-average achievement. In a similar way, we calculated domain-specific measures of gender-class-average achievement and ethnicity-class-average achievement by averaging English or math PSLE scores of students who are of the same gender (or the same ethnicity status) who are within the same class. Thus, gender- or ethnicity-class-average achievement is a more local contextual predictor than class-average achievement and even the most local contextual predictor relative to other contextual predictors considered in this study. In these analyses, the domain-specific measures of gender-average achievement (or ethnicity-average achievement) and other predictors (i.e., a dummy variable representing male/female or Chinese/non-Chinese, both linear and quadratic student achievement, stream-average achievement or class-average achievement) were entered into the models successively as predictors of ESC and MSC. These predictors were estimated as fixed effects and the intercepts at each of the four levels were estimated as random effects. Given the high correlation between domain-specific measures of stream- and class-average achievement, BFLPEs in the context of gender and ethnicity within stream and class are considered in separate models (see Table 4).

Results

Preliminary Analyses

We first performed a series of baseline or unconditional models (i.e., models with no predictors) to estimate the proportion (%) of variance in achievement and self-concept explained by differences between students (Level 1), classes (Level 2), streams (Level 3), and schools (Level 4). As shown in Table 2, in terms of English achievement, the random effects associated with Levels 1, 2, and 3 were significant and accounted for around 57%,

343

Table 2
Descriptive Statistics for the Variables in the Study

	Multilevel Proportion of Variance (in %)				High-Ability Stream		Middle-Ability Stream		Low-Ability Stream	
	Student	Class	Stream	School	M	SD	M	SD	M	SD
English achievement	57**	3**	39**	1	72.09	7.70	60.63	9.17	57.32	10.54
Math achievement	58**	9**	32**	1	71.45	10.51	54.13	13.14	52.28	14.67
English self-concept	93**	4*	2	1	3.46	1.26	3.54	1.28	3.87	1.25
Math self-concept	90**	8**	1	1	3.58	1.47	3.46	1.52	3.33	1.55

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

3%, and 39%, respectively, of student-to-student, class-to-class, and stream-to-stream differences. Similarly, the baseline model for math achievement indicated that 58%, 9%, and 32%, respectively, of the variance in math achievement was attributable to student-to-student, class-to-class, and stream-to-stream differences. The random effect associated with Level 4 was not significant and suggested that school differences accounted for only 1% of the variation in English and math achievement scores. In terms of self-concept, the random effects associated with Levels 1 (student) and 2 (class) were significant and suggested that 93% and 4%, respectively, of the variance in English self-concept and 90% and 8%, respectively, of the variance in math self-concept were explained by student-to-student and class-to-class differences. The corresponding effects associated with Levels 3 and 4 were not significant and indicated that stream and school differences explained around 2%, or lower, of the variance in English and math self-concepts. Collectively, these preliminary findings reflect the potential operation of the BFLPEs as a function of stream: Although students in different streams, on average, varied significantly in their English and math achievement scores, there is little variation across streams in their English and math self-concepts. That is, due to social comparison with peers who are academically strong, academic self-concepts of students in the high-ability stream are dampened and do not correspond with their relatively stronger academic achievement. In contrast, due to the use of academically weak students as a frame of reference, academic self-concepts of students in the low-ability stream are inflated and do not match with their own relatively weaker academic achievement.

Big-Fish-Little-Pond Effect

To test the BFLPE on English and math self-concepts, we conducted a series of 10 multilevel regression analyses in which predictors were entered sequentially (see Table 3). In Model 1, we entered student linear and quadratic achievement as predictors of academic self-concept. The effects of the linear component of student achievement on academic self-concept were significantly positive ($B = .51$, $p < .001$, for ESC and $B = .54$, $p < .001$, for MSC). These findings indicated that students whose English and math achievement levels were one standard deviation (*SD*) above the mean had ESCs and MSCs that were around .51 *SD* and .54 *SD*, respectively, above the average self-concepts in these two corresponding subjects. The effects of the quadratic component of student achievement on their ESCs and MSCs were also significantly positive, albeit much smaller than those of the linear component ($B = .06$, $p < .001$, for ESC and $B = .04$, $p < .001$, for MSC), showing that the effects of student achievement on academic self-concept were stronger for higher achieving students and weaker

Table 3

Models Assessing the Effects of School-, Stream-, Class-, and Student-Level Achievement on Academic Self-Concepts

Effect	Models Predicting English Self-Concept									
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Fixed effects										
Constant	.06	.04	−.07	−.07	−.06	−.07	−.07	−.07	−.05	−.02
English linear achievement	.51***	.51***	.54***	.54***	.54***	.54***	.54***	.54***	.54***	.51***
English quadratic achievement	.06***	.06***	.07***	.08***	.07***	.07***	.07***	.07***	.07***	.06***
School-average achievement		−.29			.19	.16		.18		
Stream-average achievement			−.62***	−.61***	−.63***	−.62***				
Class-average achievement							−.31*	−.32*		
High-ability (HA) stream							−.32*	−.32*	−.37***	
Low-ability (LA) stream										.27***
Random effects										
Level 4: School intercept	.00	.00	.01	.01	.01	.01	.01	.01	.01	.00
Level 3: Stream intercept	.16***	.15***	.00	.00	.00	.00	.01	.01	.01	.07***
Level 2: Class intercept	.04***	.04***	.05***	.04***	.05***	.05***	.04	.04***	.04***	.03**
Level 1: Student intercept	.81***	.81***	.81***	.81***	.81***	.81***	.81***	.81***	.81***	.81***
−2*log-likelihood	11956	11954	11872	11871	11870	11870	11867	11866	11879	11910
Fixed effects										

(continued)

Table 3 **(continued)**

Effect		Models Predicting English Self-Concept								
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Constant	.03	.00	−.04	−.04	−.04	−.04	−.04	−.04	−.04	.01
Math linear achievement	.54***	.54***	.57***	.56***	.57***	.56***	.56***	.56***	.57***	.54***
Math quadratic achievement	.04***	.04***	.04***	.04***	.04***	.04***	.04***	.04***	.05***	.04***
School-average achievement					.02	−.07		.02		
Stream-average achievement		−.30**	−.40***	−.31***	−.41***	−.30***	−.59***	−.60***		
Class-average achievement							.20	.20		
High-ability stream									−.25***	
Low-ability stream										.09**
Random effects										
Level 4: School intercept	.00	.00	.00	.00	.00	.00	.00	.00	.01	.00
Level 3: Stream intercept	.06**	.05**	.01	.01	.01	.01	.01	.01	.00	.05**
Level 2: Class intercept	.05***	.06***	.05***	.07***	.05***	.07***	.05***	.05***	.05***	.05***
Level 1: Student intercept	.75***	.75***	.75***	.75***	.75***	.75***	.75***	.75***	.75***	.75***
−2*log-likelihood	11,609	11,603	11,563	11,584	11,563	11,583	11,560	11,559	11,554	11,602

