Psychological Well-Being of Intellectually and Academically Gifted Students in Self-Contained and Pull-Out Gifted Programs

Trent N. Cash[1] and Tzu-Jung Lin[2]

[1]Carnegie Mellon University

[2]The Ohio State University

Author Note

Trent N. Cash, Department of Social and Decision Sciences at Carnegie Mellon University.

Tzu-Jung Lin, Department of Educational Studies at The Ohio State University.

Correspondence concerning this article should be addressed to Trent N. Cash, Department of Social and Decision Sciences, Carnegie Mellon University, 5000 Forbes Ave., 222 Porter Hall, Pittsburgh, PA, 15213. E-Mail: Tcash@andrew.cmu.edu, Phone: 330-451-9972, ORCID: 0000-0001-6624-9616.

Abstract

This study examined the psychological well-being of students enrolled in two gifted programs with different service delivery models. Participants were 292 fifth- and sixth-grade students (M_{age} = 11.70, SD_{age} = 0.65) enrolled in a gifted math pull-out program (n = 103), a self-contained gifted program (n = 90), or a program providing no gifted services, which served as a control group (n = 99). Multiple differences in psychological well-being across programs were revealed in Hierarchical Linear Models, particularly in terms of math self-concept, loneliness, and maladaptive perfectionism. Students in the two gifted programs reported different patterns of psychological well-being when compared to students in the no gifted services control group. These differences suggest distinct social phenomena underlying the two different service delivery models.

In the United States, local educational agencies (LEAs) are granted the power to determine what gifted services they want to provide and how they want to implement them. Because of this local autonomy, gifted services across the country reflect the highly variable fiscal positions and educational philosophies of individual districts' LEAs (Kettler et al., 2015; National Association for Gifted Children [NAGC] & The Council of State Directors of Programs for the Gifted, 2015). As a result, students in the United States are provided with different gifted programs that implement a wide variety of different service delivery models, many of which have not been fully evaluated in methodologically rigorous ways (Bergold et al., 2018; Dimitriadis, 2016).

Though the value of various gifted service delivery models for meeting the unique needs of gifted students (e.g., Bergold et al., 2018) has been studied and debated for many years (e.g., Plucker & Dilley, 2016), little is known about the relationship between different service delivery models, such as full-time, self-contained gifted models (e.g., Vogl & Preckel, 2014) and pull-out gifted models (e.g., Dimitriadis, 2012), and gifted students' psychological well-being. Most existing studies that have explored this concern have focused on one specific service delivery model (e.g., van der Meulen et al., 2013; Vidergor & Gordon, 2015) rather than comparing across different delivery models. Research that involves a direct comparison of different models is needed to establish a better understanding of how different service delivery models may be differentially associated with various facets of gifted students' psychological well-being (Dimitriadis, 2016) and to ensure that gifted students are not only thriving academically, but socially and emotionally as well. To begin filling this research gap, this cross-sectional survey study compared the psychological well-being of fifth- and sixth-grade students in two gifted programs that implemented different service delivery models (self-contained and pull-out) and a

program that did not provide any gifted services. Findings of this study are expected to provide practitioners, policymakers, and researchers with a deeper understanding of the psychological well-being of gifted students enrolled in programs that implement different service delivery models.

Theoretical Background

Gifted Service Delivery Models

Decades of research have shown that gifted students have unique academic (e.g., Vidergor & Gordon, 2015) and psychological needs (e.g., Bergold et al., 2018). Addressing both types of needs requires LEAs to be familiar with state and local requirements and resources regarding gifted education (NAGC & The Council of State Directors of Programs for the Gifted, 2015) and to invest significant time and resources into creating programs that are both sustainable and effective. To this end, LEAs implement a wide variety of gifted service delivery models within schools (NAGC & The Council of State Directors of Programs for the Gifted, 2015). Commonly implemented models include but are not limited to: (a) cluster models in which gifted students are concentrated in one classroom but still integrated with non-identified students; (b) pull-out models in which gifted students are instructed separately from non-identified students for set periods of time each day or week; and (c) full-time, self-contained models in which gifted students learn exclusively with other gifted students (NAGC, n.d.-a, n.d.-b). Despite their differences, each of these service models is intended to both encourage academic growth and promote psychological well-being. However, research evaluating the relationship between various service delivery models and the achievement of the two aforementioned goals is limited.

Of the two goals of gifted services (meeting gifted students' academic and psychological needs), the academic growth component has been more widely evaluated. In a seminal meta-analysis, Kulik and Kulik (1982) found that ability grouping was associated with small, positive effects on secondary students' achievement, particularly for high-ability students receiving enriched instruction. In a later meta-analysis, Kulik and Kulik (1992) found that the positive association between ability grouping and achievement also depended on the specific grouping strategy used and that gifted programs that altered curricula more substantially – such as enrichment and acceleration – were associated with greater improvements in student learning than programs with less substantial changes. More recently, Adelson et al. (2012) found that enrollment in gifted programming was not associated with gifted students' achievement or academic attitudes in the domains of math or reading. However, Adelson and colleagues noted that these results were not a condemnation of gifted education as a whole but rather evidence of the need for further investigation into specific service delivery models.

Although our understanding of the relationship between different gifted service delivery models and academic growth is far from complete (Adelson et al., 2012), the relationship between various gifted service delivery models and psychological well-being has received even less empirical attention (Dimitriadis, 2016; Plucker & Dilley, 2016). This dearth of empirical research is notable, as lower well-being is associated with a host of negative outcomes, including decreased life satisfaction and decreased academic performance (e.g., Moeller & Seehuus, 2019; Rice et al., 2014). Of the research that has been conducted, Vogl and Preckel (2014) found that gifted students in full-time, self-contained classes reported higher social self-concepts, better student-teacher relationships, and greater interest in school than their gifted counterparts who

were not in self-contained classrooms. Similarly, van der Meulen et al. (2013) found that gifted students reported higher self-concepts after enrollment in a pull-out gifted program.

Although studies of this kind generally suggest that some gifted programs may be positively associated with psychological well-being, the literature as a whole lacks evidence comparing and contrasting different service delivery models in terms of gifted students' psychological well-being, as nearly all studies that have been conducted have evaluated only one program at a time (e.g., van der Meulen et al., 2013; Vogl & Preckel, 2014). Furthermore, the existing literature is based on research that has been conducted in a variety of international contexts, limiting generalizability to American students due to variable cultural attitudes, policies, and practices regarding gifted education (e.g., Subotnik et al., 2017). The current study attempted to fill this gap in the literature by evaluating the psychological well-being of gifted students in both a part-time, pull-out gifted program and a full-time, self-contained gifted program, in comparison with students in classes providing no gifted services.

Potential Social Phenomena Underlying Different Gifted Service Delivery Models

The relationships between various gifted service delivery models and gifted students' psychological well-being have been attributed to a variety of factors including, but by no means limited to: (a) the presence of specialized teachers with knowledge of the unique needs of gifted students (NAGC & The Council of State Directors of Programs for the Gifted, 2015), (b) separation from non-identified students who think differently than gifted students (Gross, 2002), (c) the presence of like-minded peers (Herrmann et al., 2016), (d) increased curricular flexibility (Vidergor & Gordon, 2015), and (e) higher levels of academic rigor (Moon et al., 2002). Although each of these accounts has merit, the current study examined the ways in which the social phenomena associated with various gifted service delivery models may be related to gifted

students' psychological well-being. Two social phenomena that may be particularly relevant for understanding the relationship between various gifted service delivery models and gifted students' psychological well-being are social comparison theory and labeling theory.

Social Comparison Theory

Social comparison theory (Festinger, 1954) asserts that the human habit of comparing oneself to those around them is a "fundamental, ubiquitous, and robust human proclivity" (Corcoran et al., 2011, p. 134). As such, one's self-view is directly affected by the people whom they are frequently around, including classroom peers (see Dijkstra et al., 2008 for a review of social comparisons in the classroom). In the educational context, social comparison theory has commonly been interpreted through the big fish little pond effect, which suggests that a student surrounded by lower ability peers will feel more academically competent, whereas the same student surrounded by higher ability peers will tend to feel less academically competent (Marsh et al., 1995; Marsh & Parker, 1984). This social comparison process has implications for psychological well-being. For example, evidence suggests that social comparisons are related to self-esteem (Alfasi, 2019; M. S. W. Wong & Watkins, 2001) and one's ability to fulfill their basic psychological needs (Thøgersen-Ntoumani et al., 2018). Furthermore, individuals who are placed in situations in which they face unfavorable social comparisons tend to feel less competent (Kamarova et al., 2017) and less related to their peers (Steers, 2016).

The effects of social comparisons are particularly relevant for the issue of gifted service delivery models because these models alter students' reference points regarding academic standing. For example, in self-contained gifted classrooms, students are constantly surrounded by other gifted students whereas gifted students in pull-out programs spend significant amounts of time with non-identified students as well as other gifted students. The big-fish-little-pond effect

in the context of gifted education was most notably explored in two longitudinal studies by Marsh et al. (1995), which showed that gifted students enrolled in full-time, self-contained gifted programs reported larger declines in math, reading, and school self-concepts than their peers in mixed ability classrooms after one academic year. Multiple recent studies have found similar results supporting the big fish little pond effect (e.g., Preckel et al., 2008, 2010), but Herrmann et al. (2016) found that the benefit associated with belonging to a high achieving group outweighed the negative social comparison effects.

Labeling Theory

Despite its roots in sociology (Becker, 1963), labeling theory has been widely used in educational research, with evidence showing that labels can affect teachers' perceptions of students (Glock & Krolak-Schwerdt, 2014; Ohan et al., 2011; Rosenthal & Jacobson, 1968), students' perceptions of one another (Crocker et al., 1998; Matthews et al., 2015), and even students' perceptions of themselves (Norwich & Kelly, 2004). Labels, which can be defined as attributes attached to an individual (Gove, 1980), may be negatively associated with students' psychological well-being and social connectedness (Green, 2007; Walton & Carr, 2012). These labels can influence a student's self-esteem (Thomson, 2012), willingness to seek mental health support (Moses, 2009), and general perceived well-being (Green, 2007). In addition to these negative impacts, labels can have positive impacts on a student by exposing the student to opportunities that are not accessible without the label, such as gifted education classes, and by enhancing the student's sense of belonging to a group (Chambers et al., 2019). Furthermore, positive stereotypes associated with some labels, particularly those associated with student performance, have shown positive associations with academic performance (Clark et al., 2017).

Labeling theory is particularly relevant for research involving gifted students because giftedness is a label that carries a variety of stigmas and stereotypes (Baudson, 2016; Preckel et al., 2015) that are associated with anxiety and low self-esteem (Lam & Kirby, 2002; Thomson, 2012). Furthermore, gifted identification labels may lead to an in-group/out-group dynamic wherein gifted students are viewed as an out-group and therefore socially isolated from their non-identified peers (Killen et al., 2013). Educational programs can partially mitigate the negative impacts of these labels if they can help students to align the gifted label with positive characteristics (Clark et al., 2017) and reduce in-group/out-group biases.

Psychological Well-Being of Gifted Students

One of the most contentious questions in the field of gifted education is whether intellectually and academically gifted students are better or worse off than non-identified students in terms of psychological well-being. On one side of this debate is the harmony hypothesis, which suggests that gifted students' intelligence leads them to excel in nearly every domain and ultimately have improved psychological well-being as a result of said success (Bergold et al., 2018). This claim is supported by evidence showing that gifted students either outperform or do not differ from non-identified students in some dimensions of psychological well-being (e.g., Lee et al., 2012; Shechtman & Silektor, 2012).

In contrast to the harmony hypothesis, the disharmony hypothesis suggests that gifted students tend to experience lower levels of psychological well-being than their peers who are not identified as gifted because they do not fit in and are subject to stereotypes, such as being a "nerd" (Bergold et al., 2018). This concern is often framed in terms of gifted students' asynchronous development that causes them to be out of sync with their age peers (Gross, 2002). In line with this theory, Vialle et al. (2007) showed that gifted secondary school students

experienced lower levels of well-being and social support than their peers who were not identified as gifted. To date, comparative evidence of the relative psychological well-being of gifted and non-identified students is mixed (Amini, 2005; Lee et al., 2012; Vogl & Preckel, 2014), in part because psychological well-being has been inconsistently conceptualized in the literature (Pollard & Lee, 2003).

Defining Psychological Well-Being

In this study, we define psychological well-being as students' positive evaluations of themselves and their life experiences (e.g., self-concept, self-esteem), psychological health (e.g., perfectionism), and relationships with others (e.g., loneliness, attitudes towards school; Neihart, 1999; Vogl & Preckel, 2014). Specifically, the current study incorporated five constructs to capture these dimensions of psychological well-being. Two of these five constructs (*attitudes towards school* and *perfectionism*) consist of two subconstructs each, which results in seven total subconstructs. These seven subconstructs, which are discussed in the remainder of this section, will henceforth be referred to as the seven constructs.

