8-2021

Youth Profiles of Program Quality Perceptions and Their Associations with Social Competence

Scot R. Seitz
Georgia State University

YOUTH PROFILES OF PROGRAM QUALITY PERCEPTIONS AND THEIR ASSOCIATIONS WITH SOCIAL COMPETENCE

by

SCOT R. SEITZ

Under the Direction of Gabriel Kuperminc, PhD

ABSTRACT

Many youth development programs (YDPs) aim to help young people improve their social competence, defined as the ability to navigate social situations effectively. Researchers have documented that the quality of YDPs is associated with youth outcomes, including social competence. There are multiple aspects of program quality (e.g., fun, sense of belonging). Most previous research has utilized a variable-centered approach to explore how individual components of program quality are associated with young people's social competence. However, program participants do not experience individual aspects of program quality in isolation. The current study used person-centered analyses to account for young people's experiences of multiple components of program quality. Using national data from 166,640 participants of Boys

& Girls Clubs of America (BGCA), this study identified profiles of program participants based on their perceptions of seven indicators of program quality (i.e., fun, sense of belonging, adult caring relationships, physical safety, emotional safety, staff expectations, staff recognition). Youth (9-12 years old) and teens (13-18 years old) were classified into five profiles. Four of the profiles included consistent ratings for all indicators of program quality: Above Average, Below Average, Very Below Average, and Negative. Approximately 70% of youth and teens were classified into the Above Average profile. The fifth profile was small (fewer than 2% of participants) and included above average ratings for all indicators except Belonging and Emotional Safety. Young people in the Above Average and Low Belonging and Emotional Safety profiles reported higher levels of each social competence outcome variable (i.e., Teamwork, Leadership, Conflict Resolution, Concern for Others) than young people in the three other profiles. When comparing levels of social competence in 2018 controlling for baseline levels in 2017, young people in the Above Average profile and the Low Belonging and Emotional Safety profile reported higher social competence than young people in the other profiles for 21 of the 48 comparisons after Bonferroni correction. Given that program quality is associated with social competence, it is important for YDPs to prioritize improving quality, especially for young people who have low perceptions of program quality and who experience difficulties in social interaction.

INDEX WORDS: Youth development program, Program quality, Social competence

YOUTH PROFILES OF PROGRAM QUALITY PERCEPTIONS AND THEIR ASSOCIATIONS WITH SOCIAL COMPETENCE

by

SCOT R. SEITZ

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2021

YOUTH PROFILES OF PROGRAM QUALITY PERCEPTIONS AND THEIR ASSOCIATIONS WITH SOCIAL COMPETENCE

by

SCOT R. SEITZ

Committee Chair: Gabriel Kuperminc

Committee: Lisa Armistead

Christopher Henrich

Laura McKee

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

August 2021

DEDICATION

This dissertation is dedicated to my family and friends who have supported me unconditionally. I am forever grateful for the love and support of my husband, Pierce, as well as my parents, grandparents, brothers, and friends who have become my chosen family.

ACKNOWLEDGEMENTS

There are many people to acknowledge who have contributed to the success of this dissertation project and my career. Gabe Kuperminc has been my primary mentor throughout graduate school, and he has played a critical role in my professional and personal development. I would also like to thank the entire EcoLab team, both past and present, who continue to inspire me. This project would not have been possible without the support of staff at Boys & Girls Clubs of America (BGCA) who work tirelessly to promote the healthy development of young people. I am particularly grateful for Omar Guessous's guidance and support. I would also like to acknowledge the Boys & Girls Clubs staff members and young people who participated in BGCA's National Youth Outcomes Initiative. In addition, I have had countless mentors throughout high school, college, and graduate school who have gone above and beyond for me, including my clinical psychology advisor Erin Tone and my mentor and friend, Juan Leon. Finally, this project was only possible because of the community that loves and supports me, including my husband Pierce, other family members, and friends.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Youth development programs (YDPs) seek to promote youths' well-being by creating supportive environments in which young people can develop adaptive skills, attitudes, and beliefs (Roth & Brooks-Gunn, 2003, 2016). As youth programs have become more widespread, researchers and practitioners increasingly recognize the importance of program *quality* for the promotion of youth development (Yohalem & Wilson-Ahlstrom, 2010). Previous research suggests that several aspects of YDP quality (e.g., supportive youth-adult relationships, youth engagement, physical and emotional safety) may promote positive youth outcomes (e.g., Chapman, Deane, Harré, Courtney, & Moore, 2017; Durlak, Weissberg, et al., 2010; Pierce, Bolt, & Vandell, 2010). Most of this previous research has utilized variable-centered approaches that examine how individual components or overall composite measures of program quality promote youth well-being (Denault & Poulin, 2016; Laursen & Hoff, 2006). Variable-centered approaches can be helpful for better understanding the relations among program quality indicators and specific youth outcomes. To complement these approaches, it is also important to utilize person-centered analyses that examine how youth experience multiple elements of program quality. For example, some youth may feel engaged at a YDP despite a lack of close relationships with staff members, while others may find the program boring even though they have close relationships with the program staff. Within the same program, youth may have different perceptions of program quality due to a range of personal characteristics and motivations for joining the YDP (e.g., youth who join programs just to play sports may not actively work to build close relationships with staff members; Fredricks et al., 2002; Hirsch et al., 2011).

Only one study has utilized a quantitative person-centered approach to explore how young people experience multiple aspects of program quality (Denault & Poulin, 2016). However, this study only included two experiences that are considered indicators of program quality. To address the lack of research into how young people experience multiple aspects of program quality (Denault & Poulin, 2016), this study identified profiles of program participants' experiences in a national sample of YDPs affiliated with Boys & Girls Clubs of America (BGCA), using seven indicators of program quality (i.e., fun, sense of belonging, physical safety, emotional safety, adult connections, staff recognition, staff expectations).

Understanding heterogeneity in young people's distinct experiences of multiple program quality components may provide greater insights into the ways in which YDPs can help to promote positive youth outcomes (Mahoney et al., 2010). One important youth outcome that YDPs often strive to promote is social competence, which can be defined as navigating social interactions effectively (e.g., working with a team, resolving conflict with others) (Catalano et al., 2004; Rose-Krasnor, 1997). Social competence can help young people cultivate success and well-being throughout childhood and adulthood (Milligan et al., 2017; Rose-Krasnor, 1997). However, youth living in low-income communities demonstrate less social competence than their peers and are therefore in a disadvantaged position for healthy development (Gershoff et al., 2007; Lengua, 2012; Yoshikawa et al., 2012). Although researchers have demonstrated that YDPs can promote social competence for all young people, including those living in low-income communities, it is important to better understand how different aspects of program quality combine to promote increased social competence. This understanding can help programs prioritize key domains of program quality that can facilitate social competence for all youth (Durlak et al., 2010; Yohalem & Wilson-Ahlstrom, 2010). As one example, youth may need to

feel a sense of belonging, have supportive relationships, and be engaged at a YDP in order to experience gains in social competence. Previous research has not accounted for how configurations such as these are related to youth outcomes. This study explored whether different profiles of young people's experiences of program quality components were associated with social competence at one time point and over the course of one year.

1.1 Social Competence

Although social competence has been conceptualized in various ways (Milligan et al., 2017), most definitions of social competence include the concept of "effectiveness in social interaction" (Rose-Krasnor, 1997, p. 111). Rose-Krasnor defines social competence as successfully meeting both short- and long-term developmental needs in social contexts. Short-term developmental needs are those that are important for specific developmental periods (Waters & Sroufe, 1983). For example, establishing a healthy attachment relationship is an important task during infancy. Long-term developmental needs are those that promote health and well-being throughout adolescence and adulthood (e.g., maintaining relationships with parents while also having a sense of autonomy). Social competence requires specific skills, such as proficient communication and affect regulation, to navigate social situations to meet these developmental needs.

According to Rose-Krasnor's model (1997), specific domains of social competence are associated with particular goals (e.g., developing friendships, resolving conflict). Social competence in one domain (e.g., successfully resolving conflicts with friends) does not necessarily indicate social competence in a different domain (e.g., working well with others on a team). Four domains of social competence commonly examined are conflict resolution, teamwork, leadership, and concern for others. Conflict resolution skills require the ability to

consider other people's perspectives, balance personal needs with the needs of others, and determine and execute a plan that meets many of the needs of those involved in the interpersonal conflict (Chen, 2003). Teamwork is a related concept that can be defined as working cooperatively with others to achieve a common goal (Guest, 2008). Leadership also involves working with others, although it focuses on influencing other people to work towards a common purpose (Redmond & Dolan, 2016). Finally, concern for others involves sympathizing with the feelings of another person (Eisenberg, 2014). All four of these domains of social competence are related in that they involve effectively navigating social situations.

1.2 Social Competence and Youth Development

Social competence can promote healthy development by enabling young people to meet their various developmental needs, including mental health, academic achievement, and physical health (Milligan et al., 2017). Socially competent young people experience better mental health over time, likely because they are able to recruit more robust social support networks and they can effectively express themselves during times of need (Bornstein et al., 2010; Burt et al., 2008; Huber et al., 2018). In addition, young people who demonstrate social competence are more likely to achieve academic success over time because they are better-positioned to seek academic support from peers and teachers (Caprara et al., 2000; Denham & Brown, 2010). With regard to physical health, there is evidence that people who are not able to build social support networks are at an increased risk for chronic stress which is associated with a range of physical health problems (Repetti et al., 2002).

Although social competence can help promote well-being for all young people, those living in disadvantaged communities may benefit more from efforts to increase social competence (Yoshikawa et al., 2012). Elevated levels of stress associated with living in low-

income communities may negatively impact parent-child relationship quality which can, in turn, prevent children from learning how to regulate their emotions and navigate interpersonal relationships (Lengua, 2012; Martí et al., 2016). In addition, chronic stress associated with poverty can lead to dysregulated stress responses and a decreased ability to regulate behavior, two important components of social competence (Lengua, 2012). Given the role of social competence in facilitating healthy development, it is important to identify strategies for promoting social competence, especially for young people living in disadvantaged communities.

1.3 Youth Development Programs and Social Competence

Since YDP activities often take place in groups with peers and adults, YDPs are well-suited to promote social competence (Eccles, Barber, Stone, & Hunt, 2003; Fredricks, Naftzger, Smith, & Riley, 2017). Indeed, one of the goals of many YDPs is to help participants develop social competence (Catalano et al., 2004). Based on a survey of staff members from 71 YDPs, 80% listed social skills training and 78% listed leadership training as important aspects of their program (Roth & Brooks-Gunn, 2003). On a different survey, program participants, researchers, and practitioners identified opportunities for developing leadership skills and life skills as key aspects of YDPs (Urban, 2008).

Evidence suggests that YDPs are a promising approach for developing social competence (Catalano et al., 2004; Deutsch et al., 2017; Pittman, 2017). Durlak et al. (2010) conducted a meta-analysis of 75 studies examining the effectiveness of after-school programs that aim to facilitate social and academic outcomes. They found that program participants experienced greater gains in positive social behaviors (e.g., leadership, navigating interpersonal conflict, cooperation) compared to young people who were not program participants. Although this meta-analysis did not report separate findings for young people living in low-income communities,

several studies have found similar associations for such young people (e.g., Wright, John, Alaggia, & Sheel, 2006). These quantitative data are supported by qualitative evidence in which program participants and parents reported benefits from YDP participation in a range of domains, including teamwork, leadership, and communication (Dworkin, Larson, & Hansen, 2003; Wright et al., 2006).

The social development model provides insight into the specific mechanisms through which YDPs may promote social competence (Brackett et al., 2015). According to this model, individuals develop social bonds to families, schools, and youth organizations; these bonds influence individuals to behave in accordance with the social norms of the group (Catalano & Hawkins, 1996; Duerden & Witt, 2010; Hawkins & Weis, 1985). Social norms are standards for behaviors and attitudes within a particular group of people (McKeown & Taylor, 2018; Sherif, 1936), and they have been shown to influence young people's actions, including prosocial behavior (Berkowitz, 1972; McKeown & Taylor, 2018). Social norms offer a frame for people to judge whether a behavior is appropriate based on how consistent it is with other people's behavior and whether it is generally approved by other group members (Schultz et al., 2007). Thus, YDP participants may gain social competence through exposure to prosocial norms. Indeed, one of the defining characteristics of YDPs is a social environment that involves supportive relationships and expectations for positive behavior (Catalano et al., 2004; Roth & Brooks-Gunn, 2003). As evidence for this, Hansen, Larson, and Dworkin (2003) found that young people experienced more opportunities to learn about prosocial norms (e.g., learning about helping others) in youth programs compared to academic classes or unstructured time with their friends.

The social development model also draws on social learning theory to describe how behaviors, such as successfully navigating social situations, are learned and maintained over time (Hawkins & Weis, 1985). According to social learning theory, people can learn behaviors by observing others, and they are more likely to perform the learned behaviors if they observe that the behaviors result in positive outcomes (Bandura, 1977). At YDPs, participants may learn social competence by observing adults or peers engaging in prosocial behaviors (e.g., working cooperatively with a team, leading an activity, resolving interpersonal conflict) that yield positive rewards such as close relationships and collaborative goal attainment.

Although young people may learn positive social behaviors through observation, their social competence can be further developed and refined over time when they have opportunities to practice social skills. Practicing social skills has been identified as an important component of effective social skills training (Gresham, 2016). YDPs may facilitate the development of informal relationships in which program participants can practice socials skills. Supporting this idea, Bohnert, Aikins, and Edidin (2007) found that youth program participation is associated with increased friendship quality and decreased loneliness. As youth develop more relationships at YDPs, they likely have more opportunities to practice navigating social situations. Beyond having opportunities to practice skills, social learning theory suggests that learning is optimized when feedback is given (Bandura, 1977). Thus, YDPs may also facilitate gains in social competence when peers and adults at the programs provide feedback about young people's interpersonal behavior. Based on focus groups with 55 high school students discussing how they benefited from program activities, Dworkin, Larson, and Hansen (2003) found that young people learned how to work together, be successful leaders, and communicate effectively by having

opportunities to practice these skills. The program participants noted that receiving constructive feedback from others was important for learning skills related to social competence.

Finally, young people may enhance their social competence at YDPs by improving their ability to take other people's perspectives. Qualitative evidence suggests that YDP participants develop relationships with peers who are outside of their typical social networks, and that these relationships can lead to increased understanding of and empathy for peers (Dworkin et al., 2003; Wilson et al., 2017). Gains in empathy and understanding, in turn, contribute to growing perspective-taking skills which are important for many domains of social competence, including conflict resolution and teamwork (Guest, 2008; Sandy, 2014).

1.4 Experiences of Program Quality and Social Competence

Program quality may influence the trajectory of youth outcomes, including social competence (Fredricks et al., 2017; Shernoff, 2010; Yohalem & Wilson-Ahlstrom, 2010). When assessing program quality, researchers often draw on a report from the National Research Council (NRC) and the Institute of Medicine (IOM) (2002; Yohalem et al., 2009). This report identified eight components of youth programs that likely contribute to positive developmental outcomes based on theories of positive youth development and empirical research on how young people benefit from various settings (e.g., schools, families, community programs). The eight components include 1) physical and psychological safety; 2) appropriate structure; 3) supportive relationships; 4) opportunities to belong; 5) positive social norms; 6) support for efficacy and mattering (which aligns with the concept of engagement/fun; Yohalem & Wilson-Ahlstrom, 2010); 7) opportunities for skill building and; 8) integration of family, school, and community efforts. Overall, there is general agreement about these main components of program quality and

researchers often use them when developing and evaluating measures of youth program quality (Silliman & Schumm, 2013; Yohalem et al., 2009; Yohalem & Wilson-Ahlstrom, 2010).

The social development model (described above; Catalano & Hawkins, 1996; Duerden & Witt, 2010; Hawkins & Weis, 1985) offers a theoretical framework for understanding how program quality elements, and in particular, young people's experiences of those elements, may promote social competence. Specifically, some of the program quality components can be categorized as 1) contributing to the development of a social bond between YDPs and participants (hereafter referred to as "social bond-related") or 2) facilitating social norms that influence socially competent behavior (hereafter referred to as "social norms-related").

Social bond-related program quality elements include *opportunities to belong*, *supportive relationships*, and *engagement*. Sense of belonging includes feeling like a valued member of a group who can contribute in a meaningful way (NRC & IOM, 2002). Supportive relationships involve youth feeling close and connected to other people at the youth program. If participants have a sense of belonging and feel connected to their peers and staff members, they will likely develop a strong bond to the program that could facilitate social competence. In addition, young people's social bonds to YDPs likely strengthen when they find the program engaging, a multi-dimensional construct that includes effort, interest, and enjoyment/fun (Bartko, 2005; Catalano & Hawkins, 1996; Hawkins & Weis, 1985; Roth et al., 2010). Research evidence supports the idea that if young people have a sense of belonging at a YDP (McDonough et al., 2013), feel connected to their peers and staff members (Kataoka & Vandell, 2013), and are engaged in the program (Shernoff, 2010), they are more likely to experience gains in social competence. These gains may result from participants identifying with the YDP and, in turn, being influenced by the social norms of the program (Hawkins & Weis, 1985; Terry & Hogg, 1996).

Social norms-related program quality elements include *positive social norms* (e.g., expectations about and rewards for appropriate behavior) and *physical and psychological safety* (NRC & IOM, 2002). YDPs that have clear expectations about and rewards for appropriate behavior and are physically and psychologically safe are likely to cultivate social norms characterized by respect and prosocial behaviors. There is a small body of research indicating that these positive social norms at YDPs can facilitate the development of young people's social competence over time (e.g., Moore & Hamilton, 2010).

1.5 The Importance of Participants' Perceptions of Program Quality

Although youth program quality is often conceptualized as a setting-level construct and is measured through observation (Yohalem et al., 2009), it is also important to consider participants' experiences (Bean & Forneris, 2016; Silliman & Schumm, 2013). Young people who attend the same program may have different experiences with regard to program quality (Baldwin et al., 2015; Denault & Poulin, 2016; Vandell et al., 2005). For example, while visiting youth programs, Baldwin, Stromwall, and Wilder (2015) observed that certain participants did not seem engaged in program activities even when others seemed highly engaged. Measuring individual participants' perceptions of program quality can provide insight into the inter-individual variability of program quality experiences that may exist within a program.

Young people's assessment of their own experiences may also provide information that is not gathered by program observations. For instance, although participants may look engaged during an activity, they may not actually be engaged (Bartko, 2005). As Silliman and Schumm noted, "From a practical perspective, youth are the most authentic evaluators of their experience" (2013, p. 651).

1.6 Age and Other Covariates When Examining Perceptions of Program Quality

When examining perceptions of program quality, it is important to consider participants' age (Silliman & Schumm, 2013). Researchers have documented positive associations among participant age and program quality measures (e.g., Kuperminc et al., 2019; Ramey et al., 2018). This pattern may be due, in part, to teenagers often having more options and autonomy than younger youth for how they spend their time after school (Deschenes et al., 2010). Older youth who have negative experiences at YDPs may stop attending the program, resulting in a larger proportion of older participants who have positive perceptions of program quality.

In addition, because participants of different ages have different developmental needs (e.g., teenagers often have greater needs for autonomy and peer relationships; Meschke, Peter, & Bartholomae, 2012), youth programs often design different programming for teenagers and pre-teenagers (e.g., Arbreton, Bradshaw, Metz, Sheldon, & Pepper, 2008; Boys & Girls Clubs of America, 2018f). Indeed, many YDPs, including BGCA, often provide separate physical space for teenagers (Boys & Girls Clubs of America, 2020). Given that teenagers and pre-teenagers are frequently separated at YDPs and engage in different activities, they may experience various components of program quality in different ways. To account for these potential differences, in this study all analyses were conducted separately for two age groups: pre-teenagers (12 years old and younger) and teenagers (13 years old and older). These two age groups are consistent with previous research on youth programs that recognizes the unique needs and experiences of teenagers compared to pre-teenagers (e.g., Arbreton et al., 2008).