Note. HA and LA are dichotomous variables with students in the stream used to name the variable coded 1 and students in the other two streams serving as a reference group and coded 0.
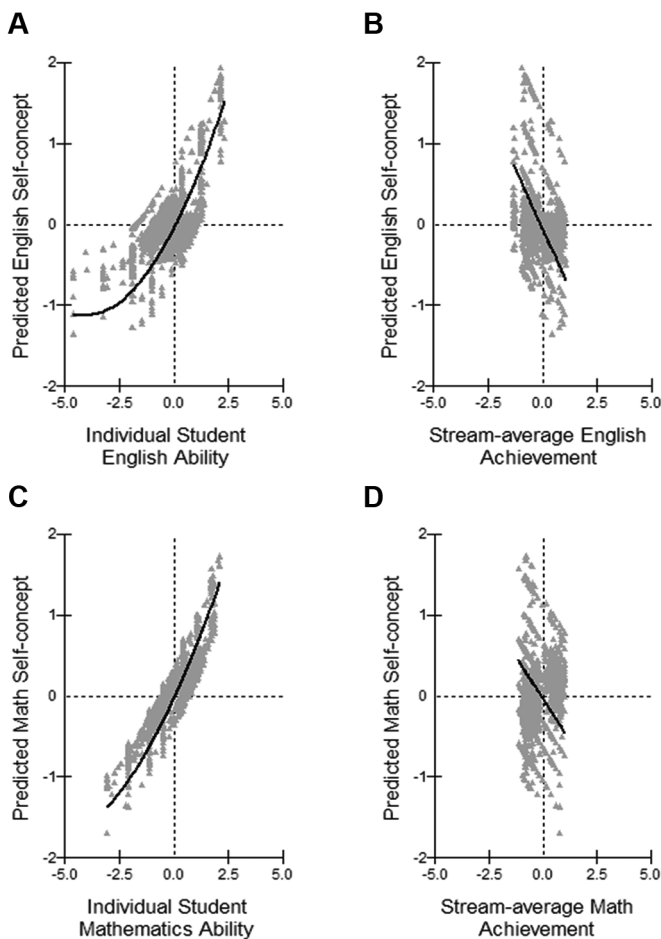*p < .05. **p < .01. ***p < .001.

347

Figure 2. **Relations between academic domain-specific measures of academic self-concept, student achievement (or ability), and stream-average achievement.** *Note*: Scatter plots of 4,461 points reflecting the relations between academic self-concept and student achievement in English (A) and math (C) and stream-average achievement in English (B) and math (D). Each grey point represents an individual student. The solid dark line represents the regression equation for each relationship across all students. Self-concept and achievement scores were standardized (*M* = 0, *SD* = 1) across all students and stream-average achievement was based on the average of the standardized achievement scores so that it is in the same metric as individual student achievement.

for lower achieving students. Figures 2A and 2C show domain-specific relations between individual student achievement and academic self-concept.

In Model 2, we added school-average achievement as a contextual predictor to examine BFLPEs as a function of the average achievement of students who are in the same school. This was a test that has typically been done in BFLPE research (e.g., Marsh, 1987). As can be seen in Table 3, the effects of domain-specific measures of school-average achievement were of a similar size. These effects, however, were significantly negative on MSC ($B = -.30$, $p < .01$) but marginally significant on ESC ($B = -.29$, $p < .10$), suggesting that students in a school with a higher average achievement tended to have lower MSCs and, to a lesser extent, ESCs. The size of these effects is of a similar range to that of school-average achievement reported in previous studies (e.g., Marsh et al., 2000).

In Model 3, we used stream-average achievement as a contextual predictor. In support of our prediction, the effects of stream-average achievement on ESC and MSC were significantly negative ($B = -.62$, $p < .001$, for ESC and $B = -.40$, $p < .001$, for MSC), showing the BFLPE as a consequence of average achievement of all students who were within the same stream. Specifically, students in streams with average achievement levels in English and math that were one *SD* above the mean had ESCs and MSCs that were .62 *SD* and .40 *SD*, respectively, below the average self-concepts in each of the two corresponding subjects. Figures 2B and 2D depict graphical representations of the domain-specific relations between stream-average achievement and academic self-concept.

In Model 4, we used class-average achievement as a contextual predictor. The results showed that the effect of class-average achievement on ESC and MSC were also significantly negative ($B = -.61$, $p < .001$, for ESC and $B = -.31$, $p < .001$, for MSC), indicating the BFLPE as a function of average achievement of students within each class.

To test the local dominance effect prediction, we examined the relative salience of school-, stream-, and class-average achievement by juxtaposing their effects in a series of nested models (see Models 5–8). In Model 5, we used both school-average achievement and stream-average achievement as contextual predictors of academic self-concept. The results showed that even after controlling for school-average achievement, the effects of stream-average achievement remained significant and large ($B = -.63$, $p < .001$, for ESC and $B = -.41$, $p < .001$, for MSC). The effect of school-average achievement on MSC, which was significant when considered as an independent predictor, became nonsignificant and substantially decreased from $B = -.30$, $p < .01$, to $B = -.02$, *ns* (see Models 2 and 5 at the lower part of Table 3). In support of the local dominance effect, these findings suggest that in the school context in which students are placed in different ability streams based on their prior achievement, the average achievement of

students within each stream is more predictive of their academic self-concepts than that within the whole school.

In Model 6, we entered both school-average achievement and class-average achievement as contextual predictors of academic self-concepts. The results indicated that even after controlling for school-average achievement, the effects of class-average achievement remained significant and large ($B = -.62$, $p < .001$, for ESC and $B = -.30$, $p < .001$, for MSC). In this model, the significant effect on MSC of school-average achievement when considered separately in Model 2 became nonsignificant and substantially decreased from $B = -.30$, $p < .01$, to $B = -.07$, ns. Also supportive of the local dominance effect hypothesis, these findings suggest that in the school context in which there are substantial class-to-class differences in average achievement (as a consequence of streaming that assigns students to different classes based on their prior achievement), the effect of average achievement of students within class is more predictive of the students' academic self-concepts than that within the whole school.

In Model 7, we used both stream-average achievement and class-average achievement as contextual predictors of academic-self-concepts. For MSC, although the effect of stream-average achievement remained significant and large ($B = -.59$, $p < .001$), the effect of class-average achievement, which was significant when considered as a separate predictor in Model 4 ($B = -.31$, $p < .001$), became nonsignificant ($B = .20$, ns) in Model 7. For ESC, the effects of both stream-average achievement ($B = -.31$, $p < .05$) and class-average achievement ($B = -.32$, $p < .05$) remained significant. To examine if the effects of stream-average achievement and class-average achievement on ESC were of a similar magnitude, we constrained the B parameters of these two contextual predictors to be equal. The difference in likelihood ratio (LHR) between the unconstrained and constrained models was not significantly different (ΔLHR $= .001$, $\Delta df = 1$, ns), suggesting that the effects of stream-average achievement and class-average achievement on ESC were not statistically different. It should be noted that given that our data were derived from an educational context in which placement of students to different classes and streams with each school was based on their PSLE scores, the achievement indicator used in this study, the two contextual predictors were highly correlated ($r = .97$, $p < .001$, for English and $r = .94$, $p < .001$, for math). Hence, in recognition that class is a more local context than stream—and, hence, class was expected to exert a stronger effect than stream in students' self-evaluations of their prior achievement—this set of findings, collectively, provide strong evidence for the salience of stream over the class context (or at least as salient as the class context) in forming student academic self-concepts.