Loneliness, which has been defined as the presence of a negative feeling resulting from a deficiency of social relationships, is a well-established component of psychological well-being (Asher & Paquette, 2003). High levels of loneliness have been linked to several negative outcomes, including depression, anxiety, and reduced life satisfaction (Moeller & Seehuus, 2019; Salimi, 2011). In a study with gifted primary school students, Kroesbergen et al. (2015) found that gifted students had lower self-worth and experienced less social acceptance than their peers who were not identified as gifted. Loneliness is frequently cited as a major challenge for gifted students (Cheng & Furnham, 2002; Ogurlu et al., 2018; Özbay & Palanci, 2011).

Self-esteem is defined as an individual's sense of worth (Rosenberg, 1965) and has often been considered as an indicator of psychological well-being. A meta-analysis conducted by Judge et al. (2002) found that self-esteem was strongly associated with overall life satisfaction and happiness. Furthermore, multiple studies have reported strong relationships between self-esteem and social and emotional experiences at school (Bajaj et al., 2016; Q. Yang et al., 2019). Previous research has found that gifted students tend to report higher self-esteem than students who are not identified as gifted (Gross, 2001; Lea-Wood & Clunies-Ross, 1995).

Perfectionism (Adaptive and Maladaptive) is a multidimensional construct (Flett & Hewitt, 2014; Stoeber & Otto, 2006) that has frequently been associated with gifted students' health and academic achievement (Rice & Ray, 2018). Perfectionism has been conceptualized in a variety of ways, with different conceptualizations identifying different underlying dimensions (Fletcher & Speirs Neumeister, 2012; Flett & Hewitt, 2014). In this study, we consider two dimensions of perfectionism: adaptive and maladaptive perfectionism (Rice et al., 2014). Maladaptive perfectionism, which can be conceptualized as a sense of discrepancy between an individual's unrealistic standards and their ability to perform (Rice et al., 2014), has been linked with depression, anxiety, and a lack of sense of purpose. Adaptive perfectionism, which can be conceptualized as setting high standards for oneself (Rice et al., 2014), has been found to be positively associated with self-esteem and academic success (Grzegorek et al., 2004; Mobley et al., 2005; Ortega et al., 2014; Park & Jeong, 2015; Rice & Slaney, 2002). The debate persists over whether any perfectionism can be adaptive in the long run (Flett & Hewitt, 2006, 2014). While perfectionism is often considered a stable trait, recent studies have suggested that perfectionism may change over time and may be malleable in response to interventions (e.g., Herman et al., 2013). A recent meta-analysis by Stricker et al. (2019) suggested that gifted

students tend to experience slightly higher levels of adaptive perfectionism than their peers who are not identified as gifted and equal levels of maladaptive perfectionism.

Attitudes towards school (school liking and school avoidance) conceptualizes the degree to which students profess to like or dislike school. School liking has been shown to be negatively associated with loneliness (Rönkä et al., 2017) and positively associated with school belonging (Honma & Uchiyama, 2014). School avoidance is positively associated with being bullied (Hutzell & Payne, 2012), feelings of interpersonal maladaptation (i.e., social dissatisfaction), and depression (Honjo et al., 2003). Although little research has been conducted on gifted students' levels of school liking and avoidance, gifted students tend to report higher levels of school belonging, a related construct (Godor & Szymanski, 2017).

Math self-concept refers to a student's self-perceptions about their own abilities in the domain of mathematics (Lohbeck, 2016; Shavelson et al., 1976). Math self-concept has been linked to school belonging for girls, student-teacher relationships for boys (T. K. Y. Wong et al., 2019), and math anxiety for both boys and girls (Bai et al., 2009; Krinzinger et al., 2009; Wu & Cheung, 2012). Gifted students report higher academic self-concepts (Litster & Roberts, 2010) than students who are not identified as gifted, but Herrmann et al. (2016) found that this difference dissipated when gifted students were frequently surrounded by other gifted students. We chose to evaluate math self-concept because the students in the pull-out gifted program only received specialized math instruction and because all students were surveyed during their math classes.

The Current Study

The current study implemented a cross-sectional survey design in which fifth- and sixth-grade students in three different programs—a program that provided no gifted services (NGS

program), a gifted math pull-out program, and a self-contained gifted program—were assessed on the aforementioned seven constructs of psychological well-being toward the end of the academic year. To limit potential confounds of subject matter, NGS students were only recruited and surveyed from math classes that did not provide gifted services and were taught by the teachers who also taught pull-out gifted math classes. We were particularly interested in students at the fifth- and sixth-grade levels because early adolescence is a phase of life that is vital for psychosocial development (Tottenham & Galván, 2016). This study was conducted within a suburban school district that provided two different gifted programs that implemented different service delivery models: a full-time, self-contained gifted program and a math pull-out program, both of which used enrollment structures dependent on students' gifted identification statuses. Although there are many ways to define whether or not a student is gifted (Mcclain & Pfeiffer, 2012), this study used the school district's guidelines for identification and only focused on students who were intellectually and academically gifted, as determined by scores on subject-specific standardized tests and assessments of cognitive ability administered by the district.

Two specific research questions were addressed in the current study: Do gifted students enrolled in self-contained gifted programs and gifted math pull-out programs differ in terms of psychological well-being? And do gifted students receiving gifted programming of either kind differ from their peers in the NGS program in terms of psychological well-being? We predicted that students in the two gifted programs would exhibit different patterns of psychological well-being in comparison to students in the NGS program. We did not make any directional hypotheses because the literature regarding gifted students' psychological well-being remains mixed and exploratory in nature (e.g., Amini, 2005; Lee et al., 2012; Vogl & Preckel, 2014).

Method

Setting

This study was conducted in a large, public, Midwestern school district that was chosen because it offered both a full-time, self-contained gifted program and a gifted math pull-out program for students enrolled in grades five and six. The self-contained gifted program was located in a different wing of the building, used enriched and/or accelerated curricula for all four core subjects, emphasized the frequent use of individual and group projects in place of direct instruction, encouraged students to learn at their own pace, and was taught by instructors who were certified as gifted intervention specialists. Due to the self-contained nature of the program, the students spent the entire school day, including specials, lunch, and recess, with gifted peers. To be enrolled in the self-contained program, students were required to perform three standard deviations above the mean on a standardized test of School Ability Index (SAI) or Intelligence Quotient (IQ) or perform two standard deviations above the mean on an SAI or IQ test while also scoring above the 95th percentile on math and reading standardized tests. These program placement criteria aligned with the district's gifted identification criteria, as students who performed two standard deviations above the mean on an SAI or IQ test were identified as having superior cognitive abilities and students who scored above the 95th percentile on math or reading standardized tests were identified as gifted in math or reading, respectively.

The math pull-out classes occurred once per day, replaced NGS math classes, and included accelerated math curricula taught by a gifted intervention specialist. To qualify for the gifted math pull-out program, students were required to perform three standard deviations above the mean on a SAI or IQ test or have a SAI or IQ score of 122 or higher while also being gifted in math (i.e., scoring above the 95th percentile on a math standardized test). Aside from the daily

math class, students enrolled in the math pull-out program were enrolled in the same classes as students in the NGS program and were integrated with their non-identified peers.

Students who qualified for the self-contained program could opt into the math pull-out program or NGS program, and students who qualified for the math pull-out program could opt into the NGS program, but this was not common in our sample. Of the 174 students for whom we were able to collect SAI and standardized test scores to validate qualifications, only six (3.4%) opted into a program with lower requirements than the one for which they qualified. All six of these students qualified for the self-contained program based on their SAI and standardized test scores, but five opted into the pull-out program and one opted into the NGS program. The NGS program was primarily made up of students who were not identified as gifted, but 16% were gifted in reading, 8% were gifted in math, and 2% were identified as having superior cognitive abilities. Other than the one student who opted out of the self-contained program, no students in the NGS program qualified for gifted services. The NGS students who were identified as gifted did not qualify for gifted services because they did not meet the SAI or IQ score thresholds.

All teachers who taught either the self-contained program or the math pull-out program at the fifth- or sixth-grade level (N = 12) were invited to have their classes participate in this study during a professional development session. All 12 teachers agreed to participate in the study and completed a brief, one-page survey about their demographic backgrounds and teaching experience. Six of the participating teachers taught the self-contained classes, and the remaining six teachers taught a combination of math pull-out classes and NGS classes. Each teacher allowed the research team to recruit all of their classes for a total of 22 classes. Of these 22 classes, 10 were NGS classes, 6 were math pull-out classes, and 6 were self-contained classes.

These classes represented all fifth- and sixth-grade classes taught by teachers who taught at least one gifted class.

Participants

Students

All students in the participating classrooms were invited to participate in the study. A total of 292 students participated in this study (158 females, 128 males, 6 provided no gender data). If a student provided assent, they were given a consent packet to be signed by a guardian and returned to the classroom teacher. The students who participated in the study were divided between the NGS program (n = 99; 10 classrooms), the gifted math pull-out program (n = 103; 6 classrooms), and the self-contained gifted program (n = 90; 6 classrooms). All students were enrolled in fifth (n = 135) or sixth (n = 157) grade, with ages ranging from 7.37 to 13.54 years (M = 11.70, SD = 0.65). The age range was wider than expected because of one anomalous 5th grader who was 7.37 years old. The next youngest student was 9.72 years old. The majority of students were White (72%, n = 205), and the remainder of the students were Black (11%), multiracial (11%), Asian (4%), or other race(s) (2%). Demographic data separated by academic program is described in Table 1.

In addition to granting consent, guardians also provided demographic information about their families and children. The child-centric portion of the survey included questions about the child's age, race, ethnicity, school history, and gifted identification status. The family-centric portion of the survey included questions about household income, guardian education, and the individuals living in the household. Guardians of 289 (99%) of the 292 participating students completed the guardian survey. According to these surveys, 79% of guardians had college degrees. For questions regarding guardian education, guardians were provided with the option to

report data for one guardian or two, if applicable, and responses were coded on a 1-5 scale ranging from "high school diploma or less" to "doctoral degree." The median household income was between $110,000 and $130,000. Household income and guardian educational attainment data separated by academic program is described in Table 1.

Of the 292 students who participated in the study, 175 (60%) received guardian permission for the district to share their school performance and gifted identification records. For these students, the district provided data on GPA (M = 3.61, SD = 0.33), SAI scores (M = 121.09, SD = 13.14), standardized reading testing scores (M = 756.79, SD = 32.27), standardized math testing scores (M = 760.56, SD = 25.61), and gifted identification statuses in various domains. Among the students who had permission to have their district records released, 47% (n = 82) were identified as having superior cognitive abilities (i.e., scored two standard deviations above the mean on an SAI or IQ test), 74% (n = 130) were identified as gifted in math (i.e., scored above the 95th percentile on a math standardized test), and 62% (n = 108) were identified as gifted in reading (i.e., scored above the 95th percentile on a reading standardized test). Performance and identification data separated by academic program is described in Table 1.

Teachers

According to their responses on the teacher survey, 92% of participating teachers were female, and 100% were White and non-Hispanic. All teachers had a master's degree as their highest level of educational attainment. The teachers who taught the self-contained classes (n = 6) were all gifted intervention specialists with gifted endorsements, meaning they were specially trained to educate gifted students. None of the teachers who taught the NGS and math pull-out classes (n = 6) had gifted endorsements. The teachers' years of teaching experience ranged from 4 to 27 years (M = 18.83, SD = 6.97), and their years of experience teaching gifted students

ranged from 2 to 22 years (M = 10.75, SD = 7.00). Reported years of teaching experience for the teachers of the self-contained classrooms ranged from 10 to 25 years (M = 19.93, SD = 5.96) and years of experience teaching gifted students ranged from 5 to 22 years (M = 14.67, SD = 7.29). Teachers who taught the NGS classes and math pull-out classes reported 4 to 27 years of teaching experience (M = 18.33, SD = 8.41) and 2 to 12 years of experience teaching gifted students (M = 6.83, SD = 4.22).

Procedure and Measures

Within a 2-week period toward the end of the academic year, all students who received guardian consent to participate completed a questionnaire evaluating their psychological well-being. The survey was completed on paper in the classroom during the students' math period. The survey contained five measures evaluating seven constructs of psychological well-being, all of which were converted to a 5-point likert scale.