To understand the associations between perceptions of program quality and the development of social competence, it is also important to account for demographic characteristics (e.g., gender, socioeconomic status) and participation (e.g., frequency of

attendance, duration of program participation) (e.g., Flett et al., 2012; Kuperminc et al., 2019; Ramey et al., 2018; Shernoff, 2010; Walker & Arbreton, 2004). These potential covariates may affect perceptions both of program quality and the development of social competence (e.g., Elias & Gordon, 2009; Fredricks et al., 2017; Rosch, Ogolsky, & Stephens, 2017).

1.7 Limitations of Previous Research on Program Quality and Developmental Outcomes

Most research on the association between experiences of program quality and young people's development has examined aspects of program quality separately. For example, Smith et al. (2017) examined whether eleven different aspects of program quality (e.g., appropriate structure, active engagement, supportive relationships with peers/adults) were associated with four different youth outcomes (i.e., collective efficacy, prosocial behavior, afterschool connectedness, respect for adults) using separate multilevel regression models for each combination of program quality indicator and youth outcome. In addition to examining the link between separate measures of program quality and youth outcomes, some researchers have utilized an overall composite measure of program quality. As one example, Smith et al. (2017) also examined whether the four youth outcomes (i.e., collective efficacy, prosocial behavior, afterschool connectedness, respect for adults) were associated with a composite measure of quality (i.e., factor scores from a factor analysis of all the separate indicators of program quality).

Although examining aspects of program quality separately can provide insight into the link between developmental outcomes and individual components of program quality, young people do not experience each aspect of program quality in isolation (Denault & Poulin, 2016). For instance, if a young person experiences a program as engaging this does not necessarily mean that she/he also experiences positive and supportive relationships within the program. These distinct program experiences may result from individual differences in how young people

experience the same program (Baldwin et al., 2015; Vandell et al., 2005). The likelihood that participants will experience the same program in different ways points to the value of person-centered research. Compared to variable-centered approaches, person-centered research shifts the focus from the associations among variables to the classification of individuals based on their unique experience of multiple variables (Laursen & Hoff, 2006). This type of analysis can provide insight into young people's varied experiences at YDPs.

However, there is a lack of person-centered research that examines young people's experiences of program quality. Indeed, in the only such study that utilized a person-centered approach, Denault and Poulin (2016) administered surveys to 413 youth in 33 different youth programs. The surveys assessed six different aspects of their program participation: feeling socially connected, receiving support from activity leaders, learning about their identities, developing social skills, learning how to take initiative, and developing emotion regulation abilities. Using latent profile analysis, Denault and Poulin identified three different categories of participants: those who provided low, average, or high ratings for all six aspects of their program participation. That is, young people tended to rate the various aspects of their program experiences similarly. It is noteworthy that this analysis included factors that are usually considered youth outcomes (i.e., learning about social skills, developing emotion regulation abilities) and only included two factors that are considered program quality elements (i.e., feeling socially connected, receiving support from activity leaders). The current study examines seven program quality elements and avoids confounding program quality with developmental outcomes. In addition, the current study uses a novel approach for organizing program quality components based on their contributions to promoting positive social bonds and social norms.

By focusing on associations between individual aspects of program quality and developmental outcomes, researchers have not explored whether young people's experiences of various program quality components facilitate healthy development. Are there optimal profiles of participants' experiences of YDP quality that are associated with the highest gains in developmental outcomes (Mahoney et al., 2010)? If so, this information could help practitioners prioritize elements of program quality when working toward quality improvements. For instance, if only participants who report both a sense of belonging and high behavioral expectations experience gains in social competence, then practitioners could focus on improving these aspects of program quality. The current study aimed to contribute to the research literature by analyzing how different experiences of program quality indicators were associated with social competence.

Finally, previous research has not examined whether profiles of program quality perceptions differ across setting-level characteristics. There is only limited evidence examining the associations among individual program quality components and setting-level characteristics. For example, previous research has documented a negative association between the number of youth served by a program and youth perceptions of quality (e.g., Kuperminc et al., 2019; Walker & Arbreton, 2004). At programs with a large number of youth, there are likely fewer opportunities for youth to interact with staff members and it can be more difficult for staff members to maintain a structured environment. In addition, preliminary evidence suggests that there may be associations among program quality perceptions and community location (e.g., rural/urban/suburban) and type of facility (e.g., school-based, community-based) (e.g., Baldwin et al., 2015; Sullivan & Larson, 2010). However, research has not examined whether unique profiles of program quality differ across these setting-level characteristics.

1.8 Current Study

The first aim of this study was to identify profiles of BGCA participants based on their experiences of seven components of program quality (fun, sense of belonging, physical safety, emotional safety, positive adult connections, staff recognition, staff expectations). Since there is a limited body of research on this topic, this aim was largely exploratory. However, the study by Denault and Poulin (2016) suggests the likelihood that profiles of young people based on how positively they experienced different quality domains (low, average, and high ratings) would emerge. Thus, in the current study it was hypothesized that BGCA participants may rate all seven components of program quality similarly, resulting in three groups that could be classified based on high, average, and low ratings. As another possibility, drawing on the social development model, program quality elements can be differentiated as social bond-related (i.e., positive adult connections, sense of belonging, fun/engagement) and social norms-related (i.e., physical and psychological safety, staff expectations, staff recognition). It was also hypothesized that young people may have similar ratings within these two categories, but different ratings across the categories. For example, one profile could include high ratings for social bond-related elements but low ratings for the social norms-related elements. These young people may feel a strong connection to their BGCA site but may not perceive positive social norms at the site.

This study also examined whether the different profiles of participants' experiences were associated with youth social competence at one time point (study aim two) and with social competence over the course of one year (study aim three). The social development model and social learning theory suggest possible profiles that may best promote social competence. For instance, young people who feel a sense of connection to a YDP may be more likely to adopt its social norms (Catalano & Hawkins, 1996; Hawkins & Weis, 1985). Thus, participants who

experience high levels of social bond-related aspects of program quality may be well-positioned to develop social competence because they feel connected to the program. Without a strong connection to the program, young people may not experience gains in social competence because they are not as influenced by the social norms of the program (even if the social norms promote social competence) (Terry & Hogg, 1996). Therefore, it was hypothesized that young people would only report high levels of social competence if they were classified into profiles with high ratings for the social bond-related indicators of program quality. Since positive social norms likely facilitate social competence, it was also hypothesized that profiles with high ratings for the social norms-related components would be associated with even higher levels of social competence.

The last aim of this study was to examine whether the proportion of youth profiles was different across different site characteristics (e.g., number of youth served each day, rural/urban/suburban community location, community/school/public housing/Native American facility type). Since activities with fewer participants are often associated with better perceptions of program quality (e.g., Kuperminc et al., 2019; Walker & Arbreton, 2004), it was expected that there would be a higher proportion of profiles with positive program quality ratings at sites that served fewer youth. The analyses involving community location (e.g., rural, urban, suburban) and facility type (e.g., school, community, public housing, Native American) were exploratory given the limited evidence for the associations among these site-level characteristics and young people's perceptions of program quality.

2 METHOD

2.1 Program Description

BGCA is a national federation of over 1,100 organizations that serves nearly four million young people each year in approximately 4,300 Clubs (Boys & Girls Clubs of America, 2018b). The mission of BGCA is "to enable all young people, especially those who need [BGCA] most, to reach their full potential as productive, caring, responsible citizens" (Boys & Girls Clubs of America, 2018e). Their focus on supporting young people who "need [BGCA] most" indicates their commitment to serving young people from disadvantaged communities.

BGCA works to accomplish their mission by developing and implementing specific programs, including programs that focus on social competence (Boys & Girls Clubs of America, 2017, 2018f). For instance, Triple Play is a health/wellness program that helps young people build healthy relationships in addition to teaching them about nutrition and physical activity. Keystone Clubs is a teen program that provides leadership opportunities related to academics, career preparation, and community service. Other programs focus on a range of topics, including sports/recreation, education, the arts, health and wellness, and career development.

According to BGCA's theory of change, Clubs promote participants' positive development by creating a "high-quality Club Experience" (Boys & Girls Clubs of America, 2017, p. 8), which involves creating a psychologically and physically safe environment, promoting fun and a sense of belonging, cultivating supportive relationships with peers and staff members, establishing high expectations for young people, and recognizing young people for their strengths, talents, and positive behaviors. A high-quality Club experience is theorized to drive healthy development, especially in the domains of academics, character/citizenship, and physical/psychological health.

2.2 Procedure

The current study used data collected in 2017 and 2018 for BGCA's National Youth Outcome Initiative (NYOI), a program that regularly collects and analyzes data from BGCA members to improve BGCA's services and document their impact (Boys & Girls Clubs of America, 2018c). The NYOI survey is administered in the spring of each year to Club members who are at least 9 years old. All BGCA Clubs were invited to participate in the 2017 and 2018 NYOI survey administration (approximately 68% and 74% of Clubs participated in 2017 and 2018, respectively).

The 2017 and 2018 NYOI projects were approved by the Abt Associates Inc. Institutional Review Board. Parental consent (either active or passive) was required for young people to participate (Boys & Girls Clubs of America, 2018a). BGCA did not provide compensation to Club members for completing the survey, although some Clubs provided small incentives to participants who completed the survey (e.g., stickers, candy, access to a pizza party). Each Club was encouraged to survey at least the average number of participants (9 years old and older) who attended the Club on a daily basis. The surveys were administered by staff members at each Club to groups of 15-20 Club members using a standardized script. The script included instructions for completing the survey and information about confidentiality (e.g., the script noted that the survey was voluntary and that young people's individual answers would not be shared with their parents, friends, or anyone at their Club).

Participants completed the surveys online or on paper while being supervised by at least one staff member. A staff member (or an automated computer program) read each question aloud for younger Club members (ages 9-12) and some older Club members (ages 13-18). The surveys included questions about demographics, perceptions of the Club experience, youth achievement,

youth attitudes and expectations, and youth behaviors. An additional module that included the Concern for Others measure was administered at a subset of clubs. The total survey completion time was approximately 30-45 minutes.

2.3 Participants and Clubs

Participants (N = 166,640) were members of BGCA Clubs and ranged in age from 9 to 18 years old (see Table 2.1). Girls comprised 47% of the sample, and most participants self-identified as Black/African American (29.2%), Caucasian/White (22.8%), Latino/Hispanic (19.9%), or Multiracial (15.3%). Most participants received free or reduced-price lunch (75.4%).

A total of 2,969 Clubs were included in the sample (see Table 2.2). Most of the Clubs were located in community settings (48.5%), and 36.5% were located in schools. The Clubs were located throughout the United States, with 46.4% in urban regions, 29.3% in suburban areas, and 24.3% in rural communities.

Table 2.1 Demographic Characteristics of Study Participants

	Mean (SD)	Frequency (%)	% Missing Data
Female		78,941 (47.4)	0.1
Race/Ethnicity			1.0
Black/African American		48,250 (29.2)	
Caucasian/White		37,702 (22.8)	
Latino/Hispanic		32,881 (19.9)	
Multiracial		25,277 (15.3)	
Other		10,963 (6.6)	
American Indian/Alaskan Native		4,844 (2.9)	
Asian		3,990 (2.4)	
Native Hawaiian/Pacific Islander		1,105 (0.7)	
Free/Reduced-Price Lunch		81,323 (75.4)	35.2
Age	11.46 (2.22)		0.3
9-12 Years Old		120,396 (72.4)	
13-18 Years Old		45,793 (27.6)	
Years Membership	1.82 (1.91)		21.2
Attendance			37.6
Highly Engaged Attender (2+x/week)		104,007 (71.0)	
Engaged Attender (1x/week)		17,131 (16.5)	
Occasional Attender (<1x/week)		13,027 (12.5)	

Note. Total N = 166,640.

Table 2.2 Club Characteristics

	Mean (SD)	Frequency (%)	% Missing Data
Community Location			0.07
Urban		1,377 (46.4)	
Suburban		869 (29.3)	
Rural		721 (24.3)	
Facility Type			0.00
Traditional/Community		1,440 (48.5)	
School		1,085 (36.5)	
Public Housing		238 (8.0)	
Other		106 (3.6)	
Native American		99 (3.3)	
Average Daily Attendance	105.21 (87.60)		0.03

Note. Total k = 2,969.

2.4 Measures

2.4.1 Demographics

Participants reported their *gender* and *grade*. *Free/reduced-price lunch* data were obtained from BGCA's member management system (MMS), a data tracking program that is managed by individual Clubs.

2.4.2 Participation Characteristics

Data for *years of membership* and *attendance* were acquired from BGCA's MMS database. *Years of membership* represents the number of years that each young person had been an active BGCA member at the time of survey completion. *Attendance* represents the average number of times per week that the member attended the Club. Three categories of attendance were created based on the MMS dataset: "occasional attenders" (attending the Club less than one time per week on average), "engaged attenders" (attending the Club one time per week on average), and "highly engaged attenders" (attending the Club two or more times per week on average).

2.4.3 Club Characteristics

Club characteristics were reported by regional BGCA organizations through an annual survey. *Average daily attendance* represents the average number of young people who participated in the Club each day during the school year. *Community location* indicates the type of community in which the Club was located (urban, suburban, or rural). *Facility type* reflects the setting in which the Club operated (traditional/community, school, public housing, Native American, and other setting).

2.4.4 Program Quality/Club Experience Indicators

All seven of the Club experience measures were developed by BGCA in collaboration with Youth Development Strategies, Inc. (Boys & Girls Clubs of America, 2018d). These measures were developed as part of a larger initiative at BGCA (the National Youth Outcome Initiative) which aims to measure program quality and document BGCA's impact. The purpose of the Club experience indicators is to measure members' perceptions of program quality across seven domains. For the current study, a composite score was generated for each measure by calculating the mean score of the items for each Club experience scale. Likert-type scales with various response options (described below) were used for each item. See Appendix A for the complete measures. Five of these measures were recently used in a previous study and were found to be reliable at both the individual and site level, and were related in expected ways with staff ratings of program quality components (Kuperminc et al., 2019).

Sense of Belonging (Cronbach's alpha = 0.76) was measured with three items that assess how much young people agree with the following statements: "I feel like I belong here [at this Club]," "I feel like my ideas count here," and "People listen to me here." The four response options ranged from "Strongly Agree" to "Strongly Disagree."

Emotional Safety (Cronbach's alpha = 0.74) included four items that examine the extent to which young people agree that they are respected by staff and other members at the Club, that the Club has rules for how people are supposed to treat each other, and that these rules are enforced. The four response options ranged from "Strongly Agree" to "Strongly Disagree."

Physical Safety (Cronbach's alpha = 0.61) was measured with two items: "I feel safe when I am at this Club" and "If someone wanted to hurt or beat me up here, someone at this

Club would stop them." This measure included four response options that ranged from "All of the time" to "Never."

Fun (Cronbach's alpha = 0.81) included three items that assess the extent to which the member has a good time at the Club, enjoys going, and has more fun there compared to other places. The four response options ranged from "Not at all true" to "Very true."

Adult Caring Relationships (Cronbach's alpha = 0.83) was measured with six items that assess perceptions of how many staff pay attention to what's going on in their life, say something nice when they do something good, could be sought out if they were mad/upset or needed help or advice, and would talk to them if something in their life wasn't going right. The four response options were "None," "One," "Two or three," and "More than three."

Staff Expectations (Cronbach's alpha = 0.75) included three items that examine whether there is an adult at the Club who believes that the young person will be a success, who expects the young person to follow the rules, and who always wants the young person to do her/his best. The four response options ranged from "Not at all true" to "Very true."

Staff Recognition (Cronbach's alpha = 0.81) was assessed by three items that measure whether staff at the Club reward young people when they do a good job, let others know when young people do a good job, and notice when young people try hard. The four response options ranged from "Not at all true" to "Very true."

2.4.5 Social Competence

All four measures of social competence were developed by BGCA and Schulman, Ronca & Bucuvalas, Inc. (SRBI). These measures were also developed as part of BGCA's National Youth Outcome Initiative to measure positive developmental outcomes and document BGCA's

impact. Items composing each measure were averaged to calculate a composite score. Likert-type scales were used for each item. See Appendix A for the complete measures.

Leadership (Cronbach's alpha = 0.74) included four items assessing whether young people agree that they are good at organizing a team of kids to complete a project, planning what needs to be done, making sure everyone in a group feels important, and standing up for what is right. The four response options ranged from "Strong Agree" to "Strongly Disagree."

Conflict Resolution (Cronbach's alpha = 0.63) included five items assessing whether young people do the following when they have problems with other people their age: talk to an adult, talk things over with the people in question, push or hit the other people (reverse coded), and fight back (reverse coded). The last item asks whether young people fight back when other people their age try to hit or push them (reverse coded). The four response options ranged from "Very True" to "Not true at all."

Teamwork (Cronbach's alpha = 0.60) was measured with two items assessing young people's willingness to do whatever the group needs them to do and whether they listen to what other people say. The four response options ranged from "Very True" to "Not true at all."

Concern for Others (Cronbach's alpha = 0.77) was assessed by three items that measure whether young people try to help when they see people in need, think about how other people will be affected when they make a decision, and want to help when they see someone having a problem. The four response options ranged from "Strongly Disagree" to "Strongly Agree." This measure was included in an optional module that some Clubs administered. Therefore, only a subset of participants completed this measure (78,092 (46.9%) of the 166,640 survey respondents in 2017, and 27,919 (61.4%) of the 45,474 survey respondents in 2018).

2.5 Data Analysis

Descriptive statistics were examined, including means, standard deviations, and correlations among all study variables. Data were assessed for outliers, non-normality, and skew using histograms and scatterplots in IBM SPSS Statistics (version 25). Since most variables were not normally distributed, maximum likelihood estimation with robust standard errors (the MLR estimator in Mplus (version 8.2)) was used to account for non-normal distributions in the study models. The local independence assumption of latent profile analysis specifies that the correlations among indicators are fully explained by the latent profile variable (Collins & Lanza, 2010). This assumption was assessed by examining bivariate residuals among each pair of indicators, and this assumption was relaxed by including significant residual associations among pairs of indicators in the final model (Asparouhov & Muthén, 2015; Breslau et al., 2005; B. Muthén, 2017; Muthén & Muthén, 2017; Vermunt & Magidson, 2002).

To address the first aim of this study, latent profile analysis was used to categorize young people into distinct subgroups based on their perceptions of the seven indicators of program quality in 2017 (Asparouhov & Muthén, 2018; Muthén & Muthén, 2017). To conduct the latent profile analysis, the manual (not automatic) Bolck, Croon, and Hagenaars (BCH; 2004) method in Mplus was used to assign cases to their most likely profiles. The BCH method is a stepwise approach to latent profile analysis in which cases are assigned to profiles in the first step, and subsequent models examine the associations among the latent profiles and predictor and/or outcome variables (Asparouhov & Muthén, 2018).

The BCH method was selected because it has been found to be one of the most robust approaches to stepwise latent profile modeling and it is recommended for analyses that involve continuous outcome variables (even when outcome variables are not normally distributed)

(Asparouhov & Muthén, 2018; Bakk & Vermunt, 2016). In addition, other stepwise latent profile modeling methods can present difficulties when predictor and/or outcome variables are included because cases can switch to different profiles than the profiles that they were originally assigned in the first model without predictor and/or outcome variables. In contrast, the BCH method maintains the same latent profiles whether or not predictor or outcome variables are included. The BCH method accounts for uncertainty of each participant's assignment to a profile by utilizing weights that are applied in the models with predictor and/or outcome variables. For all analyses, the "cluster" option was used to account for potential non-independence of the data since participants are clustered in Clubs (Mäkikangas et al., 2018; L. Muthén & Muthén, 2017). As described above, separate models were examined for two age groups (9-12 and 13-18 years old) for all study aims.

The number of profiles was determined by examining multiple criteria, including the Bayesian Information Criterion (BIC), entropy, and the Lo–Mendell–Rubin likelihood ratio test (LMR LRT) (Morgan et al., 2016; Nylund et al., 2007). BIC is a measure of model fit, and lower BIC values indicate better fitting models (Kline, 2005; Nylund et al., 2007). Entropy is a measure of how distinguishable the latent profiles are based on the data, and values closer to one represent better model fit (Morgan et al., 2016). The LMR LRT is a test of the difference in log likelihoods between the model and a similar model with one less profile ($k - 1$) (Nylund et al., 2007). This test generates a p value, and p values below 0.5 suggest that the $k - 1$ model should be rejected (and thus the k model is favored). In addition, the interpretability of the profiles was taken into account when determining the final model (Collins & Lanza, 2010).