To juxtapose the relative effects of school-, stream-, and class-average achievement on academic self-concepts, the three contextual predictors were entered simultaneously in Model 8. For MSC, the effect of stream-average

achievement remained significant and large ($B = -.60$, $p < .001$). Although the effects of school-average achievement and class-average achievement on MSC were significant when considered as separate predictors (see Models 2 and 4, respectively), their effects became nonsignificant ($B = .02$, *ns*, for school-average achievement; $B = .20$, *ns*, for class-average achievement) after controlling for the effect of stream-average achievement. For ESC, the effects of stream-average achievement ($B = -.32$, $p < .05$) and class-average achievement ($B = -.32$, $p < .05$) remained significant and of similar magnitude, whereas the effect of school-average achievement was not significant ($B = .18$, *ns*).

BFLPE as a Function of Stream Membership

To elicit further evidence of the operation of BFLPEs at both ends of the ability continuum (i.e., the high-ability and low-ability streams), we replaced context-average achievements with the centered stream membership dummy variables (HA or LA). Specifically, we entered HA as a dummy predictor in Model 9 such that students in the two lower ability streams served as a reference group and LA as a dummy predictor in Model 10 in which students in the two higher ability streams served as a reference group. In support of the BFLPE prediction, the regression coefficient for HA in Model 9 indicated that, controlling for prior academic achievement, ESCs of students in the high-ability stream were .37 *SD* lower than those of students in the two lower ability streams. Conversely, also in support of the BFLPE prediction, the regression coefficient for LA in Model 10 showed that given the same prior academic achievement, ESCs of students in the low-ability stream were .27 *SD* higher than those of students in the two higher ability streams.

For MSC, the regression coefficient for HA in Model 9 indicated that given the same prior math achievement, MSCs of students in the high-ability stream were .25 *SD* lower than those of students in the two lower ability streams. In contrast, the regression coefficient for LA in Model 10 showed that, given the same prior math achievement, MSCs of students in the low-ability stream were .09 *SD* higher than those of students in the two higher ability streams. Taken together, these findings provide evidence of the BFLPE as a function of stream membership, that is, students in the high-ability stream have lower academic self-concepts relatively to their peers in the lower ability streams, whereas students in the low-ability stream have higher academic self-concepts relatively to their peers in the higher ability streams.

In sum, the aforementioned findings provide evidence of the salience of stream-average achievement on academic self-concepts. That is, when the effects of the three contextual predictors were juxtaposed, stream-average achievement came out as the strongest negative predictor of MSC and

exerted a negative effect on ESC that was comparable with that of class-average achievement (although class is a more local setting than stream). On the whole, these findings demonstrate the salience of stream and class settings in forming student academic self-concepts. Hence, subsequent analyses to clarify the effects of gender- and ethnicity-average achievement as alternative frames of reference that students might use in evaluating their achievement focused on the stream and class contexts.

Gender-Based and Ethnicity-Based Frames of Reference in BFLPE

Gender-referenced BFLPE. The upper part of Table 4 presents the BFLPE with a specific reference to gender. Model 1 showed that the girls in our sample, on average, were lower than the boys in their ESCs ($B = -.10$, $p < .01$) and even more so in MSCs ($B = -.29$, $p < .001$). In Model 2 we included student linear and quadratic achievement as predictors. The results show that when academic achievement was controlled, the girls remained significantly lower than the boys in their ESCs ($B = -.14$, $p < .001$) and MSCs ($B = -.15$, $p < .001$). The finding that the girls were lower than the boys in MSCs may not be surprising because their PSLE math scores were also lower, albeit marginal, than those of the boys ($M_{\text{Girls}} = 61.19$, $M_{\text{Boys}} = 63.32$), $t(4,425) = -4.68$, $p < .001$. However, the finding that the girls were lower than the boys in their ESCs is interesting because they were significantly higher than the boys in English PSLE scores ($M_{\text{Girls}} = 66.98$, $M_{\text{Boys}} = 64.61$), $t(4,425) = 7.26$, $p < .001$. These results suggest that regardless of their prior achievement, the girls in our sample were more modest than the boys in evaluating their academic abilities (see Kling, Hyde, Showers, & Buswell, 1999, for a similar pattern of findings in Western settings).

In Model 3 we included gender-stream-average achievement (i.e., average achievement of same-sex students who are of the same school, the same stream, the same grade) as an additional predictor. The results showed that gender-stream-average achievement was a significantly negative predictor of academic self-concept ($B = -.56$, $p < .001$, for ESC; $B = -.28$, $p < .001$, for MSC). However, these findings should not be interpreted as evidence that students compared their academic achievements with the average achievement of same-sex students in their respective streams. Because gender-stream-average achievement is confounded with stream-average achievement, the conclusive support for the adoption of gender-based frame of reference leading to the BFLPE demanded that gender-stream-average achievement remains to be a statistically significant negative predictor of academic self-concept even after the effect of stream-average achievement was controlled. Hence, in Model 4 we included stream-average achievement as an additional predictor. The results showed that the effects of gender-stream-average achievement on academic self-concept were washed out by those of stream-average achievement such that the regression weights

Table 4

Models Assessing Gender- and Ethnicity-Based Frames of Reference in the Big-Fish-Little-Pond Effect

Gender-Referenced Effect	Models Predicting English Self-Concept						Models Predicting Math Self-Concept					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects												
Constant	.01	.06	−.07	−.07	−.05	−.07	−.02	.02	−.03	−.04	−.02	−.04
Female students	−.10**	−.14***	−.10***	−.13***	−.12***	−.13***	−.29***	−.15***	−.28***	−.10**	−.27***	−.14***
Student linear achievement		.51***	.54***	.54***	.55***	.55***		.52***	.54***	.55***	.54***	.54***
Student quadratic achievement		.06***	.07***	.07***	.07***	.07***		.04***	.04***	.04***	.04***	.04***
Gender-stream-average achievement			−.56***	−.14					−.28***	.12		
Stream-average achievement				−.48***						−.49***		
Gender-class-average achievement					−.50***	−.09					−.21***	−.01
Class-average achievement						−.51***						−.27**
Random effects												
Level 4: School intercept	.01	.00	.01	.01	.01	.01	.00	.00	.00	.00	.00	.00
Level 3: Stream intercept	.02*	.16***	.00	.01	.02	.00	.01	.01	.01	.01	.01	.01
Level 2: Class intercept	.04***	.04***	.05***	.05***	.04***	.05***	.09***	.06***	.05***	.05***	.07***	.07***
Level 1: Student intercept	.95***	.81***	.81***	.81***	.81***	.81***	.89***	.75***	.75***	.75***	.75***	.75***
−2*log-likelihood	12,588	11,933	11,865	11,848	11,876	11,849	12,287	11,579	11,555	11,537	11,565	11,557

(continued)

353

354

Table 4 (continued)

Gender-Referenced Effect	Models Predicting English Self-Concept						Models Predicting Math Self-Concept					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects												
Constant	-.02	.03	-.08*	-.08*	.07*	-.08*	-.01	.03	-.03	-.05	-.01	-.04
Chinese students	-.40***	-.38***	-.43***	-.35***	-.43***	-.35***	.13**	-.03	.08**	-.06	.03	-.02
Student linear achievement		.49***	.52***	.52***	.52***	.53***		.54***	.56***	.57***	.56***	.56***
Student quadratic achievement		.07***	.07***	.07***	.07***	.07***		.04***	.04***	.05***	.04***	.04***
Ethnicity-stream average achievement			-.49***	.01					-.33***	.15		
Stream-average achievement				-.55***						-.54***		
Ethnicity-class-average achievement					-.43***	.02					-.20***	.12
Class-average achievement						-.55***						-.42***
Random effects												
Level 4: School intercept	.01	.00	.01	.01	.01	.01	.00	.00	.00	.00	.00	.00
Level 3: Stream intercept	.01	.12***	.01	.00	.02	.01	.00	.06**	.01	.01	.02	.01
Level 2: Class intercept	.03**	.03***	.04***	.04***	.03***	.03***	.09***	.05**	.05***	.05***	.06***	.07***
Level 1: Student intercept	.92***	.79***	.80***	.79***	.79***	.79***	.90***	.75***	.75***	.75***	.75***	.75***
−2*log-likelihood	12,417	11,827	11,767	11,750	11,776	11,748	12,367	11,609	11,581	11,561	11,599	11,582