We chose the following five measures to evaluate the seven constructs for three primary reasons. First, these measures have been shown in the literature to allow psychometrically valid inferences. Second, these measures have been used with early adolescent populations similar in age to our sample (e.g., Chen et al., 2016; Eccles et al., 2005; Meeus et al., 2019; Swanson et al., 2012; H. Yang et al., 2016). Third, all of the selected measures evaluated constructs that are particularly relevant to gifted students, such as loneliness (Cheng & Furnham, 2002; Ogurlu et al., 2018; Özbay & Palanci, 2011), self-esteem (Lam & Kirby, 2002; Thomson, 2012), math self-concept (Litster & Roberts, 2010; Marsh et al., 1995), and perfectionism (Rice & Ray, 2018; Stricker et al., 2019), or constructs that are able to provide summary evaluations of students' attitudes towards their program, such as school avoidance (Godor & Szymanski, 2017; Hutzell & Payne, 2012) and school liking (Honma et al., 2014; Rönkä et al., 2017).

Loneliness

Student loneliness was measured using the 16 loneliness items (e.g., "I feel left out of things") from the Children's Loneliness and Social Dissatisfaction Scale (CLSDS) developed by Asher et al. (1984). In the literature, the CLSDS has shown high internal consistency (α = .90) and reliability (Spearman-Brown = .91; Split-half correlation between forms = .83; Guttman = .91; Asher et al., 1984). The reliability has been replicated with Chinese (α = .90) and Canadian (α = .83) students (Chen et al., 2016), and a confirmatory factor analysis by Ebesutani et al. (2012) determined that the CLSDS measures one unidimensional construct, despite some studies that have suggested a two-factor model separating the reversed and non-reversed items (e.g., Bagner et al., 2004). The 16 loneliness items also showed high internal consistency (α = .93, 95% CI [.91, .94]) within our sample, but a confirmatory factor analysis (CFA) indicated that the one-factor model using all 16 items did not fit the data well, $\chi^2(104) = 431.34$, $p < .001$, comparative fit index (CFI) = .87, root mean-square error of approximation (RMSEA) = .11, standardized root mean square residual (SRMR) = .06. In response, we removed the three items that showed the lowest factor loadings (items 3, 10, and 13 in the original survey). Once these items were removed, the remaining 13 items showed high internal consistency (α = .93, 95% CI [.91, .94]) within our sample and fit the data better, $\chi^2(65) = 265.90$, $p < .001$, CFI = .90, RMSEA = .10, SRMR = .04. The 13-item scale was used for all analyses.

Attitudes Towards School (School Liking and School Avoidance)

School liking and avoidance were measured using the School Liking and Avoidance Questionnaire (SLAQ) developed by Ladd & Price (1987). The SLAQ includes 14 items, 9 of which measure school liking (e.g., "Is school fun?") and 5 of which measure avoidance (e.g., "Do you wish you didn't have to go to school?"). To reduce repetition, we removed two of the

school liking items ("Does school make you feel like crying?" and "When you get up in the morning, do you feel happy about going to school?") from our measure. These items were removed due to similarity to other questions and because we were concerned the former would be less appropriate for our early adolescent population than for younger children, and the latter would be skewed by a dislike of waking up early. Both subscales have been shown to be internally consistent, with studies reporting high alphas for school liking (αs = .84 - .91) and school avoidance (αs = .76 - .92; Ladd, 1987; Ladd et al., 1996, 2000; Swanson et al., 2012). Both the school liking (α = .94, 95% CI [.93, .95]) and avoidance (α = .86, 95% CI [.83, .88]) subscales were internally consistent in our sample. Separate CFAs indicated that the one-factor model for the five school avoidance items fit the data well, $\chi^2(5) = 10.54$, $p = .06$, CFI = .99, RMSEA = .06, SRMR = .02, but that the one-factor model for the seven school liking items were only a moderate fit for the data, $\chi^2(14) = 138.55$, $p < .001$, CFI = .93, RMSEA = .18, SRMR = .04. Despite the unacceptable RMSEA value, we decided not to alter the model because of its high consistency, satisfactory CFI and SRMR values, and its conceptual support from the literature (Ladd & Price, 1987; Ladd et al., 1996, 2000; Swanson et al., 2012).

Self-Esteem

Self-esteem was measured using the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965). The RSES includes 10 items (e.g., "On the whole I am satisfied with myself"). The RSES has been reported to be reliable in terms of composite reliability (ρc = .84), test-retest reliability (r = .82), and internal consistency (αs = .86 - .88; Fleming & Courtney, 1984; Gray-Little et al., 1997; McKay et al., 2014). Furthermore, McKay et al. (2014) and Gray-Little at al. (1997) reported that the RSES is best modeled as a unidimensional construct, despite some studies that have suggested a two-factor model (e.g., Marsh et al., 2010; Supple et al., 2013). In our sample,

the RSES showed high internal consistency (α = .92, 95% CI [.91, .94]) and a CFA indicated that the 10 items of the RSES fit our data adequately, $\chi^2(35) = 107.73$, $p < .001$, CFI = .96, RMSEA = .09, SRMR = .04.

Perfectionism (Adaptive and Maladaptive)

Perfectionism was measured using the Short form of the Revised Almost Perfect Scale (SAPS) developed by Rice et al. (2014), a shortened version of the original 23-item measure (Slaney et al., 2001). The SAPS includes eight items, four of which measure standards (i.e., adaptive perfectionism; e.g., "I have high expectations for myself") and four of which measure discrepancy between expectations and performance (i.e., maladaptive perfectionism; e.g., "Doing my best never seems to be enough"). Both scales within the SAPS have been reported to be internally consistent (αs = .78 - .87; Rice et al., 2014; Wang et al., 2016). Furthermore, Rice et al. (2014) reported that no predictive value is lost by using the short form in place of the long form. The SAPS was internally consistent for our sample, with high alphas for both the standards (α = .84, 95% CI [.81, .87]) and discrepancy (α = .80, 95% CI [.76, .84]) subscales. Separate CFAs indicated that the one-factor model for the four adaptive perfectionism items fit our data well, $\chi^2(2) = .45$, $p = .80$, CFI = .99, RMSEA < .01, SRMR = .01, as did the model for the four maladaptive perfectionism items, $\chi^2(2) = 3.50$, $p = .17$, CFI = .99, RMSEA = .05, SRMR = .02.

Math Self-Concept

Students' math self-concept (e.g., "How good are you at math?") was measured using the 5-item math expectancy scale developed by Eccles and Wigfield (1995). This scale, which was based on earlier works in the domain of expectancy-value theory (Eccles et al., 1983, 1984), was initially used to measure math expectancies, but Eccles and Wigfield (1995) demonstrated that expectancies and self-concept load on the same factor, and thus can be treated as one construct.

The five items of the math expectancy scale have shown high internal consistency in the literature (α = .92; Eccles et al., 2005) and showed high internal consistency in our sample as well (α = .87, 95% CI [.84, .89]). A CFA indicated that the one-factor model for the five math self-concept items fit our data well, $\chi^2(5) = 6.96$, p = .22, CFI = .99, RMSEA = .04, SRMR = .02.

Analytical Strategy

Prior to analysis, composite scores of the seven psychological well-being constructs were calculated by averaging responses to all items underlying each construct. For each measure, between 7 and 14 students were excluded due to missed items. The number of students who missed items on each measure is reported in Table 2. No individual item was missed by more than 3.5% of students, and 275 students (94%) completed all items.

We conducted Hierarchical Linear Modeling (HLM) using Proc Mixed of SAS 9.0 and the maximum likelihood estimation method. Students (Level 1) were nested within classrooms (Level 2). Students' psychological well-being was predicted by program type (classroom-level predictor), after controlling for fixed effects including student characteristics (gender, race, highest level of education achieved by either of the student's parents or guardians) and grade level (classroom-level predictor). Each model included one of the psychological well-being variables as the dependent variable. Gender (female vs. male), race (white vs. non-white), and grade (6^{th} grade vs. 5^{th} grade) were coded as dichotomous variables, with male, non-white, and 5^{th} grade serving as the reference groups. Guardians' highest education was coded on a scale of 0-4 (high school diploma or less, vocational/technical training, associate's or bachelor's degree, master's degree, doctoral degree) and was treated as an uncentered interval-level variable. Program was treated as a three-level categorical variable. The model for each outcome variable

was conducted twice, once with the self-contained program serving as the reference group and once with the NGS program serving as the reference group. The model was specified as described below.

Level 1 (Student):

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{Female})_{ij} + \beta_{2j}(\text{White})_{ij} + \beta_{3j}(\text{Guardians'Highest Education})_{ij} + r_{ij}$$

Level 2 (Classroom):

Treating Self-Contained as the reference group:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(Sixth\ Grade)_j + \gamma_{02}(Pull-Out)_j + \gamma_{03}(NGS)_j + u_j$$

Treating NGS as the reference group:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(Sixth\ Grade)_j + \gamma_{02}(Pull-Out)_j + \gamma_{03}(Self-Contained)_j + u_j$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30}$$

Where $i = i^{th}$ student, $j = j^{th}$ classroom, and Y is an indicator of psychological well-being.

Intraclass correlation coefficients (ICCs) calculated from null models (models without covariates) showed that for each psychological well-being variable the percentage of the total variance explained between classrooms ranged from 1% to 12%. ICCs for each psychological well-being variable are reported in Table 4. Despite the small ICCs, we used HLMs rather than multiple regression models to account for additional dependence that may have arisen after predictors were entered into the models. Pairwise program comparisons were generated by altering the program that served as the reference group for the analysis. Following the method proposed by Raudenbush and Bryk (2002), the proportional reduction in variance that occurred due to the inclusion of the program variable was calculated for each model by comparing the

variance of reduced models that included all predictor variables except for the program variable to the variance of fully fitted models that included the program variable. These proportional reductions in variance are reported in Table 4. Missing data were handled using the listwise deletion approach because only 16-22 students (5-8%) had to be excluded from each analysis due to missing data. The demographics of the included and excluded students are described for each analysis in Supplementary Table 1. As a robustness check, we also replicated the findings using an imputed data set (40 imputations using SAS 9.0 Proc MI and Proc Mianalyze; see Supplemental Table 4).

Gender was included as a covariate because gender differences have been found in multiple constructs that evaluate psychological well-being, such as math self-concept (Seo et al., 2019), social support and loneliness (Zhang et al., 2015), and perfectionism (Margot & Rinn, 2016). Grade was included as a covariate because well-being generally decreases as students get older (Liu et al., 2016) and as they reach middle school age (Tobia et al., 2019). Race was included as a covariate because the district in this study was majority White, suggesting that Non-White students' psychological well-being may have been affected by their minority status (Akos & Galassi, 2004; Burchinal et al., 2008; Espinoza & Juvonen, 2011). Guardians' highest level of education was included as a covariate because guardian education predicts students' psychological well-being (Sutin et al., 2018) as well as their academic achievement (Ludeke et al., 2020), which in turn impacts program placement. Income was not used as a covariate because of its high rate of missing data (20.5%).

In order to address the potentially confounding effect of intelligence level across programs, we conducted an additional set of hierarchical linear models that were identical to the first set but controlled for School Ability Index (SAI) as a covariate. These analyses can at best

be supplementary to the analyses without SAI because the dataset with SAI scores oversampled gifted students and only contained 174 students (60% of the sample). As such, the models that included SAI as a covariate had less power to detect differences than the models without SAI. ICCs were not calculated for these models because of the reduced sample sizes and the oversampling of gifted students.

Scatterplots of residuals suggested homoscedasticity within all level-1 models. Levene's Test of Equality of Error Variances showed homoscedasticity for level one models that included SAI as a covariate (Fs = 0.66 – 1.21) and models that did not (Fs = 0.70 – 1.20), except for the model for loneliness that did not include SAI, which was slightly heteroscedastic ($F(23, 249) = 1.75$, $p = 0.02$). Because the scatterplot and the Levene's test for this model disagreed, we chose to proceed with the analysis but to interpret the results with caution. Q-Q plots comparing residuals and predicted values suggested that residuals were approximately normally distributed, whether SAI was included as a covariate or not. Residual-vs-fits plots suggested that the assumption of linearity was met.

Results

Preliminary Analyses

When averaged across the entire sample ($N = 292$), students reported values above the midpoint on the measures of loneliness ($M = 4.03$, $SD = 0.74$), adaptive perfectionism ($M = 3.83$, $SD = 0.83$), math self-concept ($M = 3.79$, $SD = 0.85$), self-esteem ($M = 3.61$, $SD = 0.89$), school liking ($M = 3.31$, $SD = 1.00$), and school avoidance ($M = 3.10$, $SD = 1.00$). Students only reported values below the midpoint for maladaptive perfectionism ($M = 2.81$, $SD = 0.98$). Students reported the lowest scores on the two constructs – school avoidance and maladaptive perfectionism – for which higher scores indicate lower psychological well-being. Descriptive

results of the psychological well-being survey separated by program are reported in Table 2. The bivariate correlations between the psychological well-being variables reported in Table 3 show that all of the psychological well-being constructs were correlated at a small to moderate level except for school avoidance, which was not significantly correlated with loneliness or maladaptive perfectionism.