To examine whether the latent profiles were associated with social competence (aim 2), three measures of social competence in 2017 (leadership, teamwork, and conflict resolution)

were regressed on the latent profile variable (see Figure 2.1) (Asparouhov & Muthén, 2018). This model included individual-level characteristics (e.g., demographics, participation) as covariates. Two dummy codes were generated for *attendance* with occasional attenders as the reference group (compared to engaged attenders and highly engaged attenders). A similar model was examined separately with Concern for Others as the outcome variable since only some participants were administered this measure at some Clubs (see Figure 2.2).

The third aim of this study extended the second aim by examining the *longitudinal* associations among the latent profiles and each social competence outcome variable (see Figures 2.3 and 2.4). A similar procedure was followed as for the second aim except that the sample included only young people who completed surveys in both 2017 and 2018 since only a small proportion (approximately 25%) of young people who completed the 2017 survey also completed the 2018 survey. The latent profiles generated for the first aim of this study were used for the longitudinal analysis. The 2018 data were used for each social competence outcome variable and each outcome variable was regressed on the corresponding 2017 variable to account for change across one year.

For the last aim of this study, chi-square tests were used to examine potential differences in the proportion of profiles across Club characteristics (e.g., Smith, Peck, Denault, Blazevski, & Akiva, 2010). The Club characteristics assessed were community location (urban, suburban, and rural), facility type (traditional, school, public housing, Native American, and other type), and Club size (average daily attendance separated into three categories based on BGCA's categorization of small, medium and large sites: 0-49, 50-174, and 175 or larger).

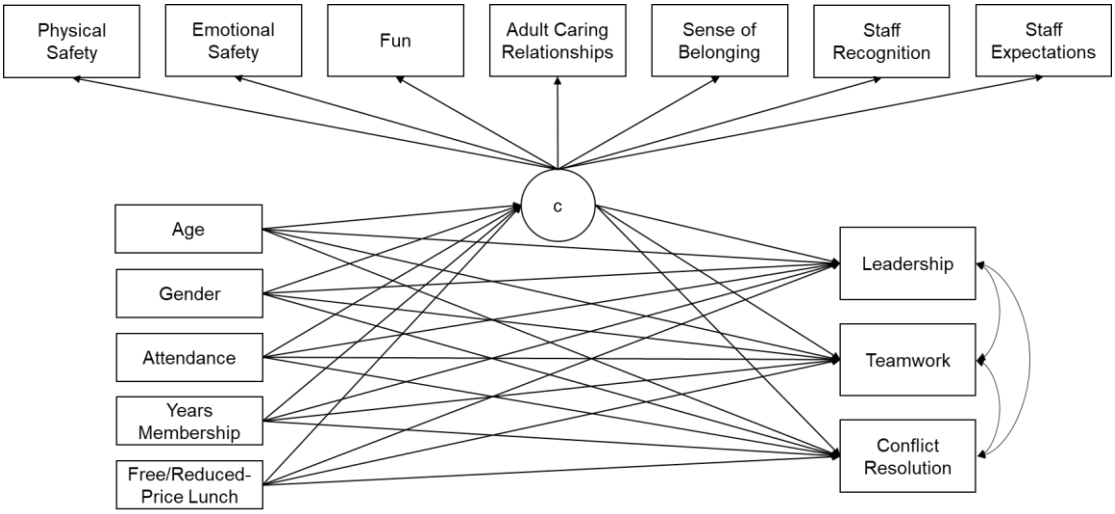

Figure 2.1 Cross-sectional mixture model with social competence variables regressed on the latent profiles of program quality and individual-level covariates. All exogenous variables were correlated in the model. Separate models were analyzed for each age group (9-12 and 13-18 years old).

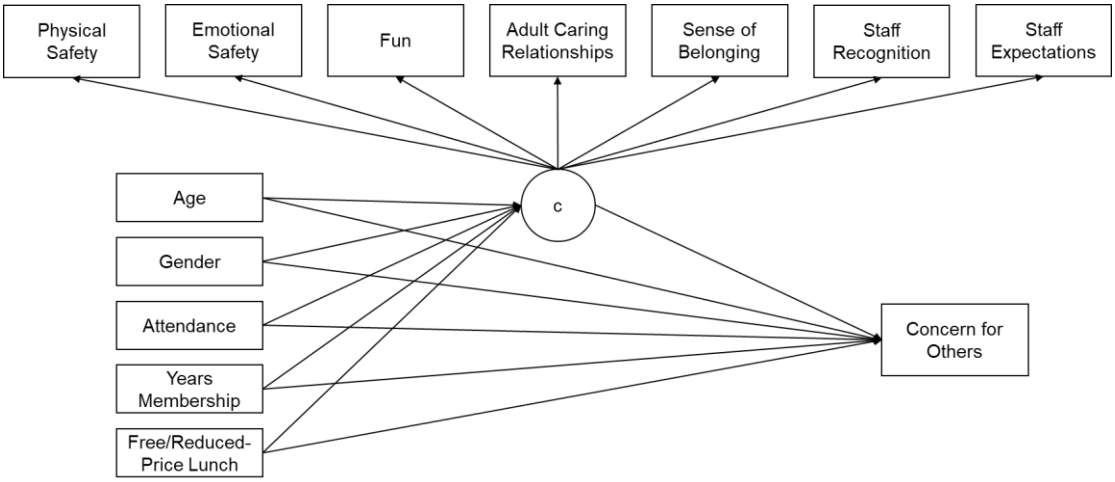

Figure 2.2 Cross-sectional mixture model with Concern for Others regressed on the latent profiles of program quality and individual-level covariates. All exogenous variables were correlated in the model. Separate models were analyzed for each age group (9-12 and 13-18 years old).

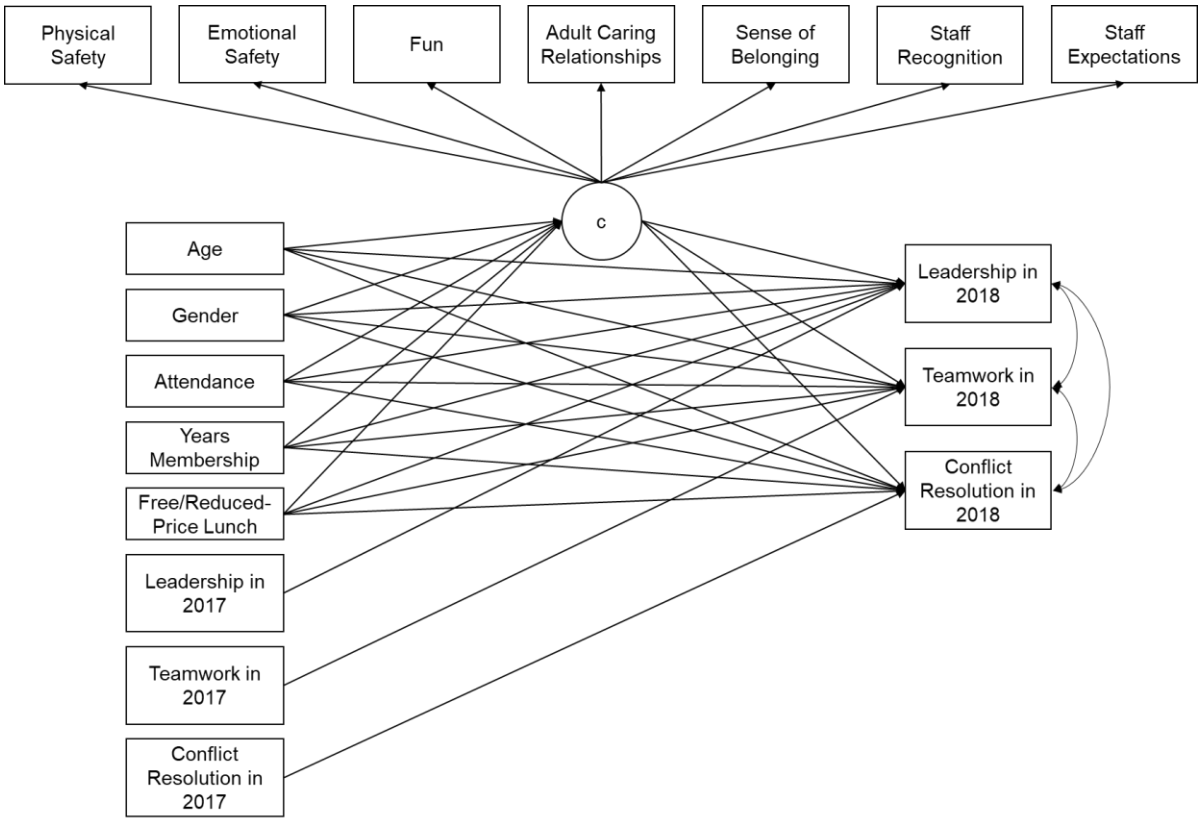

Figure 2.3 Longitudinal mixture model with social competence variables in 2018 regressed on the latent profiles of program quality and individual-level covariates (all measured in 2017). All exogenous variables were correlated in the model. Separate models were analyzed for each age group (9-12 and 13-18 years old).

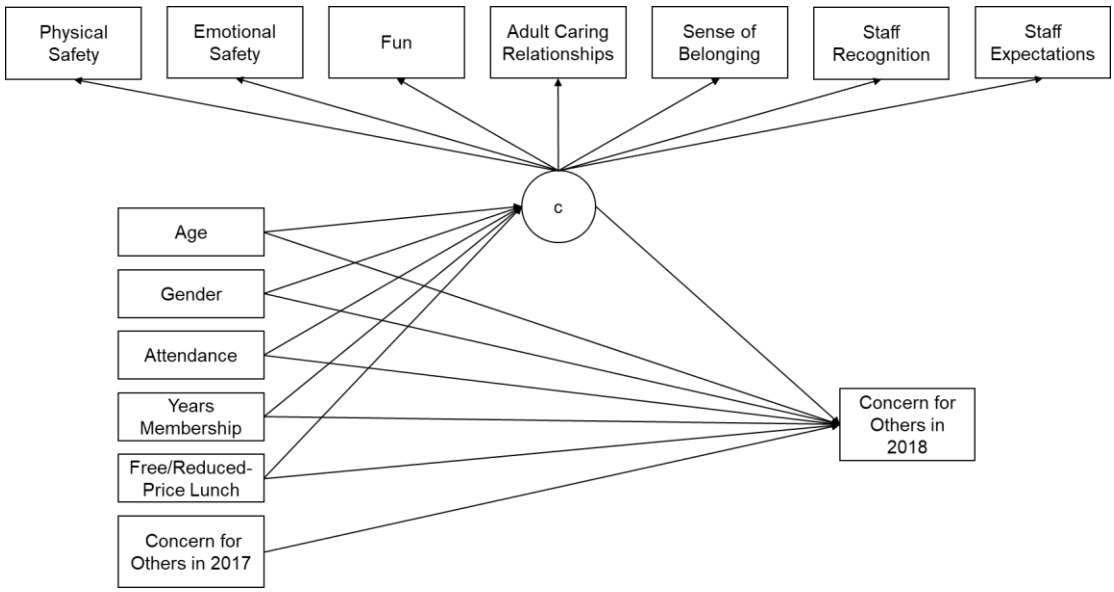

Figure 2.4 Longitudinal mixture model with Concern for Others in 2018 regressed on the latent profiles of program quality and individual-level covariates (all measured in 2017). All exogenous variables were correlated in the model. Separate models were analyzed for each age group (9-12 and 13-18 years old).

3 RESULTS

3.1 Preliminary Analyses

To provide insight into the magnitude of statistically significant results, I will include effect sizes when possible. For this study, effect sizes are categorized as small (i.e., Cohen's *d* of 0.20 – 0.49, Pearson's correlation coefficient *r* of 0.10 – 0.29), medium (i.e., Cohen's *d* of 0.50 – 0.79, Pearson's correlation coefficient *r* of 0.30 – 0.49), and large (i.e., Cohen's *d* of 0.80 and above, Pearson's correlation coefficient *r* of 0.50 and above; Cohen, 1988; Kim, 2017). When interpreting odds ratios, Chen and colleagues' (2010) guidelines were used: odds ratios from 1.50 to 2.49 were considered small in magnitude; from 2.50 to 4.09, medium in magnitude; and 4.10 and above, large in magnitude (e.g., Deshpande et al., 2020; Ruiz & Dorritie, 2020). The guidelines for Cramer's *V* depend on specific degrees of freedom (Gravetter & Wallnau, 2013). For two degrees of freedom, Cramer's *V* from 0.07 to 0.20 was considered small in magnitude; from 0.21 to 0.34, medium in magnitude; and 0.35 and above, large in magnitude. For four degrees of freedom, Cramer's *V* from 0.05 to 0.14 was considered small in magnitude; from 0.15 to 0.24, medium in magnitude; and 0.25 and above, large in magnitude. Throughout the results section I will focus on effects that are small, medium, or large in magnitude.

The means and standard deviations for the primary study variables are displayed in Table 3.1. The difference in each of the four outcome variables from 2017 to 2018 was significant for youth (9-12 years old) and teens (13-18 years old) at $p < .05$ after Bonferroni correction to account for each set of four comparisons (0.05 / 4 = 0.01). The means for leadership, conflict resolution, teamwork, and concern for others tended to decline for youth (d = 0.04 to 0.19) and

increase for teens (d = 0.03 to 0.06) However, the differences were modest and did not meet the criteria for a small effect.

Table 3.1 Means and Standard Deviations for Primary Study Measures

	All Ages		Youth (9-12 years old)		Teens (13-18 years old)	
Measure	2017	2018	2017	2018	2017	2018
Program Quality						
Belonging	3.06 (0.68)		3.02 (0.67)		3.17 (0.67)	
Emotional Safety	3.22 (0.62)		3.19 (0.61)		3.28 (0.63)	
Fun	3.22 (0.73)		3.19 (0.75)		3.33 (0.69)	
Adult Connections	2.87 (0.75)		2.84 (0.74)		2.96 (0.79)	
Staff Expectations	3.61 (0.58)		3.60 (0.59)		3.64 (0.55)	
Staff Recognition	3.12 (0.81)		3.07 (0.83)		3.24 (0.75)	
Physical Safety	3.36 (0.76)		3.30 (0.77)		3.51 (0.68)	
Leadership	3.32 (0.57)	3.28 (0.57)	3.32 (0.58)	3.27 (0.58)	3.32 (0.57)	3.33 (0.55)
Conflict Resolution	2.80 (0.71)	2.76 (0.70)	2.85 (0.73)	2.76 (0.72)	2.68 (0.63)	2.74 (0.63)
Teamwork	3.42 (0.65)	3.41 (0.64)	3.40 (0.66)	3.39 (0.65)	3.45 (0.62)	3.51 (0.59)
Concern for Others	3.34 (0.60)	3.29 (0.59)	3.34 (0.60)	3.27 (0.60)	3.35 (0.58)	3.35 (0.58)
Overall *n*	166,640	45,474	120,780	35,945	45,860	9,529
Concern for Others *n*	78,092	27,919	61,313	22,997	16,779	4,922

Note. Range of possible scores is 1-4 for each variable; Standard deviations are included in parentheses; Program quality measures in 2018 were not included in this study; Change from 2017 to 2018 was significant for all variables measured both years at $p < .05$ after Bonferroni correction to account for each set of four comparisons (0.05 / 4 = 0.01); The sample size for the Concern for Others variable is smaller than the overall sample size because this measure was included in an optional module that was only administered to youth and/or teens at some Clubs

3.1.1 Correlations

Correlations among the primary study variables can be found in Tables 3.2 and 3.3. All of the correlations among program quality indicators (e.g., sense of belonging, fun) were positive and medium to large in magnitude. The correlations among each program quality indicator and each outcome variable (e.g., leadership, conflict resolution) in 2017 and 2018 were positive, and most were small to medium. The correlations among all outcome variables in 2017 and 2018 were positive and most were small to medium, except the correlations between Leadership and Concern for Others, which were large in both years. Two of the correlations between demographic variables and 1) the program quality indicators and 2) the outcome variables reached the cut-off for a small effect size. Specifically, age was positively associated with Belonging and Physical Safety, and negatively associated with Conflict Resolution in 2017. In addition, male gender identity was negatively correlated with Conflict Resolution in 2017 and 2018. Even though the magnitude of the correlations between some of the demographic variables and program quality indicators/outcome variables did not reach the cut-off for a small effect, all demographic variables were included in the main study models as covariates for the sake of comprehensiveness.

Table 3.2 Correlations among Primary Study Variables, Part 1

	1	2	3	4	5	6	7	8	9	10	11
1.Age	1.00										
2.Free/Reduced-Price Lunch	0.05*	1.00									
3.Attend Club 1x/week	0.09*	-0.02*	1.00								
4.Attend Club 2+x/week	-0.18*	0.05*	-0.70*	1.00							
5.Years Membership	0.17*	-0.02*	0.04*	-0.02*	1.00						
6.Gender (Boy)	0.05*	-0.01	0.01	-0.02*	0.01	1.00					
7.Belonging	0.11*	0.03*	0.00	-0.03*	0.01	0.01*	1.00				
8.Emotional Safety	0.06*	0.03*	0.00	-0.03*	-0.01*	0.01*	0.72*	1.00			
9.Fun	0.08*	0.07*	0.00	-0.04*	-0.06*	0.05*	0.55*	0.45*	1.00		
10.Adult Connections	0.08*	0.04*	-0.01	0.01*	0.04*	0.02*	0.41*	0.34*	0.37*	1.00	
11.Staff Expectations	0.03*	0.04*	-0.01	0.01	-0.01	-0.03*	0.43*	0.41*	0.44*	0.41*	1.00
12.Recognition	0.08*	0.08*	0.00	-0.02*	-0.01	0.01*	0.51*	0.44*	0.54*	0.45*	0.61*
13.Physical Safety	0.13*	0.01*	0.02*	-0.05*	0.00	-0.05*	0.50*	0.49*	0.49*	0.36*	0.44*
14.Leadership 2017	0.01*	0.02*	0.00	0.00	0.00	-0.04*	0.32*	0.27*	0.28*	0.26*	0.32*
15.Conflict Resolution 2017	-0.13*	-0.09*	-0.01	0.01	-0.03*	-0.17*	0.23*	0.25*	0.23*	0.17*	0.19*
16.Teamwork 2017	0.03*	-0.03*	0.01	-0.03*	-0.01*	-0.03*	0.30*	0.28*	0.32*	0.22*	0.30*
17.Concern for Others 2017	0.00	0.01	0.00	-0.01	-0.02*	-0.05*	0.31*	0.28*	0.28*	0.26*	0.29*
18.Leadership 2018	0.04*	0.01	0.01	-0.01	0.03*	-0.06*	0.20*	0.15*	0.16*	0.17*	0.18*
19.Conflict Resolution 2018	-0.04*	-0.09*	0.01	-0.02	-0.02	-0.15*	0.16*	0.16*	0.16*	0.09*	0.12*
20.Teamwork 2018	0.08*	-0.04*	0.02	-0.03*	0.02	-0.04*	0.17*	0.15*	0.17*	0.12*	0.14*
21.Concern for Others 2018	0.04*	0.00	0.02	-0.01	0.01	-0.07*	0.18*	0.16*	0.18*	0.17*	0.17*

Note. Overall N = 166,640 and the sample size varies for each pairwise comparison due to missing data.
*Significant at $p < .05$ after Bonferroni correction (.05 / 210 = .0002).