Note. In these analyses, female and Chinese students were coded 1 whereas male and non-Chinese students—as reference groups—were coded 0.
*p < .05. **p < .01. ***p < .001.

of gender-average achievement became nonsignificant. Specifically, the regression weight of gender-stream-average achievement on ESC decreased from $B = -.56$, $p < .001$, in Model 3 to $B = -.14$, *ns*, in Model 4, and the regression weight of gender-stream-average achievement on MSC decreased from $B = -.28$, $p < .001$, in Model 3 to $B = -.12$, *ns*, in Model 4. Consistent with the earlier finding, Model 4 shows that stream-average achievement was a significantly negative predictor of ESC ($B = -.48$, $p < .001$) and MSC ($B = -.49$, $p < .001$).

Similar findings were also found with gender-class-average achievement (i.e., average achievement of same-sex students within the same class). Model 5 showed that gender-class-average achievement was a significantly negative predictor of academic self-concept ($B = -.50$, $p < .001$, for ESC and $B = -.21$, $p < .001$, for MSC). When class-average achievement was entered in Model 6, however, gender-class-average achievement became nonsignificant. The regression weight of gender-class-average achievement on ESC decreased from $B = -.50$, $p < .001$, in Model 5 to $B = -.09$, *ns*, in Model 6, and the regression weight of gender-class-average achievement on MSC decreased from $B = -.21$, $p < .001$, in Model 5 to $B = -.01$, *ns*, in Model 6. In Model 6, class-average achievement was a significantly negative predictor of ESC ($B = -.51$, $p < .001$) and MSC ($B = -.27$, $p < .01$). Taken together, these findings provide no evidence that students used the average achievement of same-sex peers in the same class or stream as a frame of reference in evaluating their academic abilities.

Ethnicity-referenced BFLPE. The lower part of Table 4 presents ethnicity-referenced BFLPE findings. Model 1 showed that the Chinese students in our sample were lower than the non-Chinese students on ESC ($B = -.40$, $p < .001$) but the reverse was true on MSC ($B = .13$, $p < .001$). However, when individual student achievement in each academic domain was controlled, while the Chinese students remained significantly lower than the non-Chinese students on ESC ($B = -.38$, $p < .001$), this difference was no longer apparent on MSC ($B = -.03$, *ns*). The finding that Chinese students were lower than the non-Chinese students on ESC was interesting given that the Chinese students were significantly higher than the non-Chinese students on their English PSLE scores ($M_{Chinese} = 66.47$, $M_{non-Chinese} = 63.89$), $t(4,459) = 7.40$, $p < .001$. Similarly, the finding that the Chinese students were only slightly higher than the non-Chinese students on their MSCs—the effect that was then eliminated when student math achievement was controlled across the two groups—was intriguing given that the Chinese students were substantially higher than the non-Chinese students on their math PSLE scores ($M_{Chinese} = 65.40$, $M_{non-Chinese} = 55.61$), $t(4,459) = 20.93$, $p < .001$. This set of findings probably reflects a modesty or self-effacement inclination among the Chinese in general (Bond, 1991).

When ethnicity-stream-average achievement (i.e., average achievement of students in the same stream, grade, and school who are of the same

ethnicity status—majority or minority) was included in Model 3, we found that ethnicity-stream-average achievement was a significantly negative predictor of academic self-concept ($B = -.49$, $p < .001$, for ESC; $B = -.33$, $p < .001$, for MSC). However, the effect of ethnicity-stream-average achievement became nonsignificant when we included stream-average achievement in Model 4. This was true for both ESC (the regression weight decreased from $B = -.49$, $p < .001$, in Model 3 to $B = .01$, *ns*, in Model 4) and MSC (the regression weight decreased from $B = -.33$, $p < .001$, in Model 3 to $B = .15$, *ns*, in Model 4). Aligned with the finding reported earlier, stream-average achievement was a significantly negative predictor of ESC ($B = -.55$, $p < .001$) and MSC ($B = -.54$, $p < .001$).

A similar pattern of results was also found with ethnicity-class-average achievement (i.e., average achievement of same-ethnicity students within the same class). Model 5 showed that ethnicity-class-average achievement was a significantly negative predictor of academic self-concept ($B = -.43$, $p < .001$, for ESC; $B = -.20$, $p < .001$, for MSC). When class-average achievement was entered in Model 6, however, the effect of ethnicity-class-average achievement was eliminated and became nonsignificant. The regression weight of ethnicity-class-average achievement on ESC decreased from $B = -.43$, $p < .001$, in Model 5 to $B = .02$, *ns*, in Model 6, and the regression weight of ethnicity-class-average achievement on MSC decreased from $B = -.20$, $p < .001$, in Model 5 to $B = .12$, *ns*, in Model 6. Class-average achievement remained to be a significantly negative predictor of ESC ($B = -.55$, $p < .001$) and MSC ($B = -.42$, $p < .001$). Collectively, these findings show no evidence of the use of students with the same ethnicity status (majority or minority) who are within the same class or stream as a frame of reference in social comparisons leading to the BFLPE.[2]

Discussion

The present study sought to examine the big-fish-little-pond effect within schools that implement streaming at low, middle, and high levels of ability. We tested this issue in an educational context (Singapore) that implements systemwide ability-streaming within schools in a consistent way based on a common, nationwide battery of achievement tests completed by all students at the end of primary school. Consistent with the BFLPE model (Marsh, 1987; Marsh et al., 2008), we found that individual students' prior achievement had positive effects on their self-concepts in English and math. After controlling for these positive effects, students in the high-ability stream had lower English and math self-concepts than students in the lower-ability stream, while students in the low-ability stream had higher ESCs and MSCs than students in the higher-ability stream.

356

Furthermore, in relation to its effects on formation of students' academic self-concept, we found that stream-average achievement played a more salient role than school-average achievement in both ESC and MSC. Although class is a more local setting than stream and, therefore, is expected to have a more salient effect on formation of students' academic self-concepts (see Rogers et al., 1978; Zell & Alicke, 2009, 2010), our findings showed that the effect of stream-average achievement was stronger than class-average achievement on MSCs and was relatively comparable with class-average achievement on ESCs.

We also sought to examine the extent to which students' formation of their academic self-concepts depends on other same-gender or same-ethnicity students in their stream or class. However, after accounting for stream- and class-average achievement (i.e., average achievement of *all* other students in their stream or class, respectively), the effects of gender- and ethnicity-average achievement became statistically nonsignificant, suggesting no evidence that students used gender-specific and ethnicity-specific frames of reference to form their MSCs or ESCs.