As a whole, students across the three programs were equivalent in their demographic characteristics. Chi-square tests of association showed no significant relationship between race and program, χ^2 (2, N = 283) = 4.13, p = .13, or gender and program, χ^2 (4, N = 289) = 5.99, p = .20. Kruskal-Wallis H Tests showed no differences between the three programs in terms of level of household income, H(2, N = 232) = 5.28, p = .07, first guardian's highest education, H(2, N = 289) = 2.52, p = .28, or highest level of education attained by either guardian, H(2, N = 289) = 1.98, p = .37. For household income, many guardians (n = 57, 20%) reported "I don't know" or "Prefer Not to Answer," but these guardians were equally distributed across groups, χ^2 (2, N = 57) = 1.16, p = .56. Finally, a one-way ANOVA showed no significant difference across programs in terms of age, F(2, 283) = 0.69, p = .51.

Students who received (n = 175) and did not (n = 117) receive permission to have their test scores and gifted identification data released showed no differences in psychological well-being. Independent sample t-tests showed no significant differences between these groups in terms of loneliness, t(280) = 1.39, p = .17, school liking, t(281) = -0.49, p = .62, school avoidance, t(280) = 1.39, p = .17, self-esteem, t(276) = 1.35, p = .18, adaptive perfectionism, t(281) = 0.15, p = .89, maladaptive perfectionism, t(281) = -0.88, p = .38, or math self-concept t(283) = 0.29, p = .77. These results remained the same when controlling for the

interdependency among students within classrooms through hierarchical linear modeling (see Supplementary Table 2).

Demographically, Chi-square tests showed that White students were more likely than Non-White students to receive parental permission to have their SAI score and gifted identification data released, χ^2 (1, N = 283) = 6.55, p = .01, but there was no significant interaction between gender and data release permission status, χ^2 (2, N = 289) = 5.61, p = .06. Kruskal-Wallis H tests showed no differences between students who did and did not have permission to have their data released in terms of level of household income, H(1, N = 232) = 1.55, p = .21, first guardian's highest education, H(1, N = 289) = 0.75, p = .39, or highest level of education attained by either guardian, H(1, N = 289) = 0.002, p = .96. Finally, a one-way ANOVA showed no differences between these two groups in terms of age F(1, 284) = 1.05, p = .31. To test these differences while controlling for the interdependencies among students within classrooms, we conducted a Hierarchical Logistic Regression (HLR) using the Proc Glimmix procedure of SAS 9.0 in which a student's likelihood of receiving permission was predicted by demographic characteristics and program. The findings remained the same except that female students' parents were more likely to give permission to release student data than male students' parents (B = 0.75, SE = 0.29, p < .01). The HLR also found that students in the self-contained program were more likely to receive permission to have their data released than NGS students (B = 1.36, SE = 0.60, p < .05). The full results of the HLR are reported in Supplementary Table 3.

Psychological Well-Being Across Programs

HLMs that did not include SAI (reported in Table 4) showed that students in the self-contained program reported significantly higher (worse) levels of maladaptive perfectionism (B = 0.43, 95% CI [0.16, 0.70], SE = 0.14, p < .01), significantly lower (worse) loneliness scores (B

= -0.25, 95% CI [-0.46, -0.03], *SE* = 0.11, $p < .05$), and significantly lower (worse) math self-concepts (*B* = -0.44, 95% CI [-0.68, -0.20], *SE* = 0.12, $p < .001$) than students in the math pull-out program. Students in the NGS program reported lower (better) levels of maladaptive perfectionism (*B* = -0.44, 95% CI [-0.71, -0.16], *SE* = 0.14, $p < .01$) and higher (better) loneliness scores (*B* = 0.30, 95% CI [-0.52, -0.09], *SE* = 0.11, $p < .01$) than students in the self-contained program, but not students in the pull-out program. The only significant difference between students in the NGS program and students in the math pull-out program was that students in the NGS program reported significantly lower (worse) math self-concepts than students in the math pull-out program (*B* = -0.28, 95% CI [-0.52, -0.05], *SE* = 0.12, $p < .05$).

In terms of the effects of the covariates on the psychological well-being variables, fifth graders reported higher (better) levels of self-esteem (*B* = 0.28, 95% CI [0.001, 0.55], *SE* = 0.14, $p < .05$) than sixth graders, as well as lower (better) levels of maladaptive perfectionism (*B* = -0.40, 95% CI [-0.62, -0.18], *SE* = 0.11, $p < .001$) and higher (better) levels of school liking (*B* = 0.37, 95% CI [0.04, 0.69], *SE* = 0.17, $p < .05$). Additionally, Non-White students reported higher (worse) levels of school avoidance (*B* = 0.32, 95% CI [0.06, 0.58], *SE* = 0.13, $p < .05$) and higher (worse) levels of maladaptive perfectionism (*B* = 0.31, 95% CI [0.06, 0.56], *SE* = 0.13, $p < .05$) than White students. These results were robust whether or not the models were estimated based on multiple imputations or listwise deletion. The results of the HLMs using multiple imputation are reported in Supplementary Table 4.

When SAI was included as a covariate, the same trends were apparent, although some of the significant effects became nonsignificant likely due to a loss of power as a result of missing SAI data. In the HLMs that included SAI (reported in Table 5), students in the NGS program reported higher (better) math self-concepts (*B* = 0.48, 95% CI [0.03, 0.93], *SE* = 0.23, $p < .05$)

than students in the self-contained program. Furthermore, students in the math pull-out program reported higher (better) math self-concepts (B = 0.57, 95% CI [0.27, 0.87], SE = 0.15, $p < .001$) and lower (better) levels of maladaptive perfectionism (B = -0.41, 95% CI [-0.76, -0.05], SE = 0.18, $p < .05$) than students in the self-contained program. No differences were found between students in the math pull-out program and students in the NGS program.

Discussion

The results of this study showed that students enrolled in the two gifted programs with different service delivery models (self-contained vs. pull-out) reported different outcomes in terms of psychological well-being, specifically with respect to maladaptive perfectionism, math self-concept, and loneliness. Students enrolled in the gifted programs also reported some differences as compared to students in the NGS program. Students in the NGS program reported lower levels of maladaptive perfectionism and less loneliness than students in the self-contained program. Students in the math pull-out program only differed from students in the NGS program in their higher math self-concepts.

When comparing students in the two different gifted programs, students in the full-time, self-contained gifted program reported significantly lower math self-concepts than students in the math pull-out program whether SAI scores were included in the model or not. One possible explanation can be drawn from social comparison theory, and more specifically the big fish little pond effect (Marsh, 1987, 1991; Marsh et al., 1995; Marsh & Parker, 1984). Because the students in the self-contained program were constantly surrounded by other gifted students, this high reference group for comparison may have made them feel as if they were “less gifted” than their classmates. These upwards comparisons (i.e., “My classmates are better than me at math”) may have decreased students’ math self-concepts in the self-contained program. In contrast,

students in the math pull-out program had frequent opportunities to interact with and compare themselves to students not identified as gifted. The non-identified students provided opportunities for downward comparisons (i.e., “I’m better than my classmates at math”) that were likely to improve the pull-out students’ math self-concepts. Although math self-concept is typically thought of as being positively correlated with motivation and achievement (Lohbeck, 2016), it is worth noting that having too high of a math self-concept has the potential to impede academic development if it discourages a student from putting effort into learning (Roelle & Rankl, 2020). As such, it is possible that the self-contained students had appropriate math self-concepts while the pull-out students’ math self-concepts were inappropriately inflated. Although exploring this scenario is beyond the scope of this study, it reflects the challenges gifted programs face in simultaneously promoting academic achievement and psychological well-being.

In addition to lower math self-concepts, students in the self-contained program reported higher levels of maladaptive perfectionism than students in the pull-out program whether SAI was included in the analysis or not. The higher level of maladaptive perfectionism may have been driven by the labeling effect, as students in the self-contained program were labeled as the highest ability students in the district, a fact that may have motivated them to set unrealistically high expectations. Furthermore, the lack of a downward comparison group of non-identified students may have prevented the self-contained gifted students from being able to accurately assess their achievements and academic progress, which may have led to chronic feelings of inadequacy. This one-two punch of high expectations and a lack of reference to non-identified students may explain the differences in maladaptive perfectionism between students in the two gifted programs.

Lastly, students in the self-contained program reported being lonelier than students in the pull-out program when SAI was not included in the model. This difference in loneliness may have been a result of the self-contained students being isolated from students other than their self-contained classmates. Additionally, being labeled as the highest ability students in the district may have caused the self-contained students to feel excluded or to find it difficult to fit in with other social groups due to the stigmas and stereotypes associated with their giftedness (Baudson, 2016; Preckel et al., 2015). This social exclusion or isolation may have made it difficult for the self-contained students to develop meaningful friendships with other students, thus leading them to feel lonelier than students in the pull-out program.

When comparing students in each of the gifted programs to students in the NGS program, the results were mixed. Most of the differences were between students in the NGS program and gifted students in the self-contained program. Gifted students in the self-contained program reported higher levels of maladaptive perfectionism and more loneliness than students in the NGS program, which could be attributed to the abovementioned labeling and isolation effects. Students in the NGS program and students in the gifted pull-out program only differed in that the students in the pull-out program had higher math self-concepts, which is unsurprising considering the pull-out students' label as being gifted in math. Notably, there is no single variable on which gifted students in each of the two gifted programs significantly differed in the same direction when compared to students in the NGS program.

The pattern of results found when comparing NGS students to gifted students in the pull-out program could have differed from the pattern of results found when comparing NGS students to gifted students in the self-contained program for two reasons. On one hand, it could indicate that the two gifted programs were in fact differentially affecting the gifted students due to their

different service delivery models. On the other hand, the results could indicate that the students enrolled in the two different gifted programs were inherently different based on their ability levels or the domains of their giftedness. Though disentangling program effects from individual differences is beyond the scope of this study, these different patterns of results for students in the two different gifted programs suggest that gifted students receiving different gifted services should not be treated as a monolithic group. Instead, the relationship between giftedness and psychological well-being should be evaluated through a nuanced lens that accounts for individual differences, program effects, and broader environmental contexts, rather than simply dichotomizing students as either gifted or non-identified. This aligns with best practices suggesting that schools should offer a continuum of gifted services to meet all gifted students' needs (Corwith et al., 2019).

In addition to the program level differences, some demographic differences appeared in our data as well. Fifth graders reported lower levels of maladaptive perfectionism and higher levels of self-esteem and school liking than sixth graders, which aligns with research showing that primary school-aged children generally report higher well-being scores than middle school-aged children (Tobia et al., 2019) and that younger students report higher subjective well-being than older students (Liu et al., 2016). Additionally, Non-White students reported higher levels of school avoidance and maladaptive perfectionism than White students, which aligns with previous research documenting Non-White students' vulnerabilities in terms of psychological well-being during middle school (Akos & Galassi, 2004; Burchinal et al., 2008; Espinoza & Juvonen, 2011). Though the purpose of this study was not to evaluate the effects of demographic characteristics on psychological well-being, these results are an important reminder that future research on

various gifted service delivery models must be conscious of the interactions of these service delivery models with demographic characteristics.

The findings of this study were limited by a few factors. First and foremost, this study was conducted on a relatively small sample of students (N = 292) taught by 12 teachers across three buildings in one school district. As such, the results found in this study could have been impacted by individual teaching styles, curricular elements of the three different programs, the educational philosophies promoted from district leadership, or resources available to the programs. Relatedly, because the various service delivery models were evaluated holistically, this study was unable to tease apart the specific aspects of each model (e.g., curriculum, learning approaches, testing strategies, social opportunities) that may have been related to students' well-being. Furthermore, the sample only included teachers who taught gifted students, which may limit the generalizability of the findings to classrooms taught by teachers who don't have experience teaching gifted students. The findings of this study do not account for the perspectives, approaches, or attitudes of the teachers, which may have impacted the ways in which they taught their students.