Table 3.3 Correlations among Primary Study Variables, Part 2

	12	13	14	15	16	17	18	19	20	21
12.Recognition	1.00									
13.Physical Safety	0.46*	1.00								
14.Leadership 2017	0.33*	0.27*	1.00							
15.Conflict Resolution 2017	0.20*	0.25*	0.17*	1.00						
16.Teamwork 2017	0.31*	0.30*	0.37*	0.30*	1.00					
17.Concern for Others 2017	0.32*	0.26*	0.54*	0.26*	0.38*	1.00				
18.Leadership 2018	0.20*	0.18*	0.35*	0.10*	0.17*	0.28*	1.00			
19.Conflict Resolution 2018	0.14*	0.17*	0.12*	0.48*	0.21*	0.18*	0.20*	1.00		
20.Teamwork 2018	0.15*	0.17*	0.18*	0.20*	0.26*	0.20*	0.40*	0.35*	1.00	
21.Concern for Others 2018	0.20*	0.18*	0.27*	0.18*	0.20*	0.32*	0.55*	0.30*	0.42*	1.00

Note. Overall N = 166,640 and the sample size varies for each pairwise comparison due to missing data.
*Significant at $p < .05$ after Bonferroni correction (.05 / 210 = .0002).

3.1.2 Missing Data

Missing data for the indicators of program quality, Teamwork, Leadership, Conflict Resolution, and site-level variables (i.e., average daily attendance, community location, facility type) ranged from 0.00% to 3.98% in 2017. The Concern for Others measure was included in an optional module that was only administered to youth (9-12 years old) and/or teens (13-18 years old) at some Clubs. In 2017, 46.9% of the survey participants completed this module (among those, 6.30% of the data were missing for Concern for Others). Missing data for demographic covariates ranged from 0.12% to 37.59%. Missing data on demographic covariates were above 20% for Frequency of Attendance, Years Membership, and Free/Reduced-Price Lunch because this data is not reported by some Clubs.

Of the 166,640 participants that completed the survey in 2017, 27.3% also completed the survey in 2018. The current study was not designed to be longitudinal, and the study is taking advantage of the sizeable number of youth who happen to be sampled in consecutive years. The longitudinal models accounted for factors that might contribute to only some youth taking the survey multiple years (described below). For the participants that completed the survey in both years, missing data for the main outcome variables in 2018 ranged from 1.41% to 2.28%. The 2018 Concern for Others module was administered to 61.4% of participants that completed the survey in 2018 (and 3.47% of the data were missing for Concern for Others in 2018).

Three different types of missing data were addressed with separate approaches. First, missing data for program quality indicators were mostly due to item non-response, and these missing data were addressed with full information maximum likelihood estimation, which is the default option in Mplus (Muthén, 2018). This approach allowed for the generation of one set of latent profiles that could be used in all of the analyses. Second, most of the missing data on

exogenous variables (i.e., covariates) is due to incomplete administrative records. Since missing data on exogenous variables cannot be addressed with full information maximum likelihood estimation using Mplus's BCH latent profile analysis method, multiple imputation in Mplus was used to create 30 datasets with complete data (Asparouhov, 2017). A multi-level imputation model was utilized to account for all individual- and Club-level study variables.

Finally, the longitudinal models only included 27.3% of the total sample and therefore may not be representative of the larger sample of youth who completed the 2017 survey. Baseline differences between participants that completed the 2017 survey only and those that completed both surveys were examined on all primary study variables using *t*-tests for continuous variables and chi-square tests for categorical variables. Only two of these comparisons were statistically significant ($p < .05$ after Bonferroni correction (0.05 / 16 = 0.003)) and met the minimum criteria for a small effect size. Specifically, participants that completed both surveys were younger ($M = 11.02$ years old, $SD = 2.04$) and were more likely to attend the Club two or more times per week (82.61%) than participants that only completed the 2017 survey ($M = 11.63$ years old, $SD = 2.26$; 65.6%), $t(166{,}187) = 53.40$, $p < .003$, d = 0.29 and $\chi^2(2, 104{,}007) = 3384.54$, $p < .003$, V = 0.18, respectively. To approximate the cross-sectional sample that completed the survey in 2017, sample weights were included in the longitudinal analyses (Seaman & White, 2013). These sample weights represent the inverse probability of completing the 2018 survey, and they were calculated using a binary logistic regression with youth covariates (e.g., gender, free/reduced-price lunch, frequency of BGCA attendance, duration of BGCA participation) and 2017 data on program quality and social competence as predictors of survey completion in 2018. The longitudinal models were analyzed both with and

without sample weights to examine potential differences in the results that may be attributed to missing data.

3.1.3 Differences between Participants Who Were and Were Not Administered the Concern for Others Module

Baseline differences between participants who were and were not administered the Concern for Others module were examined on all primary study variables. Comparisons were examined using *t*-tests for continuous variables and chi-square tests for categorical variables.

For the 2017 survey, there were significant differences between participants who were and were not administered the Concern for Others module for all baseline variables except Emotional Safety, Physical Safety, Conflict Resolution, and Teamwork ($p < .05$ after Bonferroni correction (0.05 / 15 = 0.003)). Only one of these comparisons met the minimum criteria for a small effect size. Specifically, participants who were administered the Concern for Others module were younger (M = 11.21 years old, SD = 2.08) than participants who were not administered the module (M = 11.69 years old, SD = 2.31), $t(166,187) = 43.67$, $p < .003$, d = 0.22.

For the 2018 survey, there were significant differences between participants who were and were not administered the Concern for Others module for the following baseline variables and 2018 outcomes variables: Age, Frequency of Attendance, Adult Connections, Conflict Resolution in 2017, and Leadership in 2018 ($p < .05$ after Bonferroni correction (0.05 / 18 = 0.003)). Only one of these comparisons met the minimum criteria for a small effect size: Participants who were administered the Concern for Others module were younger (M = 10.83 years old, SD = 1.91) than participants who were not administered the module (M = 11.31 years old, SD = 2.19), $t(45,390) = 23.70$, $p < .003$, d = 0.23.

3.2 Primary Analyses

3.2.1 Latent Profile Analyses

Latent profile analyses were conducted for models with two-to-eight profiles (Ferguson et al., 2019). To reduce the chance of obtaining a local maximum solution, each model was run with 2,000 random starts (Nylund-Gibson & Choi, 2018). The local independence assumption of latent profile analysis specifies that the correlations among indicators are fully explained by the latent profile variable (Collins & Lanza, 2010). This assumption was assessed by examining bivariate residual correlations among each pair of indicators for each latent profile model (Asparouhov & Muthén, 2015). All bivariate residuals among each pair of indicators were statistically significant at $p < .05$ after Bonferroni correction (.05 / 21 = .002) for the final models. Therefore, the local independence assumption was relaxed by including residual bivariate associations among all pairs of indicators in the final models (Asparouhov & Muthén, 2015; Breslau et al., 2005; B. Muthén, 2017; Muthén & Muthén, 2017; Vermunt & Magidson, 2002). As noted above, the primary analyses will be presented for youth (9-12 years old) and teens (13-18 years old) separately.

3.2.1.1 Youth Models

Table 3.4 includes fit indices and descriptive information for each latent profile analysis model that included the youth participants (9-12 years old). Entropy is above 0.80 for all solutions, suggesting minimal uncertainty when classifying individuals (Ferguson et al., 2019). As the number of latent profiles increased, the BIC and sample-size adjusted BIC decreased. Similarly, the LMR LRT was significant for each model, indicating that the *k* profile solution was significantly better than the *k*-1 profile solution. These results suggest better fitting models

as the number of profiles increased. The interpretability of the profiles was then used to select the best profile solution.

Table 3.4 Fit Statistics for Latent Profile Analysis Models for Program Quality Indicators among Youth

k	Log Likelihood	BIC	Sample-Size Adjusted BIC	LMR LRT *p*	Entropy
2	-699790.25	1400083.63	1399946.98	< .001	0.91
3	-683186.69	1366970.10	1366808.02	< .001	0.92
4	-670259.49	1341209.32	1341021.81	< .001	0.96
5	-660381.87	1321547.67	1321334.74	< .001	0.97
6	-653241.42	1306632.84	1307360.39	< .001	0.98
7	-650416.77	1301804.68	1301540.90	< .001	0.94
8[†]	-588343.04	1177750.83	1177461.62	< .001	0.98

Note. *k* = Number of profiles; BIC = Bayesian information criterion; LMR-LRT = Lo-Mendell-Rubin adjusted likelihood ratio test; The bootstrap likelihood ratio test (BLRT) is not available when accounting for clustering; *N* = 120,643. [†]The maximum log likelihood value was not replicated, suggesting that too many profiles were being extracted and/or poor model fit (Nylund-Gibson & Choi, 2018).

Based on interpretability of the profiles, the five-profile solution was most appropriate (see Figure 3.1 and Table 3.5). This solution contains four profiles in which youth ratings were similar for all program quality indicators: "Above Average" (*n* = 84,488; 70.03%), "Below Average" (*n* = 22,825; 18.92%), "Very Below Average" (*n* = 10,046; 8.33%), and "Negative" (*n* = 2,187; 1.81%). The fifth profile included youth who were above average for all program quality indicators except for Belonging and Emotional Safety ("Low Belonging and Emotional Safety"; *n* = 1,097; 0.91%). It is important to note that the profile labels reflect program quality ratings relative to the mean levels. The average scores for each indicator of program quality were quite high (see Table 3.1). Thus, even for the "Below Average" and "Very Below Average" profiles the ratings for most of the program quality indicators reflect positive-to-neutral perceptions of program quality. The solutions with additional profiles had similar profiles to the

five-profile solution, except the profiles with below average ratings for all program quality indicators were separated into additional profiles, which limited interpretability (see Appendix B for figures of all the profile solutions). Although the five-profile solution contains two profiles with less than five percent of youth participants, there was a sufficient number of youth participants in each profile to suggest generalizability (Ferguson et al., 2019).

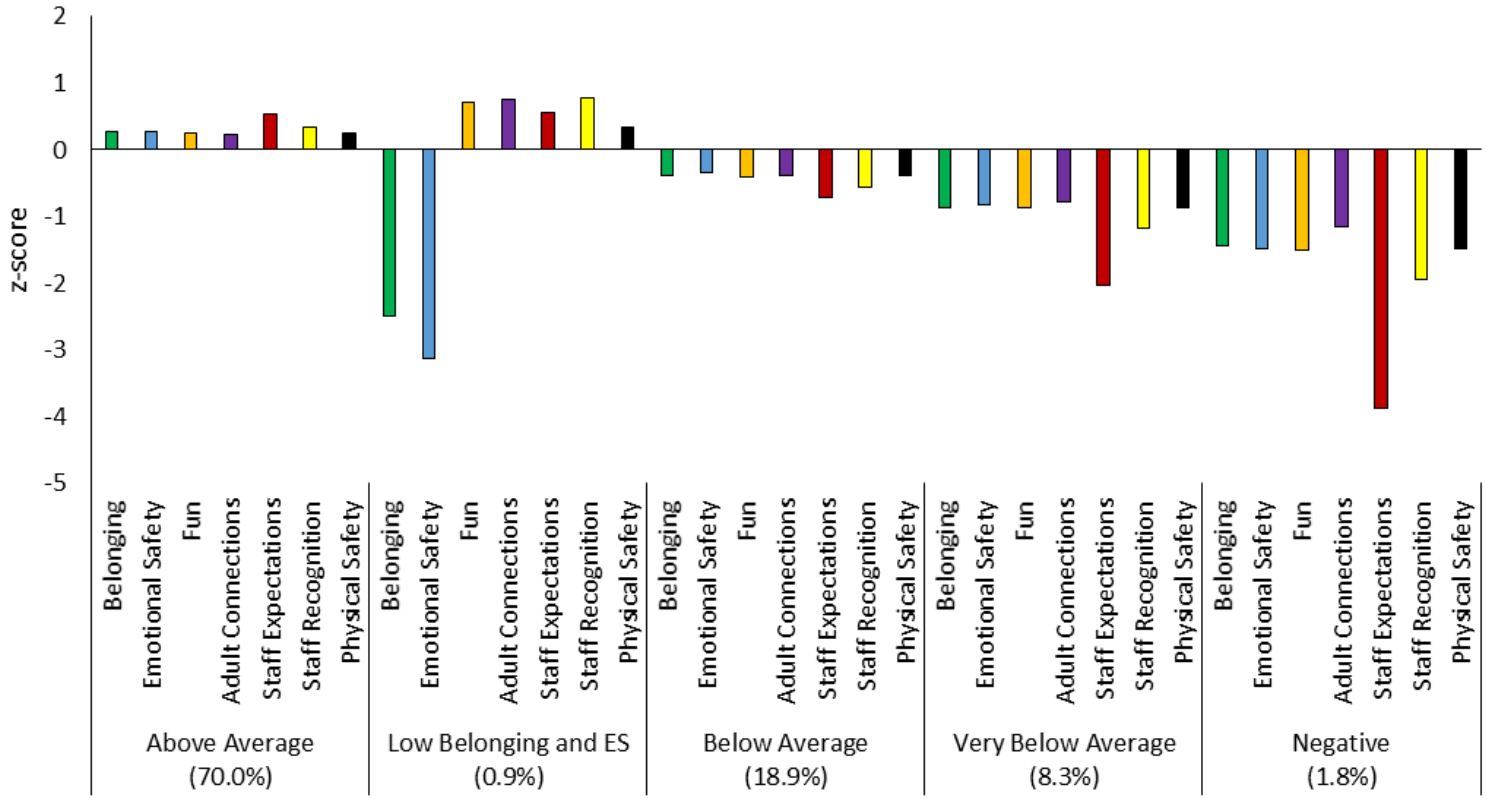

Figure 3.1 Five-profile solution for indicators of program quality among youth (9-12 years old). ES = Emotional Safety; n = 120,643.

Table 3.5 Means of Program Quality Indicators for the Five-Profile Solution for Youth (9-12 Years Old)

Variable	Profile 1: Above Average	Profile 2: Low Belonging and ES	Profile 3: Below Average	Profile 4: Very Below Average	Profile 5: Negative
Belonging	3.21 (0.59)	1.34 (0.59)	2.76 (0.59)	2.44 (0.59)	2.05 (0.59)
Emotional Safety	3.36 (0.53)	1.27 (0.53)	2.98 (0.53)	2.69 (0.53)	2.28 (0.53)
Fun	3.37 (0.68)	3.72 (0.68)	2.89 (0.68)	2.54 (0.68)	2.06 (0.68)
Adult Connections	3.01 (0.68)	3.40 (0.68)	2.55 (0.68)	2.26 (0.68)	1.99 (0.68)
Staff Expectations	3.92 (0.17)	3.94 (0.17)	3.18 (0.17)	2.40 (0.17)	1.31 (0.17)
Staff Recognition	3.35 (0.68)	3.71 (0.68)	2.61 (0.68)	2.10 (0.68)	1.46 (0.68)
Physical Safety	3.49 (0.70)	3.56 (0.70)	3.00 (0.70)	2.63 (0.70)	2.15 (0.70)
n	84,488	1,097	22,825	10,046	2,187
% of all youth	70.03%	0.91%	18.92%	8.33%	1.81%

Note. Range of possible scores is 1-4 for each variable; Standard deviations are included in parentheses; ES = Emotional Safety.

3.2.1.2 Teen Models

Table 3.6 includes fit indices and descriptive information for each latent profile analysis model that included teen participants (13-18 years old). Entropy is above 0.80 for all solutions, suggesting minimal uncertainty when classifying individuals (Ferguson et al., 2019). As the number of latent profiles increased, the BIC and sample-size adjusted BIC decreased. Similarly, the LMR LRT was significant for each model, indicating that the *k* profile solution was significantly better than the *k*-1 profile solution. These fit indices indicate that the better fitting models have more profiles. Interpretability was then used to determine the best profile solution.

Table 3.6 Fit Statistics for Latent Profile Analysis Models for Program Quality Indicators among Teens

k	Log Likelihood	BIC	Sample-Size Adjusted BIC	LMR LRT *p*	Entropy
2	-245698.31	491858.12	491721.47	< .001	0.99
3	-237773.08	476093.50	475931.42	< .001	0.97
4	-228705.94	458045.10	457857.60	< .001	0.97
5	-223572.28	447863.62	447650.70	< .001	0.97
6[†]	-218224.80	437254.53	437016.18	< .001	0.98
7	-207114.77	415120.32	414856.55	< .001	0.99
8	-187755.47	376487.59	376198.39	< .001	0.99

Note. *k* = Number of profiles; BIC = Bayesian information criterion; LMR-LRT = Lo-Mendell-Rubin adjusted likelihood ratio test; The bootstrap likelihood ratio test (BLRT) is not available when accounting for clustering; *n* = 45,816. [†]The maximum log likelihood value was not replicated, suggesting that too many profiles were being extracted and/or poor model fit (Nylund-Gibson & Choi, 2018).

Based on interpretability of the profiles, the five-profile solution was most appropriate (see Figure 3.2 and Table 3.7). This solution included similar profiles to the best profile solution for youth: “Above Average” (*n* = 31,710; 69.21%), “Below Average” (*n* = 11,271; 24.60%), “Very Below Average” (*n* = 1,576; 3.44%), “Negative” (*n* = 440; 0.96%), and “Low Belonging

and Emotional Safety" (*n* = 819; 1.79%). It is important to note that the profile labels reflect program quality ratings relative to the mean levels. Similar to youth, the average scores for each indicator of program quality were quite high among teens (see Table 3.1). Thus, even for the "Below Average" and "Very Below Average" profiles the ratings for most of the program quality indicators reflect positive-to-neutral perceptions of program quality. As with the youth models, the solutions with six or more profiles had similar profiles to the five-profile solution, except the profiles with below average ratings for all program quality indicators were separated into even more profiles, which limited interpretability (see Appendix B for figures of all the profile solutions). Although the five-profile solution contains three profiles with less than five percent of teen participants, there was a sufficient number of teen participants in each profile to suggest generalizability (Ferguson et al., 2019).

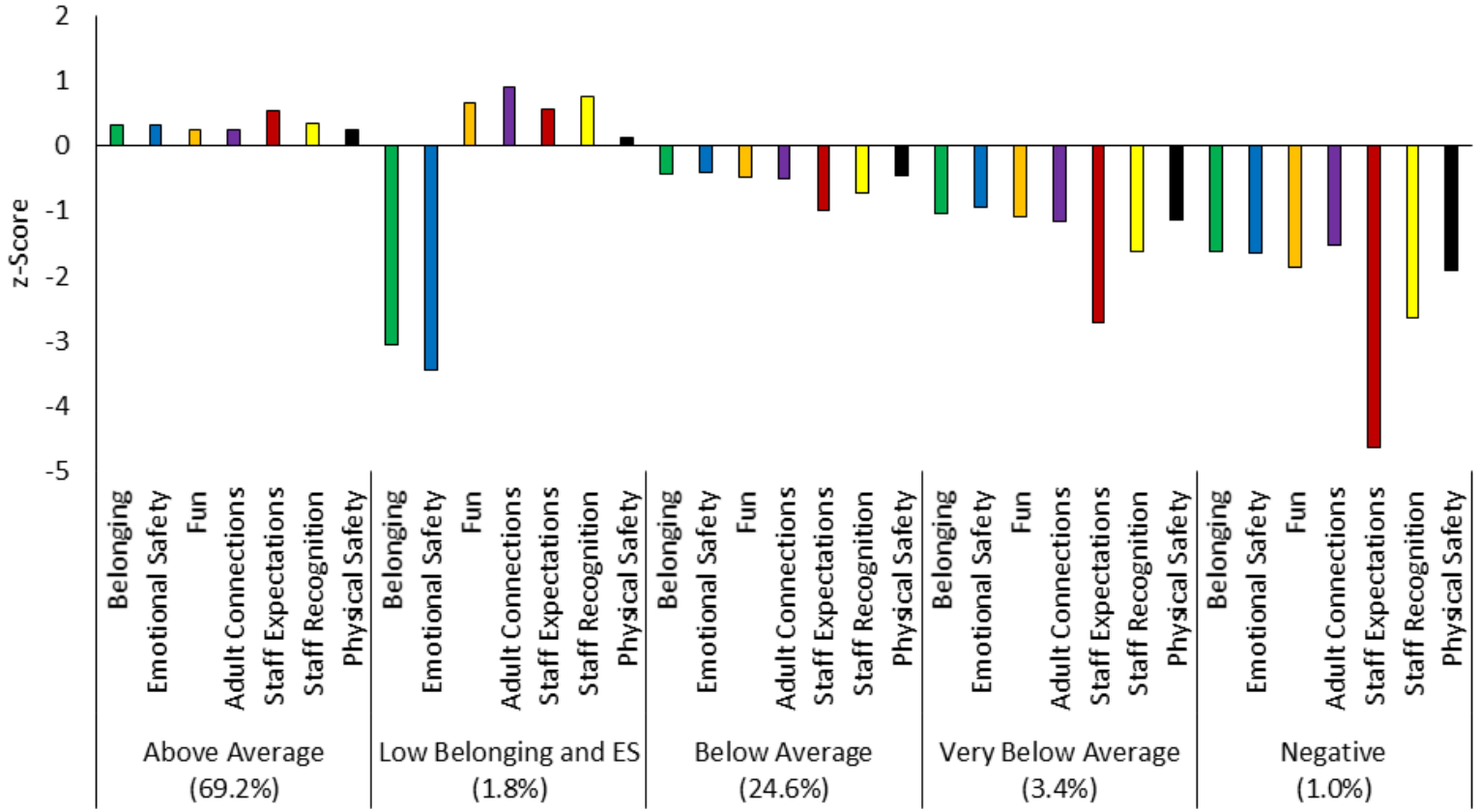

Figure 3.2 Five-profile solution for indicators of program quality among teens (13-18 years old). ES = Emotional Safety; n = 45,816.