Specific Contribution to the BFLPE Research Program: Streaming Effects

One of the most important contributions of the present investigation is the demonstration that the effects of stream-average achievement on academic self-concept were more salient than those of school-average achievement. This shows that in an educational context that implements a nationwide system for placing students into different ability streams within each school on the basis of their prior achievement, there was much more variability at the stream level than at the school level. As a consequence, when student academic achievement was statistically controlled (considered to be equal), the negative effects of stream-average achievement on students' academic self-concepts were larger than those of the school-average achievement.

We also found that stream-average achievement (as a more general context) was a stronger predictor than class-average achievement (as a more local context) of students' MSCs and it was as strong as class-average achievement in predicting students' ESCs. These results provide further evidence of the salience of stream as a context that plays an influential role in students' academic self-concepts (particularly in recognition of the more local nature of class relative to stream and the high correlation between stream-average and class-average achievement).

This is noteworthy from an applied perspective, suggesting that educational practice directed at attenuating the negative BFLPE experienced by students in the high-ability stream must take into account internal structures and grouping within schools as well as the average level of the school itself. In this instance, practice would need to be differentiated by ability group rather than—or in addition to—the whole-school level. From a design perspective,

357

this is apparently one of a very few studies that simultaneously evaluated the frame of reference effects at more than one level (i.e., class level, stream level, and school level). In particular we know of no other study that has examined this issue in a large-scale, naturalistic setting in which there is a relatively consistent basis for assigning students to streams used by all schools and that the assignment in all schools is based on scores on a common set of tests administered to all students before they started the school.

Further evidence for the BFLPE was found when stream membership was used as a predictor in the models. Given equal achievement, students in the high-ability stream had lower academic self-concepts than those in the lower-ability streams and students in the low-ability stream had higher academic self-concepts than those in the higher-ability streams. The present study demonstrated that the BFLPE indeed has effects on academic self-concepts of students at both ends of the continuum and these effects are reversed for high and low tracks (Hattie, 2002; Marsh, 1984). That is, following BFLPE predictions in relation to academic self-concept, placement of students with peers who are of similar abilities has a detrimental effect for high-ability students but a beneficial effect for students with lower abilities. This finding aligns with those showing BFLPEs in special education settings. Marsh, Chessor, Craven, and Roche (1995), for example, found that academic self-concept of gifted and talented students declined over time when they shifted from mixed-ability to academically special programs (based on pre-post score comparison) and in comparison with students matched on academic achievement who continued to attend mixed-ability program. On the other end of the spectrum, Marsh, Tracey, and Craven (2006) showed that preadolescents with intellectual disabilities perceived themselves as less academically competent when they compared themselves with mixed-ability students in regular classes—as a consequence of "mainstreaming"—than when they compared themselves with peers with the same level of achievement in special classes.

BFLPEs Specific to the Contexts of Gender and Ethnicity

To our knowledge, this is the first research investigating the BFLPE in the extent to which frames of reference in relation to academic accomplishments are specific to students of the same gender or the same ethnicity status. Implicit in this investigation is the assumption that students compare their performance with the average achievement of generalized others in the students' respective streams or classes (as posited in the BFLPE model) who share the same characteristics in terms of gender and ethnicity (as two important attributes in social comparisons: see Dijkstra et al., 2008). Our findings have provided no evidence that students used the average achievement of students of the same gender or the same ethnicity as a source of social comparison information leading to BFLPEs. Although gender-average achievement and

358

ethnicity-average achievement had negative effects on students' academic self-concepts (see Models 3 and 5, Table 4), when stream-average achievement (i.e., the average achievement of *all* students in the same stream) or class-average achievement (i.e., the average achievement of *all* students in the same class) was included in the model (Models 4 and 6, Table 4), the effects of gender- and ethnicity-average achievement became statistically nonsignificant. Thus, these findings suggested no evidence that students compared their achievements with the average achievement of other students in their streams who were of the same gender or ethnicity status.

It is also interesting to note that relative to their non-Chinese counterparts, the Chinese students in our sample tended to be more modest in evaluating their abilities. This finding might be attributable to a self-effacing inclination typically preferred by the Chinese in their self-presentations (Bond, 1991; Gudykunst, 2003). Similarly, compared with the boys, the girls in our study seemed to be more modest in judging their abilities. This pattern, however, was different from prior studies with Singaporean students. For example, in a 3-year longitudinal study, Liu, Wang, and Parkins (2005) found a less conclusive pattern of gender differences in academic self-concept—while female seventh graders in the high-ability stream reported higher academic self-concepts than their male counterparts, this pattern was reversed in the low-ability stream. Furthermore, these relatively small differences disappeared when the students were in Grade 9 (see also Liu & Wang, 2005). In our study, we found the girls were lower than the boys in both English and math self-concepts. One tentative explanation to this might be related to gender-role stereotypes within general Asian cultures, and particularly the Chinese tradition, that demand females to be socially responsible and likeable and modest in self-presentation (Crittenden, 1991). The extent to which this is the case, future studies need to measure, and appropriately control for, students' cultural beliefs about gender-role stereotypes in understanding their academic self-concept development.

Interplay of Local Dominance Effect With Variability and Salience of Frame of Reference

We have inferred from our findings that among the contextual frames of reference tested, stream-average achievement had relatively more salient effects on academic self-concepts. This inference is based on the following results. Consistent with predictions from the local dominance effect hypothesis (Zell & Alicke, 2009, 2010), the effects of stream-average achievement (a more local frame of reference) on both domain-specific academic self-concepts were larger than those of school-average achievement (i.e., a more general frame of reference). Following the local dominance effect model in self-evaluations (Zell & Alicke, 2009, 2010), one might also speculate that class might be a more dominant frame of reference than stream and

359

that average achievement of same-gender or same-ethnicity students within the same stream or class would provide an even more local frame of references than stream- or class-average achievement. However, the results showed that the effects of stream- and class-average achievement on ESCs were equally strong and that the effect of stream-average achievement on MSCs was even stronger than that of class-average achievement (although class is a more local context than stream). We also found no evidence to support the local dominance effect in relation to gender and ethnicity-specific frames of reference.

Taken together, the findings lead us to a speculation that the size of the variability of achievements across groups used as a frame of reference (or comparison standard) might be positively associated with its importance in predicting self-concepts. In the present study there was little variance at the school level (i.e., a more general frame of reference) and much more at the stream level (i.e., a more local frame of reference), thus both the variability hypothesis and the local dominance hypothesis lead to the same prediction that was supported by the results. Specifically, school-attributed differences in achievement (1% for both English and math, Table 2) were smaller than stream-related differences in achievement (39% for English and 32% for math), and this might result in the relatively more salient effect of stream-average achievement than school-average achievement on academic self-concepts—a pattern consistent with the local dominance effect prediction. The relatively small school-associated differences in achievement were not unexpected given the national policy of within-school ability streaming based on results of national standardized achievement tests taken by all pupils in the last year of their primary education. It is, however, easy to imagine situations in which students are assigned to classes within a school at random but that schools vary substantially in terms of school-average achievement. Here the variability hypothesis and the local dominance hypothesis would lead to different conclusions and we assume that the school-average achievement effect would dominate the class-average achievement effect (although, even with random assignment, there would be some variation between classes that might have some influence on self-concepts). Thus, there is a need for future research to investigate the present issue in an educational setting with a larger variability of achievement across schools.