Only a portion of guardians (n = 175; 60% of sample) permitted the district to release data to us, so intelligence measures and measures of academic achievement recorded by the district were not able to be used as covariates in the analyses without losing a great deal of statistical power, which is a particular concern because intelligence scores were one of the identification criteria for the two gifted programs. Additionally, this study was conducted at one timepoint, and the lack of a longitudinal component prevented us from measuring changes in students' psychological well-being as they matriculated through each of the programs, thus making causal claims impossible. Finally, the fact that students were placed in each of the

programs based on ability level means that the findings of this study may be driven by inherent differences in the types of students who qualified for each of the programs, making it impossible to disentangle true program effects and individual differences between the students who were enrolled in the programs.

Despite these limitations, the results of this study serve as an important starting point for acknowledging the differences in psychological well-being between students enrolled in gifted programs using different service delivery models. Since students across the country are engaged in such a wide variety of gifted programs, it is vital to continue evaluating the ways in which students in various gifted programs differ and the ways in which programs with different service delivery models interact with these individual differences to impact students' psychological well-being. In these evaluations, however, it is important to acknowledge that psychological well-being is only one of the many factors that LEAs evaluate when choosing which gifted programs to offer (Kettler et al., 2015) and to recognize that the other benefits of gifted programming may come at the expense of psychological well-being. Future research on this topic should have four major goals. First, future studies should implement longitudinal or quasi-experimental designs that allow for the disentanglement of individual differences and program effects. Second, future studies should evaluate the psychological well-being of gifted students enrolled in programs using other service delivery models (e.g., cluster classrooms, grade acceleration). Third, future research should more thoroughly integrate the perspectives of teachers and parents to better understand the impacts that gifted programs with different service delivery models have on classrooms and families. Finally, future studies should replicate the current findings using alternative measures of psychological well-being, such as the Student Subjective Wellbeing Questionnaire (Renshaw, 2020) and the Psychological Wellbeing Scale (Ryff & Keyes, 1995).

Understanding the ways in which gifted students' unique needs interact with various gifted service delivery models is essential to building gifted programs that help students to thrive not only academically, but socially and emotionally as well.

References

Adelson, J., McCoach, D. B., & Gavin, M. K. (2012). Examining the effects of gifted programming in mathematics and reading using the ECLS-K. *Gifted Child Quarterly, 56*(1), 25-39. https://doi.org/10.1177/0016986211431487

Akos, P. & Galassi, J.P. (2004). Gender and race as variables in psychosocial adjustment in middle and high school. *The Journal of Educational Research, 98,* 102-108. https://doi.org/10.3200/JOER.98.2.102-108

Alfasi, Y. (2019). The grass is always greener on my friends' profiles: The effect of Facebook social comparisons on state self-esteem and depression. *Personality and Individual Differences, 147,* 111-117. https://doi.org/10.1016/j.paid.2019.04.032

Amini, M. (2005). Identifying stressors and reactions to stressors in gifted and non-gifted students. *International Education Journal, 6,* 136-140.

Asher, S. R., Hymel, S., & Renshaw, P. D. (1984). Loneliness in children. *Child Development, 55*(4), 1456-1464. https://doi.org/10.2307/1130015

Asher, S. R., & Paquette, J. (2003). Loneliness and peer relations in children. *Current Directions in Psychological Science, 12*(3), 75-78. https://doi.org/10.1111/1467-8721.01233

Bagner, D. M., Storch, E. A., & Roberti, J. W. (2004). A factor analytic study of the Loneliness and Social Dissatisfaction Scale in a sample of African-American and Hispanic-American children. *Child Psychiatry & Human Development, 34*(3), 237-250. https://doi.org/10.1023/B:CHUD.0000014999.16111.2f

Bai, H., Wang, L., Pan, W., & Frey, M. (2009). Measuring mathematics anxiety: Psychometric analysis of a bidimensional affective scale. *Journal of Instructional Psychology, 36*(3), 185-193.

Bajaj, B., Gupta, R., & Pande, N. (2016). Self-esteem mediates the relationship between mindfulness and well-being. *Personality and Individual Differences, 94*, 96-100. https://doi.org/10.1016/j.paid.2016.01.020

Baudson, T. G. (2016). The mad genius stereotype: Still alive and well. *Frontiers in Psychology, 7*, 368. https://doi.org/10.3389/fpsyg.2016.00368

Becker, H. S. (1963). *Outsiders: Studies in the sociology of deviance.* Free Press of Glencoe.

Bergold, S., Wirthwein, L., & Steinmayr, R. (2018). Subjective well-being of intellectually gifted children and adolescents. In I. Gonzalez-Burgos (Ed.), *Psychobiological, Clinical, and Educational Aspects of Giftedness* (143-165). Nova Science.

Burchinal, M. R., Roberts, J. E., Zeisel, S. A., & Rowley, S. J. (2008). Social risk and protective factors for African American children's academic achievement and adjustment during the transition to middle school. *Developmental Psychology, 44*, 286-292. https://doi.org/10.1037/0012-1649.44.1.286

Chambers, B., Murray, C. M., Boden, Z. V. R., & Kelly, M. P. (2019). 'Sometimes labels need to exist': Exploring how young adults with Asperger's syndrome perceive its removal from the *Diagnostic and Statistical Manual of Mental Disorders* fifth edition. *Disability & Society*, 1-20. https://doi.org/10.1080/09687599.2019.1649121

Chen, X., Liu, J., Ellis, W., & Zarbatany, L. (2016). Social sensitivity and adjustment in Chinese and Canadian children. *Child Development*, *87*(4), 1115-1129. https://doi.org/10.1111/cdev.12514

Cheng, H., & Furnham, A. (2002). Personality, peer relations, and self-confidence as predictors of happiness and loneliness. *Journal of Adolescence, 25*, 327-339.

Clark, J. K., Thiem, K. C., & Kang, S. (2017). Positive stereotype validation: The bolstering effects of activating positive stereotypes after intellectual performance. *Personality and Social Psychology Bulletin, 43*(12), 1630-1642. https://doi.org/10.1177/0146167217723324

Corcoran, K., Crusius, J., & Mussweiler, T. (2011). Social comparison: Motives, standards, and mechanisms. In D. Chadee (Ed.), *Theories in Social Psychology* (pp. 119-139). Wiley-Blackwell.

Corwith, S., Johnsen, S., Lee, C. -W., Cotabish, A., Dailey, D., & Guilbault, K. (2019). *2019 Pre-K – Grade 12 Gifted Programming Standards.* National Association for Gifted Children Professional Standards Committee. https://www.nagc.org/sites/default/files/standards/Intro%202019%20Programming%20Standards%281%29.pdf

Crocker, J., Major, B., & Steele, C. (1998). Social stigma. In S. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, pp. 504-553). McGraw Hill.

Dijkstra, P., Kuyper, H., Van Der Werf, G., Buunk, A. P., & Van Der Zee, Y. G. (2008). Social comparison in the classroom: A review. *Review of Educational Research, 78*(4), 828-879. https://doi.org/10.3102/0034654308321210

Dimitriadis, C. (2012). Provision for mathematically gifted children in primary schools: An investigation of four different methods of organizational provision. *Educational Review, 64*, 241-260. https://doi.org/10.1080/00131911.2011.598920

Dimitriadis, C. (2016). Gifted programs cannot be successful without gifted research and theory. *Journal for the Education of the Gifted*, *39*(3), 221-236. https://doi.org/10.1177/0162353216657185

Ebesutani, C., Drescher, C. F., Reise, S. P., Heiden, L., Hight, T. L., Damon, J. D., & Young, J. (2012). The importance of modeling method effects: Resolving the (uni)dimensionality of the Loneliness Questionnaire. *Journal of Personality Assessment, 94*(2), 186-195. https://doi.org/10.1080/00223891.2011.627967

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). W. H. Freeman.

Eccles, J. S., Adler, T. F., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology, 46,* 26-43. https://doi.org/10.1037/0022-3514.46.1.26

Eccles, J. S., O'Neill, S. A., & Wigfield, A. (2005). Ability self-perceptions and subjective task values in adolescents and children. In K. A. Moore & L. H. Lippman (Eds.), *What do children need to flourish: Conceptualizing and measuring indicators of positive development* (pp. 237-249). Springer Science + Business Media. https://doi.org/10.1007/0-387-23823-9_15

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*(3), 215-225. https://doi.org/10.1177/0146167295213003

Espinoza, G., & Juvonen, J. (2011). Perceptions of the school social context across the transition to middle school: Heightened sensitivity among Latino students? *Journal of Educational Psychology, 103,* 749-758. https://doi.org/10.1037/a0023811

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117-140. https://doi.org/10.1177/001872675400700202

Fleming, J. S., & Courtney, B. E. (1984). The dimensionality of self-esteem: II hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology, 46*(2), 404-421. https://doi.org/10.1037/0022-3514.46.2.404

Fletcher, K. L., & Speirs Neumeister, K. L. (2012). Research on perfectionism and achievement motivation for gifted students. *Psychology in the Schools, 49*(7), 668-677. https://doi.org/10.1002/pits.21623

Flett, G. L., & Hewitt, P. L. (2006). Positive versus negative perfectionism in psychopathology: A comment on Slade and Owens's Dual Process Model. *Behavioral Modification, 30*(4), 472-495. https://doi.org/10.1177/0145445506288026

Flett, G. L., & Hewitt, P. L. (2014). A froposed framework for preventing perfectionism and promoting resilience and mental health among vulnerable children and adolescents. *Psychology in the Schools, 51*(9), 899-912. https://doi.org/10.1002/pits.21792

Glock, S. & Krolak-Schwerdt, S. (2014). Stereotype activation versus application: How teachers process and judge information about students from ethnic minorities and with low socioeconomic background. *School Psychology of Education: An International Journal, 17*(4), 589-607. https://doi.org/10.1007/s11218-014-9266-6

Godor, B. P., & Szymanski, A. (2017). Sense of belonging or feeling marginalized? Using PISA 2012 to assess the state of academically gifted students within the EU. *High Ability Studies, 28*(2), 181-197. https://doi.org/10.1080/13598139.2017.1319343

Gove, W. R. (1980). *Labeling and Juvenile Delinquency*. Sage.

Gray-Little, B., Williams, V. S. L., & Hancock, T. D. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 23*(5), 443-451. https://doi.org/10.1177/0146167297235001

Green, S. E. (2007). Components of perceived stigma and perceptions of well-being among university students with and without disability experience. *Health Sociology Review, 16*(3-4), 328-340. https://doi.org/10.5172/hesr.2007.16.3-4.328

Gross, M. U. M. (2001). Ability grouping, self-esteem, and the gifted: A study of optical illusions and optimal environments. In N. Colangelo & S. G. Assouline (Eds.), *Talent development IV: Proceedings from the 1998 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 59-88). Great Potential Press.

Gross, M. U. M. (2002). Social and emotional issues for exceptionally intellectually gifted students. In M. Neihart, S. M. Reis, N. M. Robinson, & S. M. Moon (Eds.), *The social and emotional development of gifted children: What do we know?* (pp. 19-30). Prufrock Press.