Table 3.7 Means of Program Quality Indicators for the Five-Profile Solution for Teens (13-18 Years Old)

Variable	Profile 1: Above Average	Profile 2: Low Belonging and ES	Profile 3: Below Average	Profile 4: Very Below Average	Profile 5: Negative
Belonging	3.38 (0.54)	1.13 (0.54)	2.89 (0.54)	2.47 (0.54)	2.09 (0.54)
Emotional Safety	3.47 (0.50)	1.11 (0.50)	3.02 (0.50)	2.68 (0.50)	2.24 (0.50)
Fun	3.49 (0.62)	3.77 (0.62)	2.99 (0.62)	2.58 (0.62)	2.05 (0.62)
Adult Connections	3.15 (0.71)	3.67 (0.71)	2.56 (0.71)	2.04 (0.71)	1.75 (0.71)
Staff Expectations	3.94 (0.16)	3.95 (0.16)	3.09 (0.16)	2.14 (0.16)	1.07 (0.16)
Staff Recognition	3.50 (0.59)	3.80 (0.59)	2.70 (0.59)	2.01 (0.59)	1.25 (0.59)
Physical Safety	3.68 (0.62)	3.59 (0.62)	3.19 (0.62)	2.73 (0.62)	2.20 (0.62)
n	31,710	819	11,271	1,576	440
% of all youth	69.21%	1.79%	24.60%	3.44%	0.96%

Note. Range of possible scores is 1-4 for each variable; Standard deviations are included in parentheses; ES = Emotional Safety.

Tables 3.8 and 3.9 present descriptive statistics for all demographics and outcome variables by program quality profile for youth and teens, respectively. These data will be further discussed in the following sections.

Table 3.8 Descriptive Statistics of Demographics and Outcome Variables by Program Quality Profiles for Youth

	Profile 1: Above Average		Profile 2: Low Belonging and ES		Profile 3: Below Average		Profile 4: Very Below Average		Profile 5: Negative	
	Mean (SD)	Freq. (%)	Mean (SD)	Freq. (%)	Mean (SD)	Freq. (%)	Mean (SD)	Freq. (%)	Mean (SD)	Freq. (%)
Male		42,401 (50.2)		595 (54.4)		11,977 (52.5)		5,517 (55.0)		1,216 (55.7)
Free/Reduced-Price Lunch		41,940 (75.7)		592 (83.0)		10,727 (70.7)		4,654 (69.9)		1,063 (74.7)
Attendance										
Highly Engaged Attender (2+x/week)		40,330 (76.1)		534 (78.2)		10,589 (73.5)		4,761 (74.9)		1,027 (74.4)
Engaged Attender (1x/week)		7,584 (14.3)		95 (13.9)		2,259 (15.7)		917 (14.4)		203 (14.7)
Occasional Attender (<1x/week)		5,113 (9.6)		54 (7.9)		1,560 (10.8)		675 (10.6)		150 (10.9)
Age	10.27 (1.05)		10.43 (1.07)		10.42 (1.06)		10.37 (1.04)		10.46 (1.04)	
Years Membership	1.62 (1.61)		1.66 (1.65)		1.74 (1.66)		1.70 (1.63)		1.71 (1.64)	
Leadership 2017	3.42 (0.52)		3.55 (0.65)		3.13 (0.57)		2.98 (0.66)		2.75 (0.88)	
Conflict Resolution 2017	2.94 (0.73)		2.81 (0.72)		2.70 (0.67)		2.57 (0.68)		2.44 (0.70)	
Teamwork 2017	3.51 (0.60)		3.57 (0.72)		3.23 (0.65)		3.06 (0.75)		2.70 (1.02)	
Concern for Others 2017	3.43 (0.55)		3.52 (0.73)		3.15 (0.58)		3.00 (0.70)		2.82 (0.09)	
Leadership 2018	3.32 (0.56)		3.42 (0.58)		3.13 (0.59)		3.09 (0.64)		3.00 (0.72)	
Conflict Resolution 2018	2.81 (0.72)		2.75 (0.72)		2.67 (0.68)		2.60 (0.70)		2.49 (0.68)	
Teamwork 2018	3.44 (0.63)		3.49 (0.66)		3.29 (0.65)		3.21 (0.71)		3.13 (0.78)	
Concern for Others 2018	3.33 (0.58)		3.44 (0.55)		3.13 (0.59)		3.09 (0.67)		3.04 (0.75)	

Note. N = 120,643; Freq. = Frequency; ES = Emotional Safety.

Table 3.9 Descriptive Statistics of Demographics and Outcome Variables by Program Quality Profiles for Teens

	Profile 1: Above Average		Profile 2: Low Belonging and ES		Profile 3: Below Average		Profile 4: Very Below Average		Profile 5: Negative	
	Mean (SD)	Freq. (%)	Mean (SD)	Freq. (%)	Mean (SD)	Freq. (%)	Mean (SD)	Freq. (%)	Mean (SD)	Freq. (%)
Male		17,476 (55.2)		514 (62.8)		6,513 (57.9)		927 (58.9)		263 (59.8)
Free/Reduced-Price Lunch		15,550 (78.8)		423 (85.5)		5,348 (77.1)		754 (76.6)		197 (72.4)
Attendance										
Highly Engaged Attender (2+x/week)		11,626 (59.4)		285 (56.5)		3,947 (58.2)		526 (56.1)		173 (64.8)
Engaged Attender (1x/week)		4,203 (21.5)		115 (22.8)		1,462 (21.5)		230 (24.5)		55 (20.6)
Occasional Attender (<1x/week)		3,759 (19.2)		104 (20.6)		1,378 (20.3)		181 (19.3)		39 (14.6)
Age	14.57 (1.53)		14.78 (1.60)		14.33 (1.41)		14.06 (1.30)		14.05 (1.26)	
Years Membership	2.27 (2.49)		2.37 (2.59)		2.22 (2.38)		2.15 (2.37)		2.16 (2.38)	
Leadership 2017	3.43 (0.53)		3.65 (0.54)		3.10 (0.53)		2.91 (0.65)		2.61 (1.00)	
Conflict Resolution 2017	2.75 (0.65)		2.68 (0.66)		2.52 (0.55)		2.44 (0.57)		2.44 (0.57)	
Teamwork 2017	3.56 (0.56)		3.69 (0.58)		3.22 (0.61)		2.97 (0.78)		2.52 (1.15)	
Concern for Others 2017	3.44 (0.55)		3.65 (0.51)		3.13 (0.56)		3.00 (0.66)		2.59 (1.00)	
Leadership 2018	3.38 (0.54)		3.57 (0.49)		3.16 (0.54)		3.15 (0.62)		3.14 (0.65)	
Conflict Resolution 2018	2.79 (0.64)		2.72 (0.65)		2.64 (0.57)		2.51 (0.59)		2.48 (0.51)	
Teamwork 2018	3.56 (0.57)		3.65 (0.62)		3.37 (0.61)		3.23 (0.70)		3.28 (0.66)	
Concern for Others 2018	3.41 (0.57)		3.53 (0.65)		3.19 (0.56)		3.12 (0.64)		3.16 (0.65)	

Note. N = 45,816; Freq. = Frequency; ES = Emotional Safety.

3.2.2 Cross-Sectional Analyses

As shown in Figure 2.1, the cross-sectional associations in 2017 among the program quality profiles and the three main outcome variables (i.e., leadership, conflict resolution, teamwork) were examined in the same model. A second model (Figure 2.2) examined one outcome variable (Concern for Others) separately since not all members were administered the Concern for Others module. As noted above, only statistically significant results with at least a small effect size will be discussed.

3.2.2.1 Cross-Sectional Analyses for Youth (9-12 Years Old)

See Table 3.10 and Figures 3.3 and 3.4 for results of the cross-sectional models for youth. The pattern of results suggests that youth in profiles with more positive ratings for the program quality indicators often report higher scores for social competence outcome variables. Specifically, youth in the Above Average profile and the Low Belonging and Emotional Safety profile had significantly higher scores for all outcome variables than members in the Below Average profile, Very Below Average profile, and Negative profile (except the difference in Conflict Resolution scores between youth in the Low Belonging and Emotional Safety profile and the Below Average profile did not reach the cut-off for a small effect size). Members in the Below Average profile had higher scores for all outcome variables than members in the Very Below Average profile and the Negative profile. Members in the Very Below Average profile also reported higher scores for all outcome variables than members in the Negative profile (except the difference in Conflict Resolution did not reach the cut-off for a small effect size). Finally, one result did not follow the overall pattern: youth in the Low Belonging and Emotional Safety profile reported higher Leadership scores than youth in the Above Average profile. See

Appendix C for additional model estimates (e.g., associations between covariates and outcome variables).

Table 3.10 Results of the Cross-Sectional Analyses in 2017 for Youth

	(1) Above Average Profile	(2) Low Belonging and ES Profile	(3) Below Average Profile	(4) Very Below Average Profile	(5) Negative Profile	Significant Profile Differences
Model 1						
Leadership	3.49 (0.02)	3.60 (0.03)	3.19 (0.02)	3.04 (0.02)	2.81 (0.03)	All Comparisons
Conflict Resolution	4.21 (0.03)	4.10 (0.04)	3.98 (0.03)	3.84 (0.03)	3.73 (0.03)	All Comparisons
Teamwork	3.79 (0.02)	3.85 (0.03)	3.51 (0.02)	3.33 (0.02)	2.99 (0.03)	All Comparisons Except 1 vs. 2
n	84,488	1,097	22,825	10,046	2,187	
Model 2						
Concern For Others	3.65 (0.03)	3.73 (0.04)	3.37 (0.03)	3.22 (0.03)	3.03 (0.04)	All Comparisons
n	43,507	633	11,201	4,792	1,105	

Note: Intercepts of the outcome variables (and standard errors in parentheses) are presented. For Model 1, significant differences are at $p < .05$ after Bonferroni correction (.05 / 30 = .002) and N = 120,643; For Model 2, significant differences are at $p < .05$ after Bonferroni correction (.05 / 10 = .01) and N = 61,238; ES = Emotional Safety.

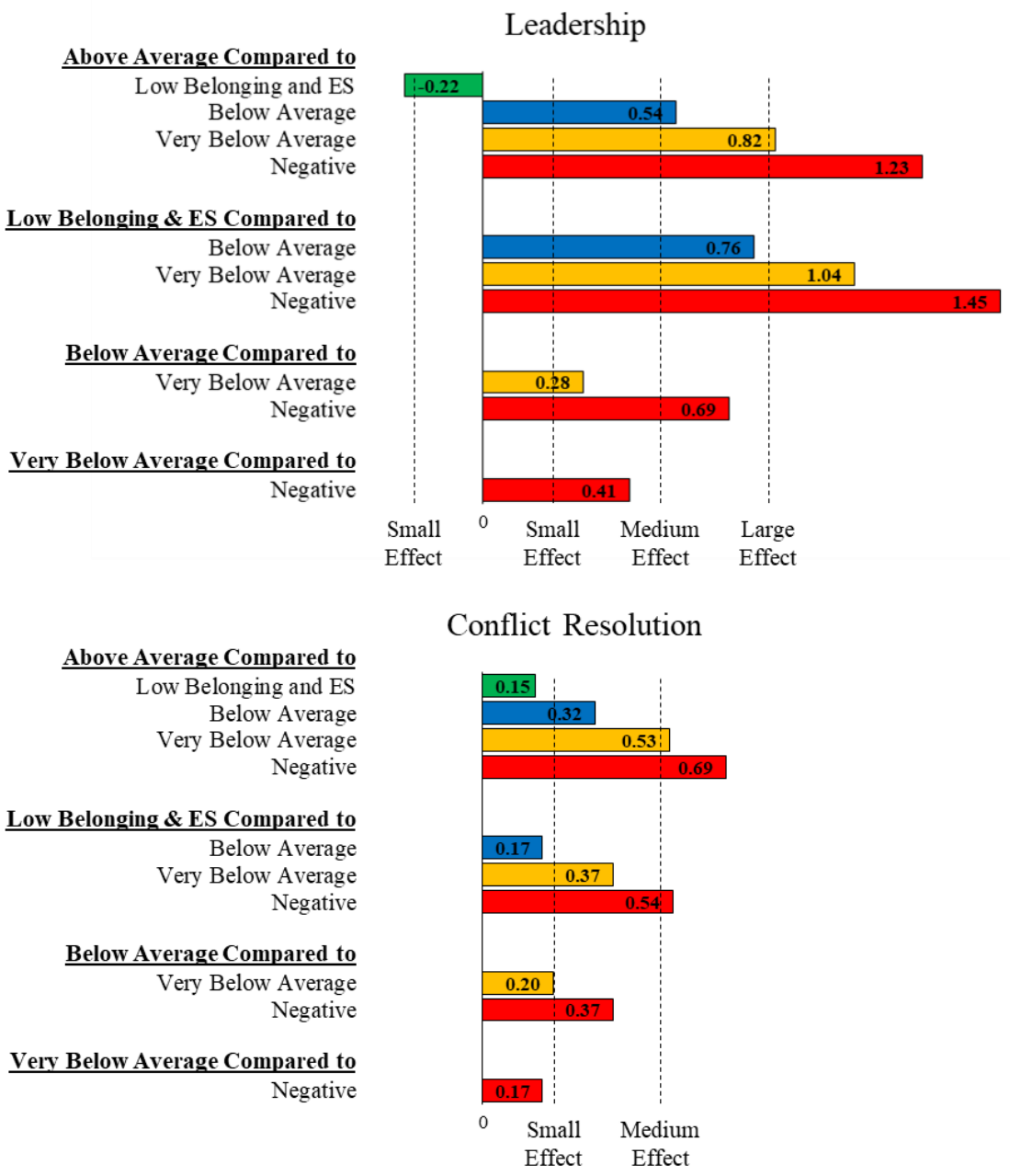

Figure 3.3 Effect sizes for profile comparisons in 2017 cross-sectional analyses for youth, part 1. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49; medium effect: 0.50 – 0.79; large effect: 0.80 and above). All comparisons were significant at $p < .05$ after Bonferroni correction. ES = Emotional Safety.

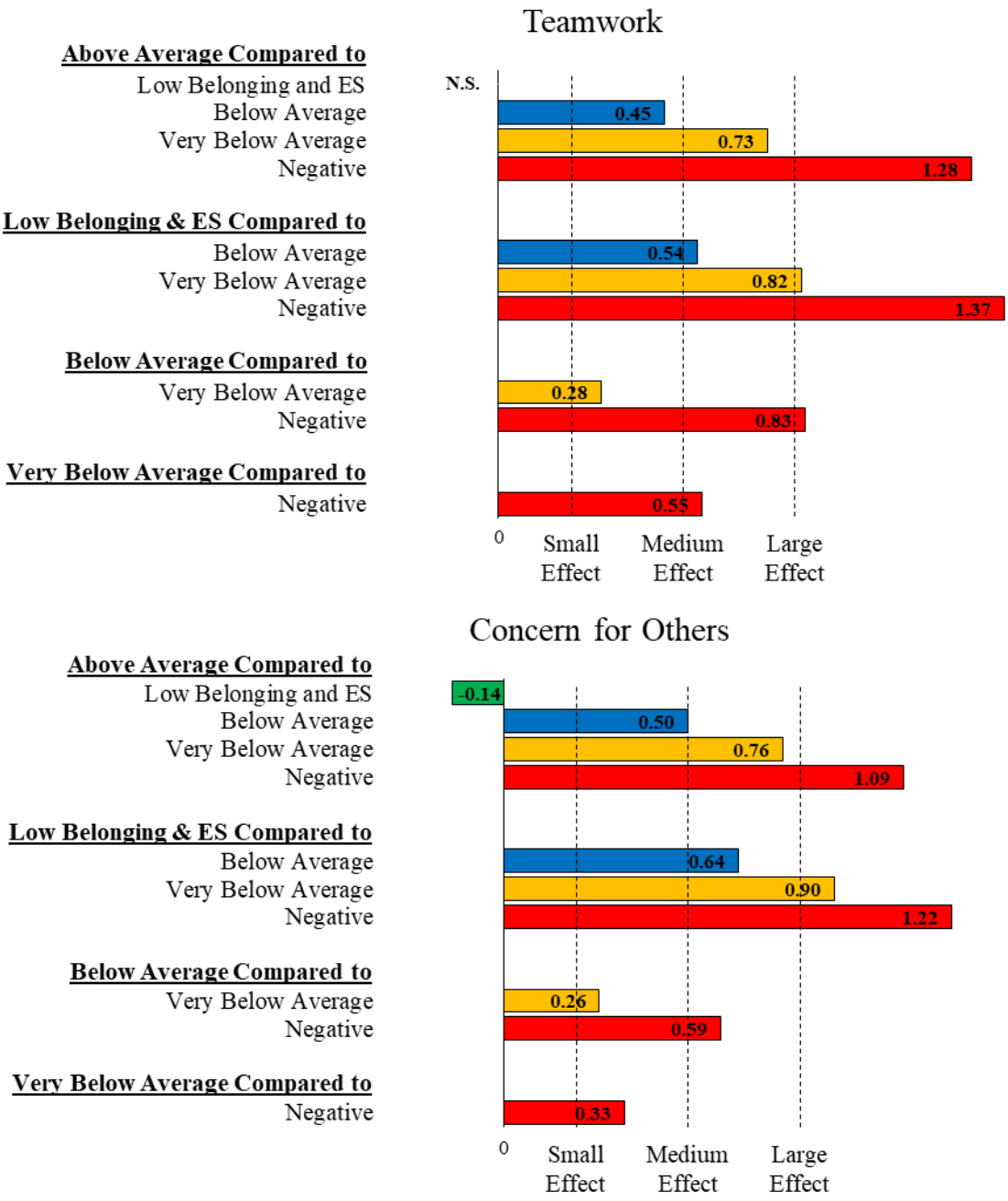

Figure 3.4 Effect sizes for profile comparisons in 2017 cross-sectional analyses for youth, part 2. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49; medium effect: 0.50 – 0.79; large effect: 0.80 and above). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

3.2.2.1 Cross-Sectional Analyses for Teens (13-18 Years Old)

The same cross-sectional models that were examined for youth were also examined for teens (see Figures 2.1 and 2.2). Similar to youth, the pattern of results suggest that teens in profiles with more positive ratings for the program quality indicators often report higher scores for the social competence outcome variables (see Table 3.11 and Figures 3.5 and 3.6). Specifically, youth in the Above Average profile and the Low Belonging and Emotional Safety profile had significantly higher scores for all outcome variables than members in the Below Average profile, Very Below Average profile, and Negative profile. Members in the Below Average profile had higher scores for all outcome variables than members in the Very Below Average profile and the Negative profile (except for Conflict Resolution the difference with the Very Below Average profile did not reach the cut-off for a small effect size and the difference with the Negative profile was not significant). Members in the Very Below Average profile also reported higher scores for all outcome variables than members in the Negative profile (except the difference for Conflict Resolution was not significant). Finally, three results did not follow the pattern: teens in the Low Emotional Safety and Belonging profile reported higher Leadership, Teamwork, and Concern for Others than the Above Average profile. See Appendix C for additional model estimates (e.g., associations between control variables and outcome variables).