Similarly, the class-attributed difference in math achievement (9%) was smaller than the stream-related difference (32%), and in support of a variability hypothesis, we found that the effect of class-average math achievement in predicting MSCs was weaker than that of stream-average achievement (see Model 7 in the lower part of Table 3). A similar interpretation may shed further light on gender reference (and ethnicity reference) effects. If there is no difference between boys and girls (or different ethnic groups) in terms of achievement, then gender-average achievement cannot predict self-concepts—even if

360

students do use a gender-specific frame of reference. In the present investigation, there were gender and ethnicity differences in achievement, but the sizes of these differences were smaller relative to differences between streams. As such, the average achievements of gender groups and ethnicity groups did not have any significant effect beyond the effects of stream- or class-average achievement, suggesting support for the variability hypothesis. Hence, our findings appear to suggest a variability hypothesis such that the most influential frame of reference might be one with the greatest variability between different groups rather than the most local one, but this is clearly a direction for further research.

In English, however, the result provided support for predictions based on both the local dominance effect hypothesis and the variability hypothesis. Although the variation in achievement attributed to class-to-class differences was small (3%) compared to that attributed to stream-to-stream differences (39%), the predictive effect of class-average achievement on ESCs was comparable with that of stream-average achievement (see Model 7 in the upper part of Table 3). This English-related finding leads us to an alternative interpretation. That is, it may be that the relative importance of different frames of reference varies with the salience of a particular domain in different settings and the characteristic/nature of the domain. In a setting like Singapore where ability grouping is such an explicit feature of the educational system and students only take classes with other students from their same ability stream, ability stream is likely to be very salient in relation to academic achievement and academic self-concept. As the ability grouping considered here is based on students' academic achievement, based on which their academic self-concepts are considered, it is reasonable to expect that academic self-concepts are more likely to be affected by the students' stream membership. Further, the fact that we found this finding specific to English also brings to the fore of the importance of the characteristic of a particular domain in considering frame-of-reference effects. It might be that the extent to which students use the achievement of others in the proximal context in forming their self-evaluations may vary across academic domains.

A similar interpretation may shed further light on gender reference effects. In the present investigation, gender-average achievement was shown to have a little effect on academic self-concepts. For illustration purposes, we may look at students' self-concepts in a physical domain. Marsh (1998), for example, has shown that for physical self-concept the frame of reference is apparently determined in relation to other students who are of a similar age and gender. He showed that if a boy and a girl had similar physical abilities in an absolute sense, the girl was likely to have a higher physical self-concept. This follows in that the girl's physical abilities were higher relative to other girls and the boy's physical abilities were lower relative to other boys. The extent that this is the case, we speculate that a gender-specific frame of reference in relation to physical abilities might be more salient

361

than a gender-specific frame of reference in relation to academic achievement, at least in part, because gender differences tend to be larger for physical characteristics than for academic achievement. Testing students' perceptions of the salience of domain-specific abilities in effecting their self-evaluations of such abilities should then also be an avenue for future research.

Educational Implications

As reviewed earlier, positive academic self-concept is an important outcome in itself as well as a facilitator of key short-term (e.g., achievement and effort) and long-term (e.g., educational and occupational aspirations, university attendance) outcomes (see e.g., Guay et al., 2004). However, group-average achievement has reversed effects on academic self-concept such that learning contexts with high average achievement (i.e., contexts comprising high-ability students) are likely to undermine students' academic self-concepts. The present study has exactly found this by showing detrimental effects of class-average and stream-average achievement on high-ability students' academic self-concepts. In this respect, our findings hold a number of important implications for developing strategies to attenuate the detrimental effect of group-average achievement among students in high-ability streams (and classes) and to enhance academic self-concepts of all students in streamed settings in general.

One key strategy, we believe, should be directed at reducing students' engagement in social comparisons. This could be done by de-emphasizing competition, which tends to reward only a relative minority of students and potentially dampens academic self-concepts of most other students who do not perform as well as the most capable minority. To this end, teachers may focus on criterion-based assessments and feedbacks in evaluating students (see Marsh et al., 2008) or encourage students to pursue their personal bests, that is, personalized and self-referenced academic goals that match or exceed their own previous best performance (Liem, Ginns, Martin, Stone, & Herrett, 2012). Aligned with this recommendation, Lüdtke, Köller, Marsh, and Trautwein (2005), for example, have shown that students' academic self-concepts are enhanced when teachers emphasized students' individualized improvement as opposed to normative assessments.

Interventions should also be targeted at guiding students to see their achievement from a more objective perspective by considering criterion- or task-based evaluative standards. In doing so, social comparison processes leading to the detrimental effect of high-ability grouping on students' academic self-concepts may be reduced. Importantly, given the longitudinal and reciprocal effects between self-concept and achievement (Marsh & Martin, 2011), intervention programs that focus on promoting students'

362

academic self-concepts should also aim at developing their learning skills with an aim to bring about changes in their actual achievement. In other words, academic self-concept interventions should target both psychological and behavioral aspects of the learners.

It is also recognized that educational interventions can be implemented at different levels, ranging from the individual student level to the whole school level. Given the bulk of variation in English and math self-concepts was attributed to student-to-student differences, a student-centered intervention is considered to be the most defensible strategy to carry out in enhancing academic self-concepts. Among the contextual factors considered here, however, our findings have shown the salient effect of stream—more so relative to school and class contexts—on the formation of students' academic self-concept. This important finding suggests that at the institutional or systemic level, efforts to enhance academic self-concepts could be efficiently and effectively implemented by designing specific programs for each of the three core ability streams rather than developing programs focusing on the class or school level.

Potential Limitations and Future Directions

This study has provided new insights by juxtaposing different possible frames of reference that students might use in the formation of their academic self-concepts, and we did so in a large-scale applied education setting. There are, however, potential limitations important to consider when interpreting the findings, which provide directions for future research. First, it is important to note that although our sample is large ($N = 4,461$) and the sampling of schools was conducted to optimize representativeness of secondary schools across all educational jurisdictions in Singapore, the students included in the present analysis were drawn from only nine schools. This may have affected between-school variability in the outcome factors examined. Furthermore, the bulk of the achievement scores were provided by students and this might potentially inflate their correlation with academic self-concept scores. Nevertheless, these self-reported achievement indicators shared a considerably substantial amount of variance (approximately 86%) with the actual achievement scores provided by some of the participating schools (see Method)—supporting evidence that the validity of students' self-reported grades are less affected by systematic bias (Dickhäuser & Plenter, 2005). Thus, it is important that further studies draw a larger sample of schools and obtain actual achievement scores from the schools involved to minimize the potential shared method bias.

Second, the present study was conducted with Singaporean students studying in a competitive education system with a nationwide implementation of ability streaming. While this setting is one of the key strengths of our study, it could also mean that the findings may not be generalizable to

363

education systems that do not implement the same policy. Furthermore, the fact that the achievement scores used in this study were derived from a high-stakes national examination that constitutes the sole basis of student placements into different streams—and thus they were perceived to be extremely important by the students—may have affected the associations between achievement and academic self-concept shown in this study. Hence, there is a need to extend the generalizability of our findings to students in other education systems and to use different types of achievement indicators, such as teacher-assigned grades, as predictors of academic self-concepts.