Grzegorek, J. L., Slaney, R. B., Franze, S., & Rice, K. G. (2004). Self-criticism, dependency, self-esteem, and grade point average satisfaction among clusters of perfectionists and nonperfectionists. *Journal of Counseling Psychology, 51*, 192-200. https://doi.org/10.1037/0022-0167.51.2.192

Herman, K. C., Wang, K., Trotter, R., Reinke, W. M., & Ialongo, N. (2013). Developmental trajectories of maladaptive perfectionism among African American adolescents. *Child Development, 84*(5), 1633-1650. https://doi.org/10.1111/cdev.12078

Herrmann, J., Schmidt, I., Kessels, U., & Preckel, F. (2016). Big fish in big ponds: Contrast and assimilation effects on math and verbal self-concepts of students in within-school gifted tracks. *British Journal of Educational Psychology, 86*(2), 222-240. https://doi.org/10.1111/bjep.12100

Honjo, S., Sasaki, Y., Kaneko, H., Tachibana, K., Murase, S., Ishii, T., Nishide, Y., & Nishide, T. (2003). Study on feelings of school avoidance, depression, and character tendencies among general junior high and high school students. *Psychiatry and Clinical Neurosciences*, *57*(5), 464-471. https://doi.org/10.1046/j.1440-1819.2003.01149.x

Honma, Y., & Uchiyama, I. (2014). Emotional engagement and school adjustment in late childhood: The relationship between school liking and school belonging in Japan. *Psychological Reports*, *114*(2), 496-508. https://doi.org/10.2466/21.10.PR0.114k19w7

Hutzell, K. L., & Payne, A. A. (2012). The impact of bullying victimization on school avoidance. *Youth Violence and Juvenile Justice*, *10*(4), 370-385. https://doi.org/10.1177/1541204012438926

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology*, *83*(3), 693-710. https://doi.org/10.1037/0022-3514.83.3.693

Kamarova, S., Chatzisarantis, N. L. D., Hagger, M. S., Lintunen, T., Hassandra, M., & Papaioannou, A. (2017). Effects of achievement goals on perceptions of competence in conditions of unfavourable social comparisons: The mastery goal advantage effect. *British Journal of Educational Psychology*, *87*(4), 630-646. https://doi.org/10.1111/bjep.12168

Kettler, T., Russell, J., & Puryear, J. S. (2015). Inequitable access to gifted education: Variance in funding and staffing based on locale and contextual school variables. *Journal for the Education of the Gifted*, *38*(2), 99-117. https://doi.org/10.1177/0162353215578277

Killen, M., Mulvey, K. L., & Hitti, A. (2013). Social exclusion in childhood: A developmental intergroup perspective. *Child Development, 84*(3), 772-790. https://doi.org/10.1111/cdev.12012

Krinzinger, H., Kaufmann, L., & Willmes, K. (2009). Math anxiety and math ability in early primary school years. *Journal of Psychoeducational Assessment, 27*(3), 206-225. https://doi.org/10.1177/0734282908330583

Kroesbergen, E. H., van Hooijdonk, M., van Viersen, S., Middel-Lalleman, M. M. N, & Reijnders, J. J. W. (2015). The psychological well-being of early identified gifted children. *Gifted Child Quarterly, 60*(1), 16-30. https://doi.org/10.1177/0016986215609113

Kulik, C.-L. C., & Kulik, J. A. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal, 19*(3), 415-428. https://doi.org/10.3102/00028312019003415

Kulik, J.A. & Kulik, C.-L. C. (1992). Meta-analytic findings on grouping programs. *Gifted Child Quarterly, 36*(2), 73-77. https://doi.org/10.1177/001698629203600204

Ladd, G. W., Buhs, E. S., & Seid, M. (2000). Children's initial sentiments about kindergarten: Is school liking an antecedent of early classroom participation and achievement? *Merrill-Palmer Quarterly, 46*(2), 255-279.

Ladd, G. W., Kochenderfer, B. J., & Coleman, C. C. (1996). Friendship quality as a predictor of children's early school adjustment. *Child Development, 67*(3), 1103-1118. https://doi.org/10.2307/1131882

Ladd, G. W., & Price, J. M. (1987). Predicting children's social and school adjustment following the transition from preschool to kindergarten. *Child Development, 58*(5), 1168-1189. https://doi.org/10.1111/j.1467-8624.1987.tb01450.x

Lam, L. T., & Kirby, S. L. (2002). Is emotional intelligence an advantage? An exploration of the impact of emotional and general intelligence on individual performance. *The Journal of Social Psychology, 142*, 133-143. https://doi.org/10.1080/00224540209603891

Lea-Wood, S. S., & Clunies-Ross, G. (1995). Self-esteem of gifted adolescent girls in Australian schools. *Roeper Review: A Journal on Gifted Education, 17*(3), 195-197. https://doi.org/10.1080/02783199509553658

Lee, S.-Y., Olzewski-Kubilius, P., & Turner Thomson, D. (2012). Academically gifted students' perceived interpersonal competence and peer relationships. *Gifted Child Quarterly, 56*(2), 90-104. https://doi.org/10.1177/0016986212442568

Litster, K., & Roberts, J. (2010). The self-concepts and perceived competencies of gifted and non-gifted students: A meta-analysis. *Journal of Research in Special Educational Needs, 11*(2), 130-140. https://doi.org/10.1111/j.1471-3802.2010.01166.x

Liu, W., Mei, J., Tian, L., & Huebner, E. S. (2016). Age and gender differences in the relation between school-related social support and subjective well-being in school among students. *Social Indicators Research, 125*(3), 1065-1083. https://doi.org/10.1007/s11205-015-0873-1

Lohbeck, A. (2016). Self-Concept and self-determination theory: Math self-concept, motivation, and grades in elementary school children. *Early Child Development and Care, 188*(8), 1031-1044. https://doi.org/10.1080/03004430.2016.1241778

Ludeke, S. G., Gensowski, M., Junge, S. Y., Kirkpatrick, R. M., John, O. P., & Andersen, S. C. (2020). Does parental education influence child educational outcomes? A developmental analysis in a full-population sample and adoptee design. *Journal of Personality and Social Psychology.* Advance online publication. https://doi.org/10.1037/pspp0000314

Margot, K. C., & Rinn, A. N. (2016). Perfectionism in gifted adolescents: A replication and extension. *Journal of Advanced Academics, 27*(3), 190-209. https://doi.org/10.1177/1932202X16656452

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology, 79,* 280-295. https://doi.org/10.1037/0022-0663.79.3.280

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. https://doi.org/10.3102/00028312028002445

Marsh, H. W., Chessor, D., Craven, R., & Roche, L. (1995). The effects of gifted and talented programs on academic self-concept: The big fish strikes again. *American Educational Research Journal, 32*(2), 285-319. https://doi.org/10.3102/00028312032002285

Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology, 47*(1), 213-231. https://doi.org/10.1037/0022-3514.47.1.213

Marsh, H. W., Scalas, L. F., & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem Scale: Traits, ephemeral artifacts, and stable response styles. *Psychological Assessment, 22,* 366-381. https://doi.org/10.1037/a0019225

Matthews, N. L., Ly, A. R., & Goldberg, W. A. (2015). College students' perceptions of peers with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(1), 90-99. https://doi.org/10.1007/s10803-014-2195-6

Mcclain, M.-C., & Pfeiffer, S. (2012). Identification of gifted students in the United States today: A look at state definitions, policies, and practices. *Journal of Applied School Psychology, 28*(1), 59-88. https://doi.org/10.1080/15377903.2012.643757

McKay, M. T., Boduszek, D., & Harvey, S. A. (2014). The Rosenberg Self-Esteem Scale: A bifactor answer to a two-factor question? *Journal of Personality Assessment, 96*(6), 654-660. https://doi.org/10.1080/00223891.2014.923436

Meeus, A., Beullens, K., & Eggermont, S. (2019). Like me (please?): Connecting online self-presentation to pre- and early adolescents' self-esteem. *New Media & Society, 21*(11-12), 2386-2403. https://doi.org/10.1177/1461444819847447

Mobley, M., Slaney, R. B., & Rice, K. G. (2005). Cultural validity of the Almost Perfect Scale-Revised for African American college students. *Journal of Counseling Psychology, 52*, 629-639. https://doi.org/10.1037/0022-0167.52.4.629

Moeller, R. W., & Seehuus, M. (2019). Loneliness as a mediator for college students' social skills and experiences of depression and anxiety. *Journal of Adolescence, 73*, 1-13. https://doi.org/10.1016/j.adolescence.2019.03.006

Moon, S. M., Swift, M., & Shallenberger, A. (2002). Perceptions of a self-contained class for fourth- and fifth-grade students with high to extreme levels of intellectual giftedness. *Gifted Child Quarterly, 46*(1), 64-79. https://doi.org/10.1177/001698620204600106

Moses, T. (2009). Self-labeling and its effects among adolescents diagnosed with mental disorders. *Social Science and Medicine, 68*(3), 570-578. https://doi.org/10.1016/j.socscimed.2008.11.003

National Association for Gifted Children (n.d.-a). *Gifted Education Strategies*. https://www.nagc.org/resources-publications/gifted-education-practices

National Association for Gifted Children (n.d.-b) *Pull-Out Programs/Specialized Classes*. https://www.nagc.org/resources-publications/gifted-education-practices/pull-out-programsspecialized-classes

National Association for Gifted Children & The Council of State Directors of Programs for the Gifted (2015). *2014-2015 State of the states in gifted education: National Policy and Practice Data.* http://www.nagc.org/sites/default/files/key%20reports/2014-2015%20State%20of%20the%20States%20%28final%29.pdf

Neihart, M. (1999). The impact of giftedness on psychological well-being: What does the empirical literature say? *Roeper Review, 22,* 10-17. https://doi.org/10.1080/02783199909553991

Norwich, B., & Kelly, N. (2004). Pupils' views on inclusion: Moderate learning difficulties and bullying in mainstream and special schools. *British Educational Research Journal, 30*(1), 43-65. https://doi.org/10.1080/0141192031000162996 5

Ogurlu, U., Yalin, H. S., & Yavuz Birben, F. (2018). The relationship between psychological symptoms, creativity, and loneliness in gifted children. *Journal for the Education of the Gifted, 41*(2), 193–210. https://doi.org/10.1177/0162353218763968

Ohan, J. L., Visser, T. A. W., Strain, M. C., & Allen, L. (2011). Teachers' and education students' perceptions of and reactions to children with and without the diagnostic label

"ADHD." *Journal of School Psychology*, *49*(1), 81-105. https://doi.org/10.1016/j.jsp.2010.10.001

Ortega, N. E., Wang, K. T., Slaney, R. B., Hayes, J. A., & Morales, A. (2014). Personal and familial aspects of perfectionism in Latino/a students. *The Counseling Psychologist*, *42*, 406-427. https://doi.org/10.1177/0011000012473166

Özbay, Y., & Palanci, M. (2011). Psychosocial characteristics of gifted children and adolescents. *Sakarya Üniversitesi Eğitim Fakültesi Dergisi [Sakarya University Journal of Education]*, *22*, 89-108

Park, H. J., & Jeong, D. Y. (2015). Psychological well-being, life satisfaction, and self-esteem among adaptive perfectionists, maladaptive perfectionists, and nonperfectionists. *Personality and Individual Differences*, *72*, 165-170. https://doi.org/10.1016/j.paid.2014.08.031

Plucker, J. A., & Dilley, A. (2016). Ability grouping and the socioemotional development of gifted students. In M. Neihart, S. I. Pfeiffer, & T. L. Cross (Eds.), *The social and emotional development of gifted children: What do we know?* 2nd ed. (pp. 231-241). Prufrock Press.

Pollard, E. L., & Lee, P. D. (2003). Child well-being: A systematic review of the literature. *Social Indicators Research, 61*, 59-78. https://doi.org/10.1023/A:1021284215801

Preckel, F., Baudson, T. K., Krolak-Schwerdt, S., & Glock, S. (2015). Gifted and maladjusted? Implicit attitudes and automatic associations related to gifted children. *American Educational Research Journal, 52*(6), 1160-1184. http://doi.org/10.3102/0002831215596413

Preckel, F., Götz, T., & Frenzel, A. (2010). Ability grouping of gifted students: Effects on academic self-concept and boredom. *British Journal of Educational Psychology, 80*(3), 451-472. https://doi.org/10.1348/000709909X480716

Preckel, F., Zeidner, M., Goetz, T., & Schleyer, E. J. (2008). Female "big fish" swimming against the tide: The "big-fish-little-pond effect" and gender-ratio in special gifted classes. *Contemporary Educational Psychology, 33*(1), 78-96. https://doi.org/10.1016/j.cedpsych.2006.08.001

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models* (2nd ed.). Newbury Park, CA: Sage.

Renshaw, T. L. (2020). *Student subjective wellbeing questionnaire (SSWQ): Measure and user guide.* Open Science Framework. https://osf.io/k25yf/

Rice, K. G., & Ray, M. E. (2018). Perfectionism and the gifted. In S. I. Pfeiffer, E. Shaunessy-Dedrick, & M. Foley-Nicpon (Eds.), *APA handbook of giftedness and talent* (pp. 645–658). American Psychological Association. https://doi.org/10.1037/0000038-042

Rice, K. G., Richardson, C. M., & Tueller, S. (2014). The Short Form of the Revised Almost Perfect Scale. *Journal of Personality Assessment, 96*(3), 368-379. https://doi.org/10.1080/00223891.2013.838172

Rice, K. G., & Slaney, R. B. (2002). Clusters of perfectionists: Two studies of emotional adjustment and academic achievement. *Measurement and Evaluation in Counseling and Development, 35*, 35-48. https://doi.org/10.1080/07481756.2002.12069046

Roelle, J., & Renkl, A. (2020). Does an option to review instructional explanations enhance example-based learning? It depends on learners' academic self-concept. *Journal of Educational Psychology, 112*(1), 131-147. https://doi.org/10.1037/edu0000365

Rönkä, A. R., Sunnari, V., Rautio, A., Koiranen, M., & Taanila, A. (2017). Associations between school liking, loneliness and social relations among adolescents: Northern Finland Birth Cohort 1986 study. *International Journal of Adolescence and Youth, 22*(1), 93-106. https://doi.org/10.1080/02673843.2015.1136659

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press. https://doi.org/10.1515/9781400876136

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. Rinehart and Winston.

Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology, 69*(4), 719-727.