Table 3.11 Results of the Cross-Sectional Analysis in 2017 for Teens

Variable	(1) Above Average Profile	(2) Low Belonging and ES Profile	(3) Below Average Profile	(4) Very Below Average Profile	(5) Negative Profile	Significant Profile Differences
Model 1						
Leadership	3.16 (0.03)	3.37 (0.04)	2.83 (0.03)	2.66 (0.04)	2.36 (0.05)	All Comparisons
Conflict Resolution	2.36 (0.05)	2.30 (0.05)	2.15 (0.05)	2.07 (0.05)	2.06 (0.06)	All Comparisons Except 1 vs. 2, 3 vs. 5, 4 vs. 5
Teamwork	3.16 (0.03)	3.28 (0.04)	2.82 (0.03)	2.59 (0.04)	2.13 (0.07)	All Comparisons
n	31,710	819	11,271	1,576	440	
Model 2						
Concern For Others	3.16 (0.05)	3.36 (0.06)	2.86 (0.06)	2.73 (0.06)	2.31 (0.10)	All Comparisons
n	11,534	282	4,191	580	180	

Note: Intercepts of the outcome variables (and standard errors in parentheses) are presented. For Model 1, significant differences are at $p < .05$ after Bonferroni correction (.05 / 30 = .002) and $N = 45{,}816$; For Model 2, significant differences are at $p < .05$ after Bonferroni correction (.05 / 10 = .01) and $N = 16{,}767$; ES = Emotional Safety.

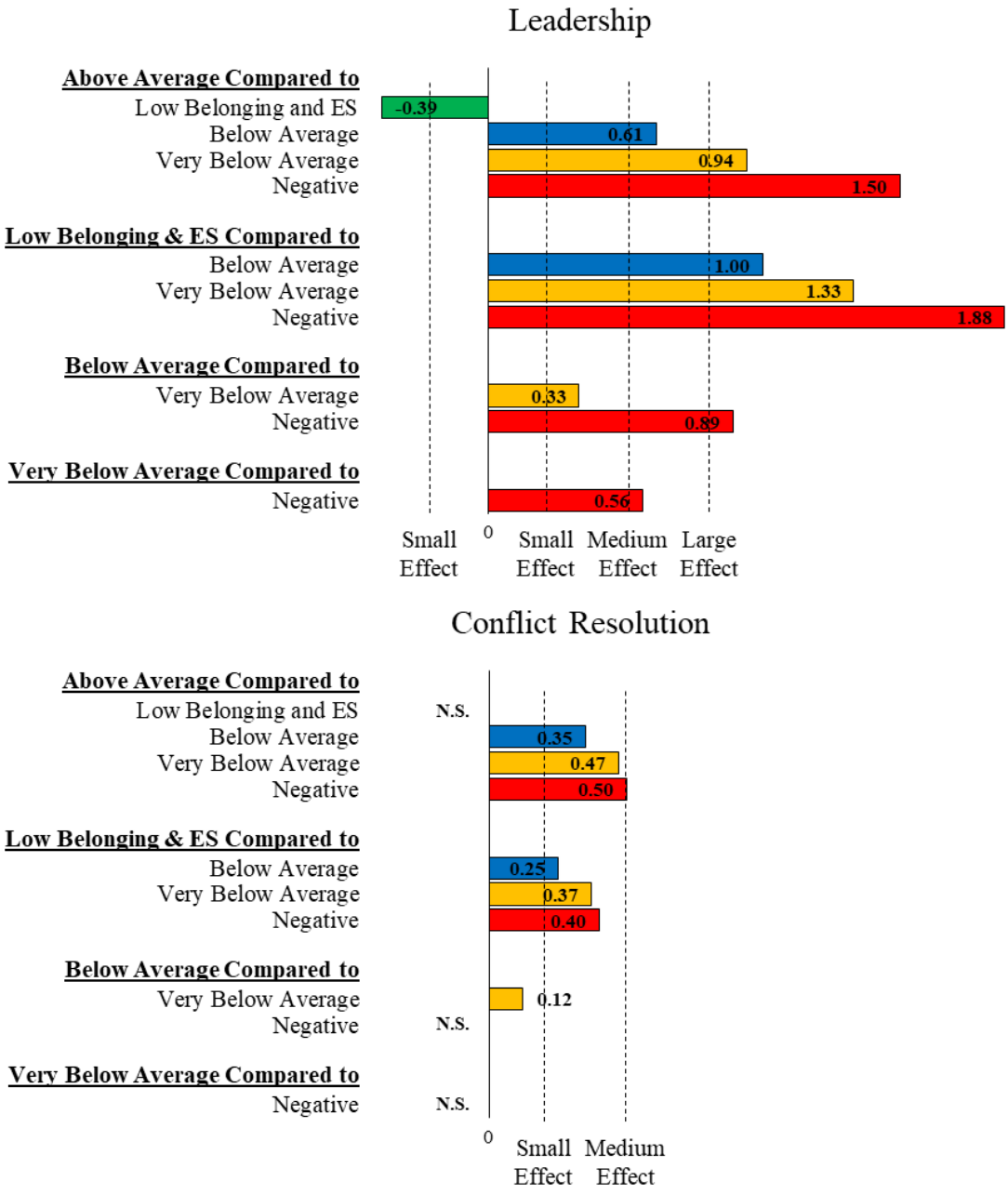

Figure 3.5 Effect sizes for profile comparisons in 2017 cross-sectional analyses for teens, part 1. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49; medium effect: 0.50 – 0.79; large effect: 0.80 and above). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

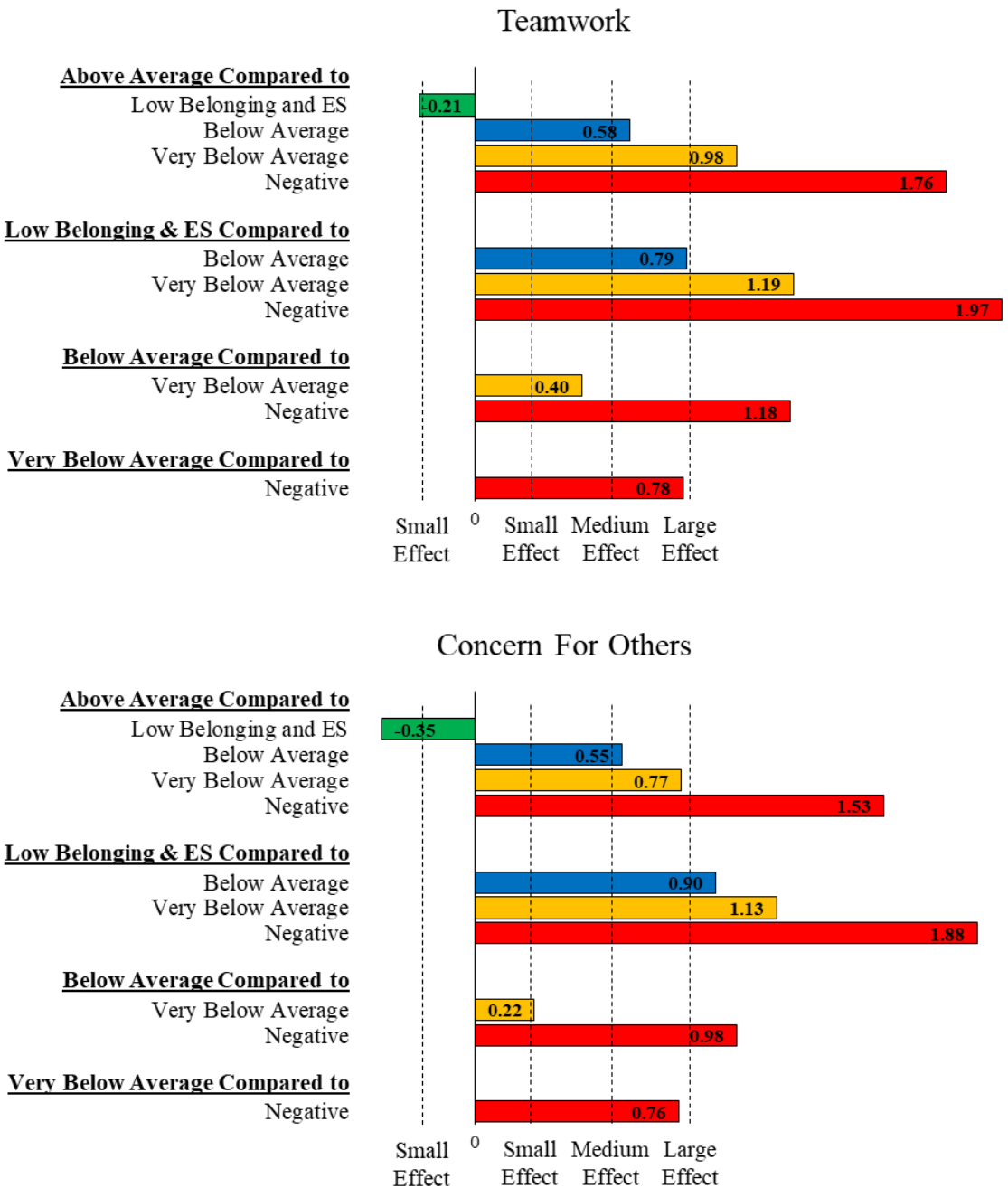

Figure 3.6 Effect sizes for profile comparisons in 2017 cross-sectional analyses for teens, part 2. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49; medium effect: 0.50 – 0.79; large effect: 0.80 and above). All comparisons were significant at $p < .05$ after Bonferroni correction. ES = Emotional Safety.

3.2.3 Associations among Covariates and Program Quality Profiles

The associations among covariates and program quality profiles are presented based on the cross-sectional analyses that incorporated the full sample of youth and teens (see Figure 2.1). Tables 3.12 and 3.13 present the results of these analyses for youth and teens, respectively. For youth, members in the Low Belonging and Emotional Safety Profile were 1.77, 1.86, and 1.64 times more likely than members in the Below Average, Very Below Average, and Negative profiles to receive free/reduced-price lunch, respectively. The significant effects of all other covariates for youth and teens did not reach the cut-off for a small effect size.

Table 3.12 Covariate Analysis Results based on the 2017 Cross-Sectional Model for Youth

	Low Belonging and ES Profile		Below Average Profile		Very Below Average Profile		Negative Profile	
	Est. (S.E.)	OR	Est. (S.E.)	OR	Est. (S.E.)	OR	Est.(S.E.)	OR
Ref.: Above Average								
Age	0.15 (0.03)*	1.16	0.13 (0.01)*	1.14	0.08 (0.01)*	1.08	0.16 (0.02)*	1.18
Male	0.17 (0.06)	1.19	0.09 (0.02)*	1.09	0.19 (0.02)*	1.21	0.22 (0.05)*	1.24
Attendance†								
1 time per week	0.15 (0.14)	1.16	0.01 (0.04)	1.01	-0.02 (0.06)	0.98	-0.05 (0.11)	0.95
2 or more times per week	0.20 (0.12)	1.22	-0.06 (0.03)	0.94	-0.04 (0.05)	0.96	-0.06 (0.09)	0.94
Years Membership	0.00 (0.02)	1.00	0.03 (0.01)*	1.03	0.03 (0.01)	1.03	0.03 (0.02)	1.03
Free/Reduced-Price Lunch	0.34 (0.09)*	1.41	-0.23 (0.02)*	0.79	-0.28 (0.04)*	0.76	-0.15 (0.06)	0.86
Ref.: Negative								
Age	-0.02 (0.04)	0.98	-0.03 (0.02)	0.97	-0.08 (0.02)*	0.92		
Male	-0.05 (0.08)	0.95	-0.13 (0.05)	0.88	-0.02 (0.05)	0.98		
Attendance†								
1 time per week	0.20 (0.17)	1.22	0.06 (0.10)	1.06	0.03 (0.11)	1.03		
2 or more times per week	0.26 (0.14)	1.30	0.00 (0.09)	1.00	0.03 (0.09)	1.03		
Years Membership	-0.03 (0.03)	0.97	0.00 (0.02)	1.00	0.00 (0.02)	1.00		
Free/Reduced-Price Lunch	0.50 (0.11)*	1.64	-0.08 (0.06)	0.92	-0.13 (0.06)	0.88		
Ref.: Very Below Average								
Age	0.07 (0.03)	1.07	0.05 (0.01)*	1.05				
Male	-0.02 (0.06)	0.98	-0.11 (0.03)*	0.90				
Attendance†								
1 time per week	0.17 (0.15)	1.18	0.03 (0.06)	1.03				
2 or more times per week	0.23 (0.12)	1.26	-0.03 (0.05)	0.97				
Years Membership	-0.03 (0.02)	0.97	0.01 (0.01)	1.01				
Free/Reduced-Price Lunch	0.62 (0.09)*	1.86	0.05 (0.03)	1.05				
Ref.: Below Average								
Age	0.02 (0.03)	1.02						
Male	0.08 (0.06)	1.09						
Attendance†								
1 time per week	0.14 (0.15)	1.15						
2 or more times per week	0.26 (0.12)	1.30						
Years Membership	-0.03 (0.02)	0.97						
Free/Reduced-Price Lunch	0.57 (0.09)*	1.77						

Note. Results are based on multinomial logistic regressions using the BCH method for latent profile analysis with distal outcomes. N = 120,643. ES = Emotional Safety; OR = Odds Ratio; Est. = Estimate; Ref. = Reference Profile. †Reference class for attendance is members that attend less than 1x/week. * $p < .05$ after Bonferroni correction (.05 / 60 = .001).

Table 3.13 Covariate Analysis Results based on the 2017 Cross-Sectional Model for Teens

	Low Belonging and ES Profile		Below Average Profile		Very Below Average Profile		Negative Profile	
	Est. (S.E.)	OR	Est. (S.E.)	OR	Est. (S.E.)	OR	Est. (S.E.)	OR
Ref.: Above Average								
Age	0.08 (0.02)	1.08	-0.12 (0.01)*	0.89	-0.26 (0.02)*	0.77	-0.25 (0.04)*	0.78
Male	0.31 (0.07)*	1.36	0.12 (0.03)*	1.13	0.18 (0.06)	1.19	0.22 (0.10)	1.24
Attendance†								
1 time per week	0.00 (0.13)	1.00	-0.03 (0.04)	0.97	0.05 (0.10)	1.05	0.12 (0.19)	1.13
2 or more times per week	-0.05 (0.11)	0.95	-0.10 (0.04)	0.91	-0.18 (0.09)	0.84	0.08 (0.16)	1.08
Years Membership	0.01 (0.02)	1.01	0.00 (0.01)	1.00	-0.01 (0.01)	1.00	-0.01 (0.02)	0.99
Free/Reduced-Price Lunch	0.29 (0.12)	1.33	-0.08 (0.04)	0.92	-0.13 (0.08)	0.88	-0.26 (0.14)	0.77
Ref.: Negative								
Age	0.33 (0.04)*	1.39	0.14 (0.04)*	1.15	-0.01 (0.04)	0.99		
Male	0.09 (0.12)	1.09	-0.09 (0.10)	0.91	-0.04 (0.11)	0.96		
Attendance†								
1 time per week	-0.12 (0.22)	0.89	-0.15 (0.19)	0.86	-0.07 (0.21)	0.93		
2 or more times per week	-0.13 (0.20)	0.88	-0.18 (0.16)	0.84	-0.26 (0.18)	0.77		
Years Membership	0.02 (0.03)	1.02	0.01 (0.02)	1.01	0.01 (0.03)	1.01		
Free/Reduced-Price Lunch	0.54 (0.18)	1.72	0.18 (0.14)	1.19	0.13 (0.15)	1.14		
Ref.: Very Below Average								
Age	0.34 (0.03)*	1.40	0.15 (0.02)*	1.16				
Male	0.13 (0.09)	1.14	-0.05 (0.06)	0.95				
Attendance†								
1 time per week	-0.05 (0.15)	0.95	-0.08 (0.10)	0.92				
2 or more times per week	0.13 (0.13)	1.14	0.08 (0.09)	1.08				
Years Membership	0.02 (0.02)	1.02	0.00 (0.01)	1.00				
Free/Reduced-Price Lunch	0.41 (0.14)	1.51	0.05 (0.08)	1.05				
Ref.: Below Average								
Age	0.19 (0.03)*	1.21						
Male	0.18 (0.07)	1.20						
Attendance†								
1 time per week	0.03 (0.13)	1.03						
2 or more times per week	0.05 (0.11)	1.05						
Years Membership	0.02 (0.02)	1.02						
Free/Reduced-Price Lunch	0.37 (0.13)	1.44						

Note. Results are based on multinomial logistic regressions using the BCH method for latent profile analysis with distal outcomes. N = 45,816. ES = Emotional Safety; OR = Odds Ratio; Est. = Estimate; Ref. = Reference Profile. †Reference class for attendance is members that attend less than 1x/week. * $p < .05$ after Bonferroni correction (.05 / 60 = .001).

3.2.4 Longitudinal Analyses

Before conducting the longitudinal analyses, the association between program quality profiles and 2018 survey completion was examined for youth and teens separately using chi-square tests. The association between program quality profiles and 2018 survey completion was significant at $p < .05$ for youth, $\chi^2(4, 120{,}643) = 163.93$, $p < .001$, V = 0.04, and teens, $\chi^2(4, 45{,}816) = 49.98$, $p < .001$, V = 0.03. These comparisons did not meet the minimum criteria for a small effect size.

Figure 2.3 presents the longitudinal model examining the associations among program quality profiles and the three main outcome variables (i.e., Leadership, Conflict Resolution, Teamwork). A second model (Figure 2.4) examined one outcome variable (Concern for Others) separately since not all members were administered the Concern for Others module. These models were run with sample weights to approximate the entire 2017 cross-sectional sample. The models were also run without sample weights to determine whether the missing data influenced the results. In general, the results were similar for youth and teens when sample weights were and were not used, although the results with weights were more conservative (i.e., the results with sample weights included fewer significant findings). The results for the models with sample weights are presented in this section (and the results for models without sample weights are presented in Appendix D), and key differences between the models are noted in the text.

3.2.4.1 Longitudinal Analyses for Youth with Sample Weights

See Table 3.14 and Figures 3.7 and 3.8 for the results of the longitudinal analyses for youth. Controlling for baseline levels in 2017, youth in the Above Average profile reported higher 2018 levels of Leadership and Teamwork compared to the Negative profile, and higher 2018 levels of Teamwork compared to the Very Below Average profile. In addition, youth in the

Low Belonging and Emotional Safety profile reported higher levels of Leadership and Teamwork in 2018 compared to youth in the Below Average, Very Below Average, and Negative profiles when controlling for baseline levels. The model without sample weights included these same results, as well as the following significant findings that met criteria for a small effect size: 1) Youth in the Above Average profile reported higher 2018 levels (controlling for 2017 baseline levels) of Leadership and Concern for Others compared to youth in the Below Average profile, and these youth also reported higher 2018 levels of Concern for Others than youth in the Very Below Average profile; and 2) youth in the Low Belonging and Emotional Safety profile reported higher 2018 levels of Concern for Others than youth in the Below Average and Very Below Average profiles when controlling for baseline levels. See Appendix E for additional model estimates (e.g., associations between control variables and outcome variables).

Table 3.14 Results of the Longitudinal Analysis with Sample Weights for Youth

Variable	(1) Above Average Profile	(2) Low Belonging and ES Profile	(3) Below Average Profile	(4) Very Below Average Profile	(5) Negative Profile	Significant Profile Differences
Model 1						
Leadership	2.37 (0.04)	2.41 (0.06)	2.27 (0.04)	2.27 (0.04)	2.23 (0.05)	1 > 3,4,5 2 > 3,4,5
Conflict Resolution	2.00 (0.05)	2.01 (0.06)	1.98 (0.05)	1.97 (0.05)	1.95 (0.06)	None
Teamwork	2.73 (0.05)	2.78 (0.06)	2.65 (0.05)	2.59 (0.05)	2.60 (0.06)	1 > 3,4,5 2 > 3,4,5 3 > 4
n	25,029	343	6,806	3,067	672	
Model 2						
Concern For Others	2.48 (0.07)	2.53 (0.09)	2.40 (0.07)	2.39 (0.07)	2.35 (0.10)	1 > 3,4
n	10,258	141	2,655	1,077	237	

Note: Intercepts of the outcome variables (and standard errors in parentheses) are presented. For Model 1, significant differences are at $p < .05$ after Bonferroni correction (.05 / 30 = .002) and $N = 35{,}918$; For Model 2, significant differences are at $p < .05$ after Bonferroni correction (.05 / 10 = .01) and $N = 14{,}367$; ES = Emotional Safety.