Third, the study has focused on the relative salience of school-, stream-, and class-average achievement in academic self-concepts. As reported previously, controlling for SES-related indices in all the models tested did not affect the size, significance, and direction of the BFLPEs as a function of class, stream, and school (see also Marsh et al., 2008). These robust findings may relate to the fact that the achievement indicators we used were students' PSLE scores, which were the key basis for placement into different streams. However, this does not mean that other contextual and individual factors including those related to parents, teachers, and peers do not play a role in the academic self-concept formation. Ireson and Hallam (2001), for example, found that teachers typically prefer to teach high-ability groups to low-ability groups and they expect students in high-ability groups to be more analytical than those in low-ability groups. This attitude and expectation variation is in turn manifested in instructional methods, material development, enthusiasm, classroom management, and the extent teachers expend effort in their preparation time—all of which are key factors associated with students' academic performance and self-concepts (Rubie-Davies, 2006). Furthermore, there is also evidence that teachers' expectations differ according to the students' ethnicity (Tenenbaum & Ruck, 2007). As such, future research should also include other factors relevant to academic self-concepts to examine if the effects of contextual predictors found here hold true with effects of these other relevant factors controlled.

Fourth, our findings provided no evidence that students used *average achievements* of their same-gender or same-ethnicity class/stream-mates as social referents of their academic self-concepts. However, when asked to nominate an *individual classmate* as a target comparison, prior studies (e.g., Meisel & Blumberg, 1990; Preckel & Brüll, 2008) have shown that students would typically choose a same-sex or same-ethnicity classmate. Future studies therefore need to examine and juxtapose the effects of *individual* and *aggregate* referents as target comparisons in forming self-evaluations of students' achievement. Moreover, Oakes (1985) observed an overrepresentation of ethnic minority students in lower-ability classes. The extent to which that this pattern may interact with students' cultural beliefs and values and teachers' ethnicity and expectations in forming academic self-concepts

of students in tracked settings is also an important issue worth addressing in future research.

Conclusion

The present study has contributed to the big-fish-little-pond effect program of research in the following ways. First, the study has lent further support to the BFLPE model in an ability streaming setting by identifying the salience of stream-average achievement in the formation of students' English and mathematics self-concepts over and above the effects of school-average achievement and class-average achievement (or, at least, equally as strong as class-average achievement). Second, the study has shed light on the potential role of gender and ethnicity as frames of reference in social comparison processes relevant to BFLPEs on academic self-concepts. Third, the study has extended the generalizability of the BFLPE to the Singaporean education context—a unique setting that implements a nationwide practice of within-school ability streaming based on students' scores in a high-stakes national examination at the end of their primary education. Fourth, the study has highlighted the potential interplay of a local dominance effect with variability and/or salience of frames of reference in the formation of students' academic self-concepts. Taken together, findings from this investigation contribute to a better understanding of the formation of academic self-concept, hold implications for educators and students in streamed systems, and provide substantive and methodological directions for future research.

Notes

[1]The overall academic performance obtained at Primary-4 is the basis of placing Primary-5 and Primary-6 pupils into the EM1, EM2, and EM3 streams. This placement system was experienced by secondary students in the present sample as this study was conducted in 2007. From 2008 onwards, however, Primary-5 and Primary-6 pupils in Singapore are tracked into either a foundation (weaker) or standard (stronger) level in individual academic subjects (English, math, science, mother tongue). Consequently, it is possible, for example, that a student takes a standard-level English and mother tongue but he or she takes math and science at the foundation level.

[2]We also performed the same set of analyses reported in Tables 3 and 4 with student, class-average, stream-average, and school-average socioeconomic status (SES) included as

365

a covariate in the models. SES index was formed using information provided by students about their father's and mother's highest educational background. For the additional 80 analyses performed to test each of the models reported in Table 3 (10 models $\times$ 2 academic domains [English, math] $\times$ 4 analyses [student, class-average, stream-average, and school-average SES were separately included as a covariate]), SES-related factors were found to have small significant effects ($p < .05$) in only four models predicting English self-concept (ESC): class-average SES in Models 1 and 2 (both Bs = $-.10$) and stream-average SES in Models 1 and 2 ($B = -.17$ and $B = -.16$, respectively). Similarly, for the additional 96 analyses conducted to test models reported in Table 4 (6 models $\times$ 2 academic domains [English, math] $\times$ 2 frames of reference [gender, ethnicity] $\times$ 4 analyses [with student, class-average, stream-average, or school-average SES as a covariate]), SES-related factors were found to have significant effects ($p < .05$) in only four of the gender-context models: class-average SES ($B = -.10$) and stream-average SES ($B = -.17$) on ESC and stream-average SES ($B = .14$) and school-average SES ($B = .24$) on math self-concept (MSC). In all of these additional models tested, the size, significance, and direction of the parameters of central interest (i.e., the regressions weights of student achievement and contextual predictors on ESC and MSC) remained the same, suggesting that findings reported in Tables 3 and 4 are robust and unaffected by SES of the students or their peers—a set of findings consistent with those found in previous studies (see Marsh, 1987; Marsh et al., 2008).

References

Aboud, F. E. (1976). Self-evaluation: Information seeking strategies for interethnic social comparison. *Journal of Cross-Cultural Psychology*, *7*, 289–300.

Aboud, F. E. (2003). The formation of in-group favouritism and out-group prejudice in young children: Are they distinct attitudes? *Developmental Psychology*, *39*, 48–60.

Aiken, L., & West, S. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage Publications.

Alicke, M. D., Zell, E., & Bloom, D. L. (2010). Mere categorization and the frog-pond effect. *Psychological Science*, *21*, 174–177.

Bachman, J. G. (2002). *Volume I of the documentation manual*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

Blanton, H., Buunk, B. P., Gibbons, F. X., & Kuyper, H. (1999). When better-than-others compare upwards: Choice of comparison and comparative evaluations as independent predictors of academic performance. *Journal of Personality and Social Psychology*, *76*, 420–430.

Bond, M. H. (1991). *Beyond the Chinese face: Insight from psychology*. New York, NY: Oxford University Press.

Branden, N. (1994). *Six pillars of self-esteem*. New York, NY: Bantam.

Byrne, B. (1996). *Measuring self-concept across the lifespan*. Washington, DC: American Psychological Association.

Cleveland, C., Blascovich, J., Gangi, C., & Finez, L. (2011). When good teammates are bad: Physiological threat on recently formed teams. *Small Group Research*, *42*, 3–31.

Crittenden, K. S. (1991). Asian self-effacement or feminine modesty? Attributional patterns of women university students in Taiwan. *Gender & Society*, *5*, 98–117.

Dickhäuser, O., & Plenter, I. (2005). "Letztes Halbjahr stand ich zwei." Zur Akkuratheit selbst berichteter Noten ["Last term I had a B." On the accuracy of self-reported school grades]. *Zeitschrift für Pädagogische Psychologie*, *19*, 219–224.

Dijkstra, P., Kuyper, H., van der Werf Buunk, A. P., & van der Zee, Y. G. (2008). Social comparison in the classroom: A review. *Review of Educational Research*, *78*, 828–879.

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

Goldstein, H. (1995). *Multilevel statistical models*. London: Arnold.

Graham, J. W., & Hoffer, S. M. (2000). Multiple imputation in multivariate research. In T. D. Little, K. U. Schnable & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 201–218). Mahwah, NJ: Erlbaum.

Guay, F., Larose, S., & Boivin, M. (2004). Academic self-concept and educational attainment level: A ten-year longitudinal study. *Self and Identity*, *3*, 53–68.

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*, 124–136.