Salimi, A. (2011). Social-emotional loneliness and life satisfaction. *Procedia-Social and Behavioral Sciences, 29,* 292-295. https://doi.org/10.1016/j.sbspro.2011.11.241

Seo, E., Shen, Y., & Alfaro, E. C. (2019). Adolescents' beliefs about math ability and their relations to STEM career attainment: Joint consideration of race/ethnicity and gender. *Journal of Youth and Adolescence, 48*(2), 306-325. https://doi.org/10.1007/s10964-018-0911-9

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407-441. https://doi.org/10.3102/00346543046003407

Shechtman, Z., & Silektor, A. (2012). Social competencies and difficulties of gifted children compared to nongifted peers. *Roeper Review, 34*, 63-72. https://doi.org/10.1080/02783193.2012.627555

Slaney, R. B., Rice, K. G., Mobley, M., Trippi, J., & Ashby, J. S. (2001). The Revised Almost Perfect Scale. *Measurement and Evaluation in Counseling and Development, 34*(3), 130-145. https://doi.org/10.1080/07481756.2002.12069030

Steers, M. L. N. (2016). "It's complicated": Facebook's relationship with the need to belong and depression. *Current Opinion in Psychology, 9,* 22-26. https://doi.org/10.1016/j.copsyc.2015.10.007

Stoeber, J., & Otto, K. (2006). Positive conceptions of perfectionism: Approaches, evidence, challenges. *Personality and Social Psychology Review*, *10*(4), 295-319. https://doi.org/10.1207/s15327957pspr1004_2

Stricker, J., Buecker, S., Schneider, M., & Preckel, F. (2019). Intellectual giftedness and multidimensional perfectionism: A meta-analytic review. *Educational Psychology Review, 32,* 391-414. https://doi.org/10.1007/s10648-019-09504-1

Subotnik, R. F., Stoeger, H., & Olszewski-Kubilius, P. (2017). Talent development research, policy, and practice in Europe and the United States: Outcomes from a summit of international researchers. *Gifted Child Quarterly, 61*(3), 262-269. https://doi.org/10.1177/0016986217701839

Supple, A. J., Su, J., Plunkett, S. W., Peterson, G. W., & Bush, K. R. (2013). Factor structure of the Rosenberg Self-Esteem Scale. *Journal of Cross-Cultural Psychology*, *44,* 748-764. https://doi.org/10.1177/0022022112468942

Sutin, A. R., Stephan, Y., & Terracciano, A. (2018). Parental educational attainment and offspring subjective well-being and self-beliefs in older adulthood. *Personality and Individual Differences, 128,* 139-145. https://doi.org/10.1016/j.paid.2018.01.023

Swanson, J., Valiente, C., & Lemery-Chalfant, K. (2012). Predicting academic achievement from cumulative home risk: The mediating roles of effortful control, academic relationships, and school avoidance. *Merrill-Palmer Quarterly*, *58*(3), 375-408. https://doi.org/10.1353/mpq.2012.0014

Thøgersen-Ntoumani, C., Ntoumanis, N., Carey, R., Dodos, L., Quested, E. J., & Chatzisarantis, N. (2018). A diary study of appearance social comparisons and need frustration in young women. *Personality and Individual Differences*, *122*, 120–126. https://doi.org/10.1016/j.paid.2017.10.020

Thomson, M. M. (2012). Labelling and self-esteem: Does labelling exceptional students impact their self-esteem? *British Journal of Learning Support, 27*(4), 158-165. https://doi.org/10.1111/1467-9604.12004

Tobia, V., Steca, P., & Marzocchi, G. M. (2019). Children's wellbeing at school: A multi-dimensional and multi-informant approach. *Journal of Happiness Studies, 20*(3), 841-861. https://doi.org/10.1007/s10902-018-9974-2

Tottenham, N., & Galván, A. (2016). Stress and the adolescent brain: Amygdala-prefrontal cortex circuitry and ventral striatum as developmental targets. *Neuroscience & Biobehavioral Reviews, 70*, 217-227. https://doi.org/10.1016/j.neubiorev.2016.07.030

van der Meulen, R. T., van der Bruggen, C. O., Spilt, J. L., Verouden, J., Berkhout, M., & Bögels, S. M. (2013). The pullout program day a week school for gifted children: Effects on social-emotional and academic functioning. *Child and Youth Care Forum*, *43*(3), 287-314. https://doi.org/10.1007/s10566-013-9239-5

Vialle, W., Heaven, P. C. L., & Ciarrochi, J. (2007). On being gifted, but sad and misunderstood: Social, emotional, and academic outcomes of gifted students in the Wollongong Youth

Study. *Educational Research and Evaluation*, *13*(6), 569-586. https://doi.org/10.1080/13803610701786046

Vidergor, H. E., & Azar Gordon, L. (2015). The case of a self-contained elementary classroom for the gifted: Student, teacher, and parent Perceptions of existing versus desired teaching–learning aspects. *Roeper Review*, *37*(3), 150-164. https://doi.org/10.1080/02783193.2015.1047549

Vogl, K., & Preckel, F. (2014). Full-time ability grouping of gifted students: Impacts on social self-concept and school-related attitudes. *Gifted Child Quarterly, 58*(1), 51-68. https://doi.org/10.1177/0016986213513795

Walton, G. M., & Carr, P. B. (2012). Social belonging and the motivation and intellectual achievement of negatively stereotyped students. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application* (pp. 89-106). Oxford University Press.

Wang, K. T., Permyakova, T. M., & Sheveleva, M. S. (2016). Assessing perfectionism in Russia: Classifying perfectionists with the Short Almost Perfect Scale. *Personality and Individual Differences*, *92*, 174-179. https://doi.org/10.1016/j.paid.2015.12.044

Wong, M. S. W., & Watkins, D. (2001). Self-esteem and ability grouping: A Hong Kong investigation of the big fish little pond effect. *Educational Psychology*, *21*(1), 79-87. https://doi.org/10.1080/01443410123082

Wong, T. K. Y., Konishi, C., & Tao, L. (2019). A social-emotional pathway to promoting math self-concept: The moderating role of sex. *Educational Psychology, 39*(9), 1119-1135. https://doi.org/10.1080/01443410.2019.1621994

Wu, J., & Cheung, H. Y. (2012). The relations of academic cheating with academic self-concept and perceived autonomy support among Chinese university students. In K. Gana (Ed.), *Psychology of self-concept* (pp. 61–71). Nova Science.

Yang, H., Guo, W., Yu, S., Chen, L., Zhang, H., Pan, L., & Wang, C. (2016). Personal and family perfectionism in Chinese school students: Relationships with learning stress, learning satisfaction and self-reported academic performance level. *Journal of Child and Family Studies, 25*(12), 3675-3683. https://doi.org/10.1007/s10826-016-0524-4

Yang, Q., Tian, L., Huebner, E. S., & Zhu, X. (2019). Relations among academic achievement, self-esteem, and subjective well-being in school among elementary school students: A longitudinal mediation model. *School Psychology Quarterly, 34*(3), 328-340. https://doi.org/10.1037/spq0000292

Zhang, B., Gao, Q., Fokkema, M., Alterman, V., & Liu, Q. (2015). Adolescent interpersonal relationships, social support and loneliness in high schools: Mediation effect and gender differences. *Social Science Research, 53*, 104-117. https://doi.org/ 10.1016/j.ssresearch.2015.05.003

Table 1

Students' Demographic Information by Program

Variable	Missing Rate	NGS Students (*n* = 99)	Pull-Out Students (*n* = 103)	Self-Contained Students (*n* = 90)
Percentage of 5th graders	0.00%	54.55%	44.67%	38.89%
Percentage of Female	2.05%	58.59%	45.63%	58.89%
Age[a] (*SD*)	2.05%	11.67 (0.55)	11.76 (0.62)	11.67 (0.75)
Median Household Income[b]	20.55%	$110 - $120	$110 - $120	$130 - $140
Race	3.08%			
White		65.31%	74.49%	78.16%
Black		15.31%	11.22%	5.75%
Multi-Racial		12.24%	8.16%	11.49%
Asian		4.08%	4.08%	3.45%
Other Race		3.06%	2.04%	1.15%
Hispanic	2.05%	4.10%	1.00%	3.40%
Guardians' Highest Education	1.03%			
High School Diploma		15.14%	12.44%	8.14%
Vocational School		8.65%	8.46%	4.07%
Associate's or Bachelor's Degree		42.70%	54.23%	55.23%
Master's Degree		22.16%	20.90%	27.33%
Doctoral Degree		5.41%	3.98%	5.23%
Unknown/Not Applicable		5.95%	0.00%	0.00%
GPA (*SD*)	57.19%	3.55 (0.37)	3.73 (0.26)	3.57 (0.33)
SAI Score (*SD*)	40.41%	105.55 (12.30)	122.78 (6.76)	130.25 (6.67)
Reading Test Score (*SD*)	40.07%	742.19 (31.28)	751.19 (27.63)	771.67 (30.78)
Math Test Score (*SD*)	40.41%	745.46 (27.53)	769.67 (24.55)	763.44 (20.13)
Superior Cognitive Ability	40.07%	2.08%	22.41%	98.55%
Gifted Math	40.07%	8.33%	98.28%	100.00%
Gifted Reading	40.07%	16.67%	53.45%	100.00%

[a] In Years
[b] In Thousand Dollars
Note. SAI = School Ability Index

Table 2

Means and Standard Deviations of Psychological Well-Being Measures (*N* = 292)

Variable	Missing Rate[1]	Total		NGS		Pull-Out		Self-Contained	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
[a]Loneliness	3.42%	4.03	0.73	4.15	0.70	4.09	0.64	3.83	0.83
[a]School Liking	3.08%	3.31	1.00	3.29	1.01	3.22	1.00	3.43	0.99
[b]School Avoidance	3.42%	3.10	1.00	3.22	1.02	3.19	0.97	2.86	0.98
[a]Self-Esteem	4.79%	3.61	0.89	3.73	0.85	3.71	0.85	3.37	0.95
[a]Adaptive Perfectionism	3.08%	3.83	0.83	3.74	0.81	3.80	0.82	3.94	0.85
[b]Maladaptive Perfectionism	3.08%	2.81	0.98	2.66	0.91	2.66	0.87	3.13	1.11
[a]Math Self-Concept	2.40%	3.79	0.85	3.74	0.89	4.01	0.67	3.58	0.93

[1] Students are considered missing if they miss any individual item on the measure for a given variable.
[a] Higher scores indicate higher psychological well-being
[b] Higher scores indicate lower psychological well-being

Table 3

Correlations between Psychological Well-Being Measures

Variable	L	SE	AP	MP	MSC	SL	SA
L	-	-	-	-	-	-	-
SE	.68**	-	-	-	-	-	-
AP	.13*	.13*	-	-	-	-	-
MP	-.49**	-.74**	.13*	-	-	-	-
MSC	.30**	.42**	.32**	-.29**	-	-	-
SL	.31**	.30**	.23*	-.20**	.29**	-	-
SA	-.10	-.13*	-.22**	.08	-.20**	-.75**	-

Note. Reported values are Pearson correlation coefficients; Variables are abbreviated as follows: L = Loneliness; SE = Self-Esteem; AP = Adaptive Perfectionism; MP = Maladaptive Perfectionism; MSC = Math Self-Concept; SL = School Liking; SA = School Avoidance. *p<.05, **p <.01

Table 4

Hierarchical Linear Models Predicting Psychological Well-Being by Program Type

Variable	Adaptive Perfectionism		Maladaptive Perfectionism		School Liking		School Avoidance	
ICC	.01		.08		.12		.09	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Student-Level								
Intercept	3.79***	0.19	3.28***	0.22	3.35***	0.26	3.23***	0.25
Gender (Female vs. Male)	0.04	0.10	-0.04	0.11	0.18	0.11	-0.20	0.12
Race (White vs. Non-White)	-0.16	0.11	-0.31*	0.13	0.21	0.13	-0.32*	0.13
Guardians' Highest Education	0.07	0.05	-0.06	0.06	0.01	0.06	-0.04	0.06
Classroom-Level								
Grade (6^{th} vs. 5^{th})	0.12	0.10	0.40***	0.11	-0.37*	0.17	0.19	0.15
Pull-Out vs. Self-Contained	-0.08	0.12	-0.43**	0.14	-0.19	0.21	0.27	0.20
Pull-Out† vs. NGS	0.13	0.12	0.01	0.13	-0.08	0.20	-0.03	0.19
Self-Contained† Vs. NGS	0.21	0.12	0.44**	0.14	0.11	0.20	-0.30	0.19
Random Effects								
Intercept	< 0.01	0.02	0	-	0.08*	0.04	0.05	0.04
Residual	0.65***	0.06	0.84***	0.07	0.83***	0.07	0.87***	0.08
Proportional Reduction in Variance								
Intercept	.99		1.00		.06		.28	
Residual	.00		-.01		.00		.00	

Note. All dichotomous variables are reported such that the category after the "vs." is the comparison group. Self-Contained, Pull-Out, and NGS are shortened names for the three academic programs. Results are reported for models that used Self-Contained as the reference program except for variables marked with †, which used NGS as the reference program. *p<.05, **p <.01, ***p <.001.