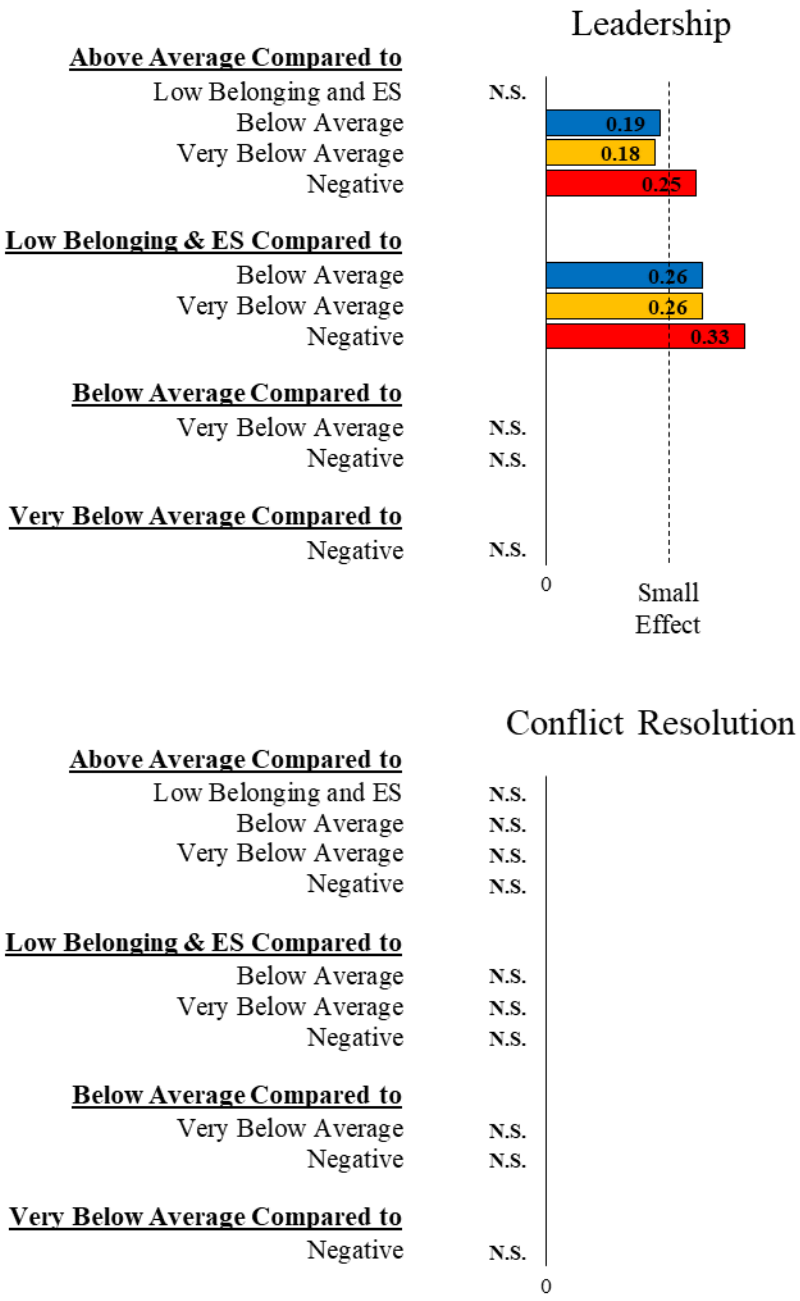

Figure 3.7 Effect sizes for profile comparisons in longitudinal analyses with sample weights for youth, part 1. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

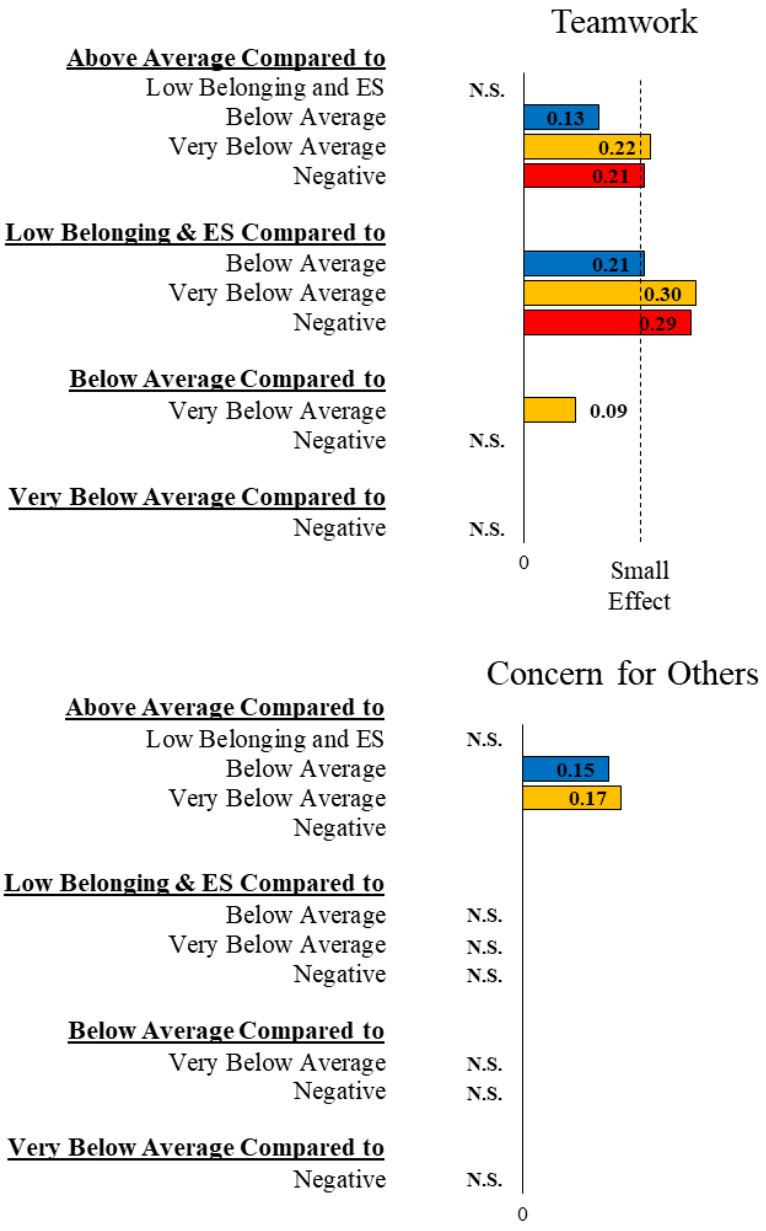

Figure 3.8 Effect sizes for profile comparisons in longitudinal analyses with sample weights for youth, part 2. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

3.2.4.2 Longitudinal Analyses for Teens with Sample Weights

See Table 3.15 and Figures 3.9 and 3.10 for the results of the longitudinal analyses for teens. Controlling for 2017 baseline levels, teens in the Above Average profile reported higher 2018 levels of Conflict Resolution and Teamwork compared to teens in the Very Below Average profile. Teens in the Low Belonging and Emotional Safety profile also reported higher 2018 levels of Leadership compared to the Below Average profile, and higher 2018 levels of Teamwork compared to the Very Below Average profile when controlling for baseline levels in 2017. The model without sample weights included these same results, as well as the following significant findings that met criteria for a small effect size: Teens in the Above Average profile reported higher 2018 levels of each outcome variable (controlling for baseline levels in 2017) compared to teens in the Below Average profile. See Appendix E for additional model estimates (e.g., associations between control variables and outcome variables).

Table 3.15 Results of the Longitudinal Analysis with Sample Weights for Teens

Variable	(1) Above Average Profile	(2) Low Belonging and ES Profile	(3) Below Average Profile	(4) Very Below Average Profile	(5) Negative Profile	Significant Profile Differences
Model 1						
Leadership	1.81 (0.08)	1.92 (0.09)	1.72 (0.08)	1.76 (0.08)	1.83 (0.14)	1 > 3 2 > 3
Conflict Resolution	0.94 (0.08)	0.92 (0.09)	0.90 (0.08)	0.82 (0.09)	0.87 (0.11)	1 > 4
Teamwork	2.39 (0.09)	2.43 (0.10)	2.29 (0.09)	2.20 (0.10)	2.33 (0.14)	1 > 3,4 2 > 4
n	6,618	166	2,334	330	77	
Model 2						
Concern For Others	1.42 (0.16)	1.42 (0.18)	1.33 (0.15)	1.26 (0.18)	1.41 (0.21)	1 > 3
n	1,808	44	551	86	21	

Note: Intercepts of the outcome variables (and standard errors in parentheses) are presented. For Model 1, significant differences are at $p < .05$ after Bonferroni correction (.05 / 30 = .003) and $N = 9{,}525$; For Model 2, significant differences are at $p < .05$ after Bonferroni correction (.05 / 10 = .01) and $N = 2{,}510$; ES = Emotional Safety.

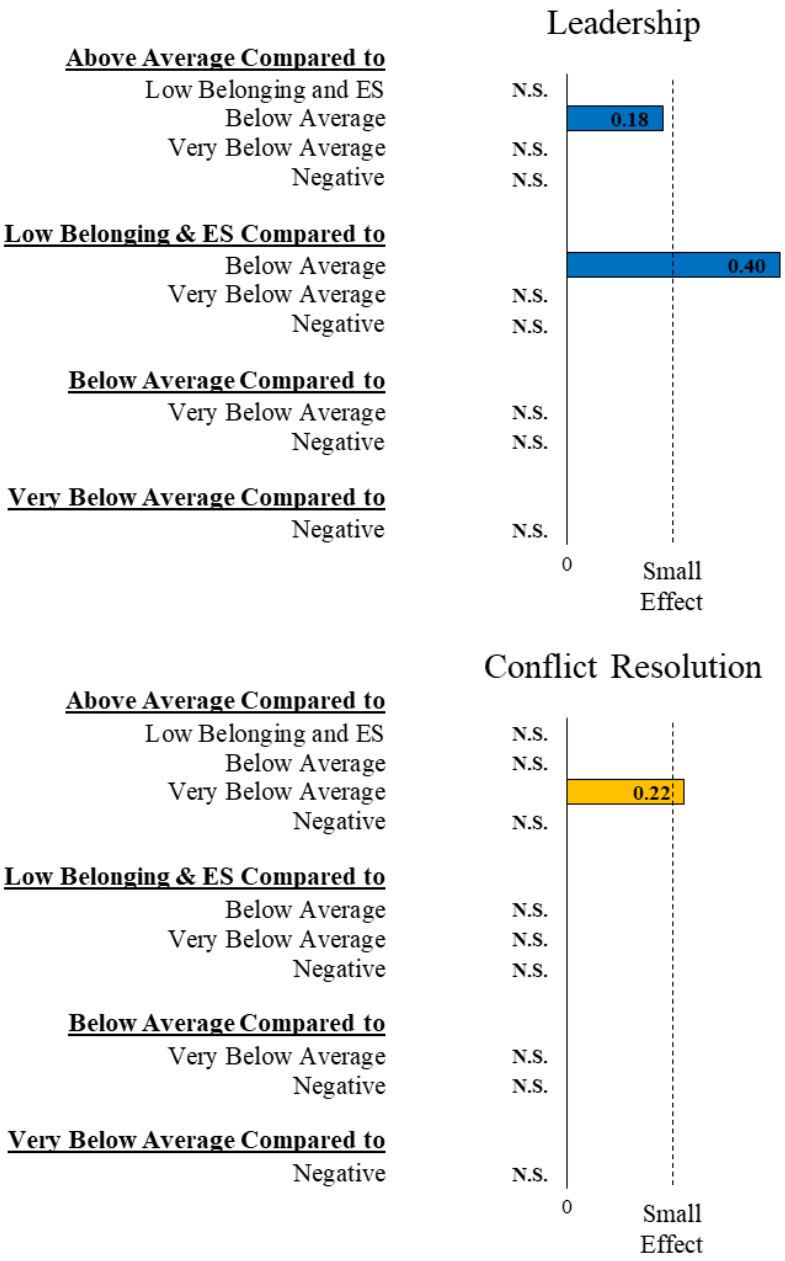

Figure 3.9 Effect sizes for profile comparisons in longitudinal analyses with sample weights for teens, part 1. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

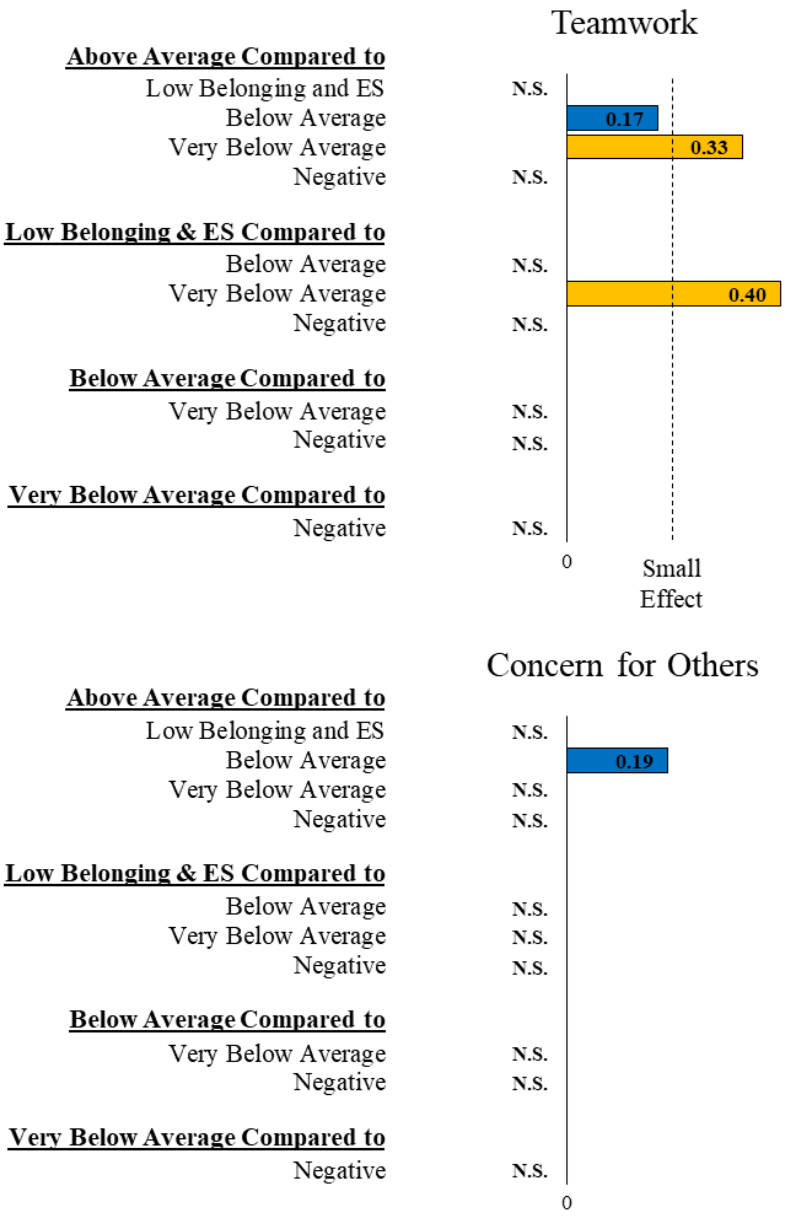

Figure 3.10 Effect sizes for profile comparisons in longitudinal analyses with sample weights for teens, part 2. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

3.2.5 Site-Level Characteristics

Chi-square tests were used to examine potential differences in the proportion of profiles across community location (urban, suburban, and rural), facility type (traditional, school, public housing, Native American, and other type), and Club size (average daily attendance of 0-49, 50-174, and 175 or larger; see Tables 3.16 and 3.17). For youth, there were significant differences in the proportion of profiles across each site-level characteristic, although these differences did not reach the cut-off for a small effect (community location: $\chi^2(8, N = 119{,}787) = 177.32$, $p < .001$, Cramer's $V = 0.03$; facility type: $\chi^2(16, N = 119{,}806) = 283.85$, $p < .001$, Cramer's $V = 0.02$; average daily attendance: $\chi^2(8, N = 119{,}794) = 757.75$, $p < .001$, Cramer's $V = 0.06$). For teens, there were also significant differences in the proportion of profiles across each site-level characteristic that did not reach the cut-off for a small effect (community location: $\chi^2(8, N = 45{,}477) = 36.42$, $p < .001$, Cramer's $V = 0.03$; facility type: $\chi^2(16, N = 45{,}477) = 103.72$, $p < .001$, Cramer's $V = 0.03$; average daily attendance: $\chi^2(8, N = 45{,}469) = 130.29$, $p < .001$, Cramer's $V = 0.04$).

Table 3.16 Percentages and Adjusted Standardized Residuals for Profiles by Site-Level Characteristics among Youth

	(1) Above Average Profile	(2) Low Belonging and ES Profile	(3) Below Average Profile	(4) Very Below Average Profile	(5) Negative Profile
Community Type					
Rural	21% (4.09)*	20% (-0.44)	19% (-3.37)*	18% (-4.75)*	16% (-4.07)*
Suburban	29% (-5.04)*	27% (-1.96)	33% (6.26)*	33% (4.48)*	34% (2.90)
Urban	50% (1.34)	53% (1.81)	48% (-2.76)	49% (-0.48)	50% (0.33)
Site Type					
Native American	2% (2.46)	2% (0.07)	2% (-1.20)	1% (-4.14)*	1% (-2.61)
Other	3% (2.61)	2% (-1.09)	3% (-1.57)	2% (-4.43)*	2% (-0.86)
Public Housing	6% (5.66)*	9% (4.80)*	5% (-6.33)*	4% (-6.83)*	4% (-3.49)*
School	28% (0.03)	26% (-1.27)	29% (2.48)	27% (-2.05)	25% (-2.89)
Traditional	61% (-2.70)	61% (-0.36)	62% (0.76)	66% (5.10)*	68% (3.64)*
ADA					
0-49	8% (6.77)*	8% (-0.11)	7% (-6.23)*	6% (-7.66)*	5% (-5.46)*
50-174	65% (4.50)*	69% (2.06)	62% (-3.41)*	59% (-5.99)*	54% (-5.60)*
175+	26% (-10.30)*	23% (-3.04)*	31% (8.38)*	35% (12.99)*	41% (11.26)*
n	83,876	1,091	22,673	9,984	2,170

Note. Percentage of youth within each profile are presented, along with adjusted standardized residuals in parentheses; The proportion of profiles across each site-level characteristic was significantly different at $p < .05$, and the effect sizes did not meet criteria for a small effect; Positive standardized residuals indicate a larger percentage of youth than expected by chance; Negative standardized residuals indicate a smaller percentage of youth than expected by chance; ES = Emotional Safety; ADA = Average Daily Attendance. *$p < .05$ after Bonferroni correction (.05 / 15 = .003 for Community Type and ADA, and .05 / 25 = .002 for Site Type).