Gudykunst, W. B. (Ed.). (2003). *Cross-cultural and intercultural communication*. Thousand Oaks, CA: Sage.

Hattie, J. C. (2002). Classroom composition and peer effects. *International Journal of Educational Research*, *37*, 449–481.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

Howes, C., & Wu, F. (1990). Peer interactions and friendships in an ethnically diverse school setting. *Child Development*, *61*, 537–541.

Huguet, P., Dumas, F., Marsh, H. W., Regner, 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.

Ireson, J., & Hallam, S. (2001). *Ability grouping in education*. London: Paul Chapman Publishing.

Ireson, J., & Hallam, S. (2009). Academic self-concepts in adolescence: Relations with achievement and ability grouping in schools. *Learning and Instruction*, *19*, 201–213.

Ireson, J., Hallam, S., & Plewis, I. (2001). Ability grouping in secondary schools: Effects on pupils' self-concepts. *British Journal of Educational Psychology*, *71*, 315–326.

Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin*, *125*, 470–500.

Kulik, C. L. (1985). *Effects of inter-class ability grouping on achievement and self-esteem*. Paper presented at the annual meeting of the American Psychological Association, Los Angeles.

Kulik, C. L., & Kulik, J. A. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal*, *21*, 799–806.

Liem, A. D., Lau, S., & Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship and English achievement outcome. *Contemporary Educational Psychology*, *33*, 486–512.

Liem, G. A. D., Ginns, P., Martin, A. J., Stone, B., & Herrett, M. (2012). Personal best goals and academic and social functioning: A longitudinal perspective. *Learning and Instruction*, *22*, 222–230.

Liem, G. A. D., & Martin, A. J. (2011). Peer relationships and adolescents' academic and non-academic outcomes: Same-sex and opposite-sex peer effects and the

mediating role of school engagement. *British Journal of Educational Psychology*, *81*, 183–206.

Liu, W. C., & Wang, C. K. J. (2005). Academic self-concept: A cross-sectional study of grade and gender differences in a Singapore secondary school. *Asia Pacific Education Review*, *6*, 20–27.

Liu, W. C., Wang, C. K. J., & Parkins, E. J. (2005). A longitudinal study of students' academic self-concept in a streamed setting: The Singapore context. *British Journal of Educational Psychology*, *75*, 567–585.

Lüdtke, O., Köller, O., Marsh, H. W., & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology*, *30*, 263–285.

Marsh, H., & Martin, A. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*, 59–77.

Marsh, H. W. (1984). Self-concept, social comparison, and ability grouping: A reply to Kulik and Kulik. *American Educational Research Journal*, *21*, 799–806.

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

Marsh, H. W. (1991). The failure of high ability high schools to deliver academic benefits: The importance of academic self-concept and educational aspirations. *American Educational Research Journal*, *28*, 445–480.

Marsh, H. W. (1992). *Self-Description Questionnaire (SDQ) III: A theoretical and empirical basis for the Measurement of multiple dimensions of late adolescent self-concept: A test manual and a research monograph*. Sydney: Publication Unit, SELF Research Centre, University of Western Sydney.

Marsh, H. W. (1998). Age and gender effects in physical self-concepts for adolescent elite-athletes and non-athletes: A multi-cohort-multi-occasion design. *Journal of Sport & Exercise Psychology*, *20*, 237–259.

Marsh, H. W. (2007). *Self-concept theory, measurement and research into practice: The role of self concept in educational psychology—25th Vernon-Wall lecture series*. London: British Psychological Society.

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

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science*, *1*, 133–163.

Marsh, H. W., Kong, C.-K., & Hau, K.-T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of Personality and Social Psychology*, *78*, 337–349.

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin*, *34*, 542–552.

Marsh, H. W., & O'Mara, A. J. (2010). Long-term total negative effects of school-average ability on diverse educational outcomes: Direct and indirect effects of the big-fish-little-pond effect. *Zeitschrift für pädagogische psychologie (The German Journal of Educational Psychology)*, *24*, 51–72.

Marsh, H. W., & Rowe, K. J. (1996). The negative effects of school-average ability on academic self-concept—An application of multilevel modeling. *Australian Journal of Education*, *40*, 65–87.

Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K.-T., O'Mara, A. J., & Craven, R. G. (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., Tracey, D. K., & Craven, R. G. (2006). Multidimensional self-concept structure for preadolescents with mild intellectual disabilities: A hybrid multi-group-mimic approach to factorial invariance and latent mean differences. *Educational and Psychological Measurement*, *66*, 795–818.

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., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*, 297–416.

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

Meisel, J. C., & Blumberg, C. J. (1990). The social comparison choices of elementary and secondary school students: The influence of gender, race, and friendship. *Contemporary Educational Psychology*, *15*, 170–182.

Nagengast, B., & Marsh, H. W. (2011). The negative effect of school-average ability on science self-concept in the UK, the UK countries and the world: The Big-Fish-Little-Pond-Effect for PISA 2006. *Educational Psychology*, *31*, 629–656.

Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.

O'Connell, A. A., & McCoach, D. B. (2008). (Eds.). *Multilevel modeling of educational data*. Charlotte, NC: Information Age Publishing.

Pakir, A. (1993). Two tongue tied: Bilingualism in Singapore. *Journal of Multilingual and Multicultural Development*, *14*, 73–90.

Preckel, F., & Brüll, M. (2008). Grouping the gifted and talented: Are gifted girls most likely to suffer the consequences? *Journal for the Education of the Gifted*, *32*, 54–85,

Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2009). *A user's guide to MLwiN*. Bristol, UK: Centre for Multilevel Modelling.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rogers, C. M., Smith, M. D., & Coleman, J. M. (1978). Social comparison in the classroom: The relationship between academic achievement and self-concept. *Journal of Educational Psychology*, *70*, 50–57.

Rubie-Davies, C. M. (2006). Teacher expectations and student self-perceptions: Exploring relationships. *Psychology in the Schools*, *43*, 537–552.

Seta, J. J., & Seta, C. E. (1996). Big fish in small ponds: A social hierarchy analysis of intergroup bias. *Journal of Personality and Social Psychology*, *71*, 1210–1221.

Singapore Department of Statistics. (2010). *Yearbook of statistics*. Singapore: Ministry of Trade and Industry.

Singapore Ministry of Education. (2009). *2009 education statistics digest*. Singapore: Author.

Skaalvik, E. M. (1997). Issues in research on self-concept. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (*Vol. 10*, pp. 51–98). Greenwich, CT: JAI.

Skaalvik, E. M., & Skaalvik, S. (2002). Internal and external frames of reference for academic self-concept. *Educational Psychologist*, *37*, 233–244.

Suls, J., & Wheeler, L. (2008). A reunion for approach-avoidance motivation and social comparison. In A. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 585–600). New York, NY: Psychology Press.

Tarrant, M. (2002). Adolescent peer groups and social identity. *Social Development*, *11*, 110–123.

Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, *99*, 253–273.

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: using group composition and status to predict self-concept and interest in ninth grade mathematics. *Journal of Educational Psychology*, *98*, 788–806.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*, 111–133.

Zell, E., & Alicke, M. D. (2009). Contextual neglect, self-evaluation, and the frog-pond effect. *Journal of Personality and Social Psychology*, *97*, 467–482.

Zell, E., & Alicke, M. D. (2010). The local dominance effect in self-evaluation: Evidence and explanations. *Personality and Social Psychology Review*, *14*, 368–384.