Table 4 (cont.)

Hierarchical Linear Models Predicting Psychological Well-Being by Program Type

Variable	Self-Esteem		Loneliness		Math Self-Concept	
ICC	.11		.05		.02	
	B	*SE*	*B*	*SE*	*B*	*SE*
Student-Level						
Intercept	3.58***	0.23	3.94***	0.17	3.49***	0.19
Gender (Female vs. Male)	-0.14	0.10	-0.15	0.09	0.07	0.10
Race (White vs. Non-White)	0.03	0.12	0.02	0.10	-0.05	0.11
Guardians' Highest Education	0.03	0.05	0.02	0.05	0.04	0.05
Classroom-Level						
Grade (6^{th} vs. 5^{th})	-0.28*	0.14	-0.10	0.09	-0.01	0.10
Pull-Out vs. Self-Contained	0.26	0.18	0.25*	0.11	0.44***	0.12
Pull-Out† vs. NGS	-0.02	0.17	-0.06	0.11	0.28*	0.12
Self-Contained† vs. NGS	-0.28	0.17	-0.30**	0.11	-0.16	0.12
Random Effects						
Intercept	0.04	0.03	< 0.01	0.01	0	-
Residual	0.68***	0.06	0.51***	0.04	0.65***	0.06
Proportional Reduction in Variance						
Intercept	.30		.93		1.00	
Residual	.00		.00		.01	

Note. All dichotomous variables are reported such that the category after the "vs." is the comparison group. Self-Contained, Pull-Out, and NGS are shortened names for the three academic programs. Results are reported for models that used Self-Contained as the reference program except for variables marked with †, which used NGS as the reference program. *p<.05, **p <.01, ***p <.001.

Table 5

Hierarchical Linear Models Predicting Psychological Well-Being by Program Type, With SAI as a Covariate

Variable	Adaptive Perfectionism		Maladaptive Perfectionism		School Liking		School Avoidance	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Student-Level								
Intercept	1.74	0.94	2.59*	1.04	2.00	1.07	4.78***	1.09
Gender (Female vs. Male)	0.09	0.13	-0.02	0.14	0.16	0.14	-0.23	0.15
Race (White vs. Non-White)	-0.12	0.15	-0.40*	0.17	0.34*	0.17	-0.41*	0.17
Guardians' Highest Education	0.02	0.07	0.06	0.07	-0.09	0.07	0.01	0.08
SAI Score	0.02*	0.01	< 0.01	0.01	0.01	0.01	-0.01	0.01
Classroom-Level								
Grade (6^{th} vs. 5^{th})	0.11	0.13	0.59***	0.15	-0.36*	0.18	0.14	0.17
Pull-Out vs. Self-Contained	0.08	0.15	-0.41*	0.18	- < 0.01	0.23	0.09	0.21
Pull-Out† vs. NGS	-0.12	0.20	-0.09	0.23	-0.14	0.27	0.07	0.25
Self-Contained† vs. NGS	-0.20	0.23	0.32	0.26	-0.14	0.29	-0.02	0.28
Random Effects								
Intercept	0	-	0.01	0.03	0.06	0.05	0.03	0.05
Residual	0.62***	0.07	0.75***	0.09	0.76***	0.09	0.81***	0.10

Note. All dichotomous variables are reported such that the category after the "vs." is the comparison group. Self-Contained, Pull-Out, and NGS are shortened names for the three academic programs. Results are reported for models that used Self-Contained as the reference program except for variables marked with †, which used NGS as the reference program. *p<.05, **p<.01, ***p<.001

Table 5 (cont.)

Hierarchical Linear Models Predicting Psychological Well-Being by Program Type, With SAI as a Covariate

Variable	Self-Esteem		Loneliness		Math Self-Concept	
	B	*SE*	*B*	*SE*	*B*	*SE*
Student-Level						
Intercept	1.97	1.06	3.96***	0.88	1.67	0.92
Gender (Female vs. Male)	-0.15	0.13	-0.22	0.12	0.11	0.12
Race (White vs. Non-White)	0.22	0.15	0.22	0.14	-0.05	0.14
Guardians' Highest Education	-0.10	0.07	-0.10	0.06	-0.04	0.06
SAI Score	0.01	0.01	< 0.01	0.01	0.02*	0.01
Classroom-Level						
Grade (6^{th} vs. 5^{th})	-0.43**	0.16	-0.18	0.13	-0.08	0.12
Pull-Out vs. Self-Contained	0.28	0.20	0.12	0.17	0.57***	0.15
Pull-Out† vs. NGS	-0.24	0.25	-0.10	0.20	0.09	0.20
Self-Contained† vs. NGS	-0.52	0.28	-0.22	0.23	-0.48*	0.23
Random Effects						
Intercept	0.04	0.04	0.02	0.02	0	-
Residual	0.63***	0.07	0.51***	0.06	0.59***	0.06

Note. All dichotomous variables are reported such that the category after the "vs." is the comparison group. Self-Contained, Pull-Out, and NGS are shortened names for the three academic programs. Results are reported for models that used Self-Contained as the reference program except for variables marked with †, which used NGS as the reference program. *p<.05, **p<.01, ***p <.001

Supplementary Materials

Supplementary Table 1

Demographic Characteristics of Students Included and Excluded from HLMs Due to Missing Data

Outcome Variable (# of Students Excluded from model due to missing data)	**% Female**		**% White**		**Mean Guardians' Highest Education (SD)**		**% in 5th Grade**	
	Inc.	*Excl.*	*Inc.*	*Excl.*	*Inc.*	*Excl.*	*Inc.*	*Excl.*
Loneliness (19)	56%	46%	73%	60%	2.43 (0.95)	2.44 (0.82)	46%	47%
School Liking (18)	55%	67%	73%	67%	2.44 (0.96)	2.20 (0.56)	46%	50%
School Avoidance (19)	55%	62%	73%	60%	2.43 (0.95)	2.38 (0.89)	46%	53%
Self-Esteem (22)	55%	56%	73%	69%	2.44 (0.95)	2.32 (0.89)	46%	50%
Adaptive Perfectionism (18)	55%	58%	73%	56%	2.44 (0.96)	2.20 (0.56)	46%	44%
Maladaptive Perfectionism (18)	55%	50%	73%	56%	2.44 (0.96)	2.20 (0.56)	46%	44%
Math Self-Concept (16)	55%	60%	73%	57%	2.44 (0.95)	2.23 (0.60)	46%	50%

Note. "Inc." columns denote the demographic characteristics of the students who were included in the HLM predicting the psychological well-being variable in the given row, while "Excl." columns denote the demographic characteristics of the students who were not included in the HLM due to missing data. These HLMs were conducted using listwise deletion, so students were excluded if they were missing data for the outcome variable (e.g., loneliness) or any of the covariates (gender, race, guardians' education, grade, program).

Supplementary Table 2

Hierarchical Linear Models of Permission Status Predicting Psychological Well-Being Variables

Variable	Adaptive Perfectionism		Maladaptive Perfectionism		School Liking		School Avoidance	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects								
Intercept	3.86***	0.08	2.79***	0.11	3.25***	0.12	3.22***	0.12
Permission Status[a] ("Yes" vs. "No")	-0.07	0.10	-0.01	0.12	0.17	0.12	-0.23	0.13
Random Effects								
Intercept	0.01	0.02	0.08*	0.05	0.13*	0.06	0.09*	0.05
Residual	0.66***	0.06	0.86***	0.08	0.83***	0.07	0.88***	0.08

Variable	Self-Esteem		Loneliness		Math Self-Concept	
	B	*SE*	*B*	*SE*	*B*	*SE*
Fixed Effects						
Intercept	3.67***	0.11	4.12***	0.08	3.83***	0.08
Permission Status[a] ("Yes" vs. "No")	-0.03	0.11	-0.12	0.09	-0.05	0.10
Random Effects						
Intercept	0.08*	0.04	0.02	0.02	0.01	0.02
Residual	0.69***	0.06	0.51***	0.05	0.68***	0.06

Note. *p<.05, **p<.01, ***p <.001.

[a] Permission status refers to whether or not parents provided the school with permission to release their children's SAI scores and gifted identification data. Permission status is reported such that "No" is the reference group.

Supplementary Table 3

Hierarchical Logistic Regression Predicting Permission Status

Variable	Likelihood of Receiving Permission to Have Data Released	
	B	*SE*
Student-Level		
Intercept	0.57	0.69
Gender (Female vs. Male)	0.75**	0.29
Race (White vs. Non-White)	0.61*	0.31
Guardians' Highest Education	-0.13	0.14
Classroom-Level		
Grade (6^{th} vs. 5^{th})	0.30	0.49
Pull-Out vs. Self-Contained	-1.00	0.65
Pull-Out† vs. NGS	0.36	0.58
Self-Contained† vs. NGS	1.36*	0.60

Note. All dichotomous variables are reported such that the category after the "vs." is the comparison group. Self-Contained, Pull-Out, and NGS are shortened names for the three academic programs. Results are reported for models that used Self-Contained as the reference program except for variables marked with †, which used NGS as the reference program..
*p<.05, **p <.01, ***p <.001.

Supplementary Table 4

Hierarchical Linear Models Predicting Psychological Well-Being by Program Type with Multiple Imputations

Variable	Adaptive Perfectionism		Maladaptive Perfectionism		School Liking		School Avoidance	
	B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Student-Level								
Intercept	3.79***	0.19	3.29***	0.22	3.28***	0.26	3.27***	0.25
Gender (Female vs. Male)	0.03	0.10	-0.01	0.11	0.20	0.12	-0.21	0.12
Race (White vs. Non-White)	-0.16	0.11	-0.29*	0.13	0.22	0.13	-0.31*	0.13
Guardians' Highest Education	0.08	0.05	-0.07	0.06	0.04	0.06	-0.07	0.06
Classroom-Level								
Grade (6^{th} vs. 5^{th})	0.12	0.10	0.40***	0.11	-0.37*	0.17	0.20	0.16
Pull-Out vs. Self-Contained	-0.13	0.12	-0.46***	0.14	-0.21	0.22	0.29	0.20
Pull-Out† vs. NGS	0.08	0.12	-0.02	0.14	-0.07	0.20	-0.02	0.19
Self-Contained† vs. NGS	0.21	0.12	0.44**	0.14	0.14	0.21	-0.32	0.19

Note. All dichotomous variables are reported such that the category after the "vs." is the comparison group. Self-Contained, Pull-Out, and NGS are shortened names for the three academic programs. Results are reported for models that used Self-Contained as the reference program except for variables marked with †, which used NGS as the reference program. This model was conducted as a robustness check of the analysis using listwise deletion, so only fixed effects are reported. *p<.05, **p <.01, ***p <.001.

Supplementary Table 4 (cont.)

Hierarchical Linear Models Predicting Psychological Well-Being by Program Type with Multiple Imputations

Variable	Self-Esteem		Loneliness		Math Self-Concept	
	B	*SE*	*B*	*SE*	*B*	*SE*
Student-Level						
Intercept	3.56***	0.22	3.90***	0.17	3.46***	0.20
Gender (Female vs. Male)	-0.15	0.11	-0.14	0.09	0.06	0.10
Race (White vs. Non-White)	0.01	0.12	0.01	0.10	-0.06	0.11
Guardians' Highest Education	0.04	0.05	0.03	0.05	0.06	0.05
Classroom-Level						
Grade (6^{th} vs. 5^{th})	-0.31*	0.14	-0.10	0.09	-0.02	0.10
Pull-Out vs. Self-Contained	0.28	0.17	0.24*	0.11	0.44***	0.12
Pull-Out† vs. NGS	-0.01	0.16	-0.08	0.11	0.28*	0.12
Self-Contained† vs. NGS	-0.29	0.17	-0.32**	0.11	-0.16	0.12

Note. All dichotomous variables are reported such that the category after the "vs." is the comparison group. Self-Contained, Pull-Out, and NGS are shortened names for the three academic programs. Results are reported for models that used Self-Contained as the reference program except for variables marked with †, which used NGS as the reference program. This model was conducted as a robustness check of the analysis using listwise deletion, so only fixed effects are reported. *p<.05, **p <.01, ***p <.001.