Table 3.17 Percentages and Adjusted Standardized Residuals for Profiles by Site-Level Characteristics among Teens

	(1) Above Average Profile	(2) Low Belonging and ES Profile	(3) Below Average Profile	(4) Very Below Average Profile	(5) Negative Profile
Community Type					
Rural	16% (2.01)	12% (-2.83)	15% (-1.93)	14% (-1.38)	14% (-0.84)
Suburban	25% (-1.26)	23% (-1.02)	26% (1.18)	29% (2.53)	29% (1.35)
Urban	59% (-0.21)	65% (2.13)	59% (0.22)	57% (-0.95)	57% (-0.45)
Site Type					
Native American	2% (-0.51)	1% (-1.74)	2% (0.96)	3% (1.91)	1% (-1.73)
Other	3% (1.16)	3% (0.45)	3% (-1.98)	3% (0.05)	3% (-0.51)
Public Housing	7% (3.63)*	7% (-0.43)	6% (-5.09)*	6% (-1.14)	4% (-2.23)
School	22% (-3.33)*	23% (-0.21)	26% (5.12)*	24% (0.82)	26% (1.02)
Traditional	65% (0.65)	66% (0.48)	64% (-1.16)	64% (-0.47)	67% (0.54)
ADA					
0-49	13% (3.89)*	9% (-2.59)	10% (-4.41)*	10% (-2.44)	8% (-2.49)
50-174	61% (1.20)	64% (1.20)	59% (-1.64)	58% (-1.40)	57% (-0.88)
175+	27% (-4.31)*	28% (-0.07)	30% (5.30)*	33% (3.66)*	35% (2.92)
n	31,461	813	11,197	1,560	438

Note. Percentage of teens within each profile are presented, along with adjusted standardized residuals in parentheses; The proportion of profiles across each site-level characteristic was significantly different at $p < .05$, and the effect sizes did not meet criteria for a small effect; Positive standardized residuals indicate a larger percentage of teens than expected by chance; Negative standardized residuals indicate a smaller percentage of teens than expected by chance; ES = Emotional Safety; ADA = Average Daily Attendance. *$p < .05$ after Bonferroni correction (.05 / 15 = .003 for Community Type and ADA, and .05 / 25 = .002 for Site Type).

4 DISCUSSION

Researchers have documented a positive association between youth program quality and positive youth outcomes, including social competence (e.g., Chapman et al., 2017; Durlak et al., 2010; Dworkin et al., 2003; Wright et al., 2006; Yohalem & Wilson-Ahlstrom, 2010). Most of this research has utilized a variable-centered approach to examine how individual aspects of program quality, such as supportive relationships and youth engagement, are associated with developmental outcomes (Denault & Poulin, 2016). The primary aims of this study were to 1) utilize a person-centered approach to identify profiles of program participants based on their perceptions of multiple indicators of program quality; 2) examine cross-sectional and longitudinal associations among profiles of program quality and social competence outcomes; and 3) determine whether profiles of program quality were associated with site-level characteristics (e.g., average daily attendance, community location).

The first aim of this study was largely exploratory given the lack of research into how young people experience multiple aspects of program quality. For both youth (9-12 years old) and teens (13-18 years old), five profiles of program quality were identified: Above Average, Low Belonging and Emotional Safety, Below Average, Very Below Average, and Negative. For the second aim of this study, the hypothesis was based on the social development model which suggests that youth who develop a positive bond to an important institution, such as a YDP, will be more likely to adopt its social norms (Catalano & Hawkins, 1996; Hawkins & Weis, 1985). Thus, it was hypothesized that young people would only report high levels of social competence if they had profiles with high ratings for the social bond-related indicators of program quality (i.e., positive adult connections, sense of belonging, fun/engagement). Furthermore, it was hypothesized that youth profiles with high ratings for social bond-related indicators of program

quality *and* social norms-related indicators of program quality (i.e., physical and emotional safety, staff expectations, staff recognition) would be associated with even higher levels of social competence. These hypotheses were partially supported. Specifically, profiles with higher ratings for social-bond and social-norms related indicators of program quality tended to be associated with higher levels of social competence. However, there was one profile of participants that endorsed high ratings for all components of program quality except for Belonging (one of the social bond-related indicators of program quality) and Emotional Safety (one of the social norms-related indicators of program quality), and these participants still reported high levels of social competence.

The hypothesis for the last aim of this study was also partially supported. As hypothesized, there was a higher proportion of youth (9-12 years old) and teens (13-18 years old) that were categorized in the Above Average profile at sites that serve the smallest number of participants (up to 49 per day on average). Although statistically significant in the direction that was hypothesized, the effect size did not reach criteria for practical significance. Hypotheses were not specified for the other site-level characteristics (i.e., community location, facility type) given the limited evidence related to these site-level characteristics and youth perceptions of program quality.

4.1 Profiles of Program Quality

The majority of youth (9-12 years old) and teens (13-18 years old) were classified in the "Above Average" profile (70.0% of youth and 69.2% of teens), indicating that the majority of participants have positive experiences across all aspects of program quality examined in this study. This is consistent with one previous study that examined profiles of young people's experiences within organized extracurricular activities. Denault and Poulin (2016) examined

experiences related to program quality (such as social integration and support from activity leaders) and experiences related to youth outcomes (such as the development of emotion regulation and social skills). These researchers found that 60.9% of 412 participants were classified in a "High" category (with all ratings about 0.5 standard deviations above average). In Denault and Poulin's study and the current study, around two-thirds of participants provided above average ratings for all indicators of the latent profiles. These findings provide evidence that the majority of youth program participants have positive experiences across multiple domains. This is promising given that program quality and positive youth program experiences are associated with positive outcomes (Roth & Brooks-Gunn, 2016; Smith et al., 2017; Yohalem & Wilson-Ahlstrom, 2010).

Three of the other profiles included below average ratings for all aspects of program quality. However, since the average score for each indicator of program quality was quite high, for the Below Average and Very Below Average profiles the ratings for most of the program quality indicators still reflect positive-to-neutral perceptions of program quality. In contrast, the "Negative" profile included negative perceptions of each program quality indicator. It is noteworthy that a very small percentage of teens were classified into the Negative profile (about 1.0% of teens). The low percentage of teens classified in the Negative profile is consistent with previous research documenting a positive association between age and program quality ratings (e.g., Kuperminc et al., 2019; Ramey et al., 2018). One potential explanation for this finding is that teens who have negative experiences may be more likely to stop attending the program because they often have more autonomy and more options for how they spend their time afterschool (Deschenes et al., 2010).

The vast majority of youth and teens were classified into profiles with similar ratings for all indicators of program quality (i.e., Above Average, Below Average, Very Below Average, Negative). This is consistent with the one previous study that examined young people's experiences within organized extracurricular activities. Specifically, Denault and Poulin (2016) found that young people were classified into three categories with similar ratings for multiple experiences (e.g., social integration, support from activity leaders): "High" (with all ratings about 0.5 standard deviations above average), "Average" (with all ratings about 0.5 standard deviations below average), and "Low" (with all ratings from 0.5 to 2.3 standard deviations below average). Although the body of evidence is small, the available data suggest that most young people tend to rate their experiences within youth programs in a similar way.

Participants' similar ratings across multiple aspects of program quality is also consistent with a confirmatory factor analysis that examined five of the seven indicators of program quality included in this study (i.e., fun, sense of belonging, emotional safety, physical safety, adult connections; Kuperminc et al. 2019). Kuperminc and colleagues found that a one-factor model fit the data better than a two-factor model (one factor was composed of emotional safety and physical safety, and a second factor was composed of the three other indicators of program quality). This suggests that the five indicators of program quality may reflect one underlying construct (Kline, 2005): participants' overall experience of program quality. Although other researchers have identified multiple-factor models when examining other program quality items, the factors are often highly correlated. For example, Bean and Forneris (2016) conducted an exploratory factor analysis using 19 items related to program quality and they found 4 factors: adult support and structure, empowered skill-building, expanding horizons, and negative experiences. The magnitude of the correlations among the three positive factors was large.

Taken together, research findings indicate that most participants experience multiple aspects of program quality in a similar way. Conceptually, the different components of program quality reinforce each other (Bean & Forneris, 2016). For example, if participants have a lot of fun at a program, they likely feel a sense of belonging and have positive relationships. In addition, feeling emotionally and physically safe can provide the foundation for participants to relax and have fun. On the other hand, if participants do not feel safe at a program, they may be less likely to have fun, feel a sense of belonging, have positive relationships with staff, etc.

Although the data suggest that youth experience most aspects of program quality in a consistent way, it is also possible that this finding is exaggerated by response bias (Serido et al., 2011). Specifically, young people are more likely (compared to adults) to agree with items without regard to the items' content ("acquiescence"; Navarro-González et al., 2016). Social desirability may also play a role. Both acquiescence and social desirability may contribute to the high percentage of youth that rate all aspects of program quality highly. Other response biases, such as the tendency to disagree with items ("counter- acquiescence"; Rammstedt & Farmer, 2013), could also contribute to profiles in which participants rated all indicators of program quality below average. Future research should examine the role of response bias in program participants' ratings of multiple program quality indicators.

One of the five profiles for both youth (9-12 years old) and teens (13-18 years old) did not include consistent ratings for all indicators of program quality. Specifically, there was one profile with low ratings for Belonging and Emotional Safety, and high ratings for all other program quality components. The percentage of participants in this profile was very small for youth (0.9%) and teens (1.8%), and therefore it is important for future research to replicate this finding in samples large enough for such a small profile to be detected. However, the fact that

this profile emerged in both subsamples and included over 1,000 youth and 800 teens suggests that this profile reflects a real subgroup of program participants.

It is surprising that some participants lack a sense of belonging and do not feel emotionally safe, yet they have close relationships with adults, have fun, feel physically safe, and perceive that adults have high expectations and recognize their positive behaviors. There are a few potential explanations for these participants' experiences. Many of the items that make up the Belonging and Emotional Safety measures may reflect participants' perceptions of the Club as an institution, rather than relationships with individual people within the Club. For example, the Belonging measure asks whether participants "feel like [their] ideas count here" and the Emotional Safety measure asks whether "this Boys & Girls Club has rules for how people are supposed to treat each other." In contrast, many of the other measures ask more directly about staff or adults at the Club, rather than perceptions of the Club as a whole. It is also noteworthy that a larger percentage of teens (1.8%) than youth (0.9%) were classified into the Low Belonging and Emotional Safety category, suggesting that teens may be better able to differentiate these more subtle dimensions of program quality.

Another potential explanation for the existence of the Low Belonging and Emotional Safety profile is based on the covariate analyses. Youth in the Low Belonging and Emotional Safety profile were significantly more likely to receive free/reduced-price lunch than youth in all other profiles (the effect size was not practically significant for the comparison with the Above Average profile), and this same trend was present for teens. Since free/reduced-price lunch is an indicator of low-income status and living in a low-income community increases the risk of trauma exposure (Kiser & Black, 2005), it is possible that participants in the Low Belonging and Emotional Safety profile have experienced higher rates of traumatic stressors. Young people who

have experienced trauma or other difficulties in their lives may struggle to develop a sense of emotional safety and belonging (Corrales et al., 2016; Overstreet & Mathews, 2011; Shonkoff et al., 2012; Villalta et al., 2018). Despite such challenges, these youth may be able to display resilience by forming close relationships, having fun with their peers, etc. (Khanlou & Wray, 2014). Unfortunately, the available data did not permit an assessment of the extent to which participants had experienced prior trauma; thus, such a possibility cannot be further explored in this study. Additional research, including qualitative research and studies with direct data on trauma exposure, is needed to better understand the experiences of these young people.

4.2 Individual- and Site-Level Characteristics Associated with Profiles of Program Quality

Most individual-level characteristics (i.e., age, gender, years of membership, frequency of attendance) did not meaningfully predict classification into profiles of program quality for youth or teens. Indeed, free/reduced-price lunch status was the only individual-level characteristic that was associated with profiles of program quality and had a practically significant effect size (described above).

However, these results do not necessarily indicate that individual-level characteristics are unrelated to perceptions of program quality at all youth programs. Instead, the results suggest that when examining these associations in aggregate across 2,969 Clubs, there are no clear patterns of practical significance. It is possible that individual-level characteristics may influence program quality perceptions in different ways at specific programs. For example, Flett, Gould, and Lauer (2012) observed 45 sports programs and found that boys' sports programs were characterized by a less welcoming and supportive atmosphere compared to girls' sports programs. As this demonstrates, participants at certain Clubs may be treated in different ways

based on individual-level characteristics. However, when examined in aggregate there is not a clear pattern of associations among individual-level characteristics and perceptions of program quality (Kievit et al., 2013). Roth and Brooks-Gunn (2016) highlight the importance of considering the interaction between individual participants and specific program contexts, and future research in this area is important for advancing understanding of the role of individual-level characteristics on perceptions of program quality.

Similarly, none of the effect sizes for the associations among site-level characteristics (i.e., number of young people served, community location, facility type) and profiles of program quality reached criteria for practical significance (i.e., an effect size that is at least small in magnitude). The limited previous research examining the association between average daily attendance and program quality has documented negative correlations that were small in magnitude. For example, Kuperminc and colleagues (2019) found small negative associations among perceptions of program quality and the number of young people served. Overall, the lack of moderate-to-strong associations among site-level characteristics and program quality may be due to the multiple factors that can affect program quality. For example, at programs with a large number of participants, it may be difficult for staff to build close relationships with participants and maintain a structured environment. However, if a Club is very large yet also well-funded with a low participant-to-staff ratio, the potentially negative effects of a large Club may be minimized. In this way, bivariate associations among individual site-level characteristics and profiles of program quality may obscure the combined influence of multiple site-level characteristics on program quality (Kievit et al., 2013). Future research that includes multiple site-level characteristics (e.g., staff-youth ratios, staff training, funding levels) could explore the

role of more nuanced program contexts and individual characteristics on program quality perceptions (Deutsch et al., 2017; Roth & Brooks-Gunn, 2016).

4.3 Cross-Sectional and Longitudinal Associations among Profiles of Program Quality and Social Competence

For the cross-sectional and longitudinal analyses, the overall pattern of results is consistent with previous variable-centered research that has found positive associations among individual components of program quality and developmental outcomes, including social competence (e.g., Kataoka & Vandell, 2013; McDonough et al., 2013; Moore & Hamilton, 2010; Shernoff, 2010). This same pattern was found in the current study: young people whose profiles had more positive ratings for program quality in 2017 tended to report higher levels of social competence in 2017 and higher levels of some aspects of social competence in 2018 (controlling for baseline levels in 2017). Even after accounting for the cross-sectional associations, profiles with more positive program quality ratings were associated with higher levels of some aspects of social competence over the course of a year. These findings suggest that positive YDP experiences can have lasting positive effects on the development of social competence. Along with previous variable-centered research, this study provides support for the growing focus among youth program researchers and practitioners on improving program quality as a strategy for promoting positive youth outcomes (Fredricks et al., 2017; Yohalem & Wilson-Ahlstrom, 2010).

The social development model would predict that profiles with more positive ratings for social bond-related and social norms-related components of program quality would be associated with higher levels of social competence, and the findings in this study are mostly consistent with these predictions. Specifically, profiles with more above average ratings for social bond-related

and social norms-related indicators of program quality were frequently associated with higher levels of social competence. However, there were four notable exceptions. Youth classified into the Low Belonging and Emotional Safety profile reported higher leadership levels than those classified into the Above Average profile, and teens in the Low Belonging and Emotional Safety profile reported higher levels of leadership, teamwork, and concern for others than their peers in the Above Average profile. The percentage of youth and teens in the Low Belonging and Emotional Safety profile was very small, and there may be something unique about these young people that explains their high reports of social competence. As noted above, one potential explanation is that these young people may have difficulty feeling a sense of belonging and feeling emotionally safe due to difficult past experiences, yet they may be particularly resilient in other domains (e.g., the ability to develop personal relationships) (Corrales et al., 2016; Khanlou & Wray, 2014; Oshri et al., 2017; Overstreet & Mathews, 2011; Shonkoff et al., 2012; Villalta et al., 2018). In addition, youth and teens in the Low Belonging and Emotional Safety profile rated other aspects of program quality very high. Thus, it is possible that young people in this profile still felt an overall positive bond to the YDP and were therefore still influenced by the YDP's positive social norms. Follow-up qualitative research with this small, but unique, subset of participants might provide valuable insights into the ways that experiences at YDPs are associated with the development of social competence.

4.4 Additional Implications for Research

The current study is one of the first to examine program quality using a person-centered approach, and additional research can build on this study's findings. Although the current study accounted for clustering at the site level, future research can conduct *multi-level* mixture modeling to examine whether there are profiles of *Clubs/sites* based on quality indicators

(Flunger et al., 2019; Van Eck et al., 2017). For example, there may be Clubs that facilitate fun and a sense of belonging, but struggle with setting high expectations and reinforcing positive behavior. In this way, multi-level mixture modeling can identify Clubs that only facilitate high quality experiences in certain areas, thereby helping practitioners prioritize improvement efforts at specific sites. In addition, multi-level mixture modeling can provide additional insight into how multiple configurations of program quality at the site level are associated with developmental outcomes.

Future research could also examine the bidirectional relations among social competence and program quality perceptions. The longitudinal results of this study suggest that program quality can facilitate the development of social competence. At the same time, social competence may also promote more positive experiences of program quality. Development involves the complex interplay among individuals and their contexts (Cantor et al., 2019; Lerner et al., 2013). If young people with high levels of social competence attend a YDP, they may be better able to navigate the program's social dynamics which can lead to more positive experiences and, in turn, more positive perceptions of program quality (Domitrovich et al., 2017; Milligan et al., 2017). Future studies with multiple wave of data could employ statistical techniques, such as cross-lagged panel analysis, to study these bidirectional associations (e.g., Klein et al., 2018).

Some of the literature examining the link between YDPs and developmental outcomes has emphasized the importance of program activities that focus on skill-building (Deutsch et al., 2017). For example, in a meta-analysis examining whether afterschool programs facilitate personal and social skills, Durlak et al. (2010) found that participants in programs utilizing skill-building activities that are sequenced, active, focused, and explicit (SAFE) experienced gains in all outcomes examined (e.g., positive social behaviors, school grades). However, participants in

programs that did not utilize SAFE skill-building activities did not experience gains in any of the outcomes examined. The findings of the current study suggest that perceptions of setting characteristics (e.g., fun, sense of belonging, positive expectations) can contribute to gains in social competence independent of such programming. Future research can explore how perceptions of program quality interact with intentional skill-building programming to facilitate the development of youth outcomes.

4.5 Implications for Practice

The most critical implication for practice is that participants' experiences of program quality matter for the development of social competence; these findings highlight the importance of youth program staff working to improve program quality (e.g., Chapman et al., 2017; Durlak et al., 2010; Dworkin et al., 2003; Wright et al., 2006; Yohalem & Wilson-Ahlstrom, 2010). One strategy for improving program quality is the use of continuous quality improvement initiatives that are being implemented by some YDPs, including BGCA (e.g., Hill, 2019). These initiatives involve YDPs examining data on program quality, developing improvement plans (e.g., staff trainings, the use of new curricula), and implementing the plans. This process is repeated over time to continuously improve program quality.

The current study found that about 5-10% of young people report very below average/negative experiences among all aspects of program quality, and it may be especially important to develop strategies to better meet the needs of these young people. Qualitative research suggests that it only takes a few young people to disrupt the positive functioning of a youth program (Larson & Walker, 2010). It is possible that the disruptive young people are often those 5-10% of participants who are classified in the Very Below Average and Negative profiles. Identifying these participants, asking about their experiences at the program, and then working to

improve the program to better meet these participants' needs may lead to positive changes for these young people and the overall program culture. In addition, since the 5-10% of young people in the Very Below Average and Negative profiles reported negative perceptions of *all* components of program quality, a multi-pronged approach for improving program quality may be most effective.

4.6 Study Strengths, Limitations, and Future Directions

One of this study's main strengths is the large sample size of over 165,000 program participants. The large sample size allowed for the identification of profiles that included only a small percentage of youth, since these profiles still contained a substantial number of individuals (Ferguson et al., 2019). The large sample size and national sample also contribute to the generalizability of the findings.

One of the limitations of the current study is that it relies solely on young people's perceptions of program quality. Young people's perceptions are important because they can give insight into experiences of program quality that may be missed by other methods, such as observational approaches (Baldwin et al., 2015; Bartko, 2005; Silliman & Schumm, 2013). However, observational measures are also important because they utilize standardized benchmarks when examining multiple programs (Oh et al., 2015). In addition, trained observers may be better able to assess certain aspects of program quality, such as the use of structured skill-building approaches. Utilizing multiple methods to assess program quality in the same study is critical for gathering a more complete understanding of quality and how it is associated with developmental outcomes (Deutsch et al., 2017).

In addition, these findings may not generalize to all Clubs within the BGCA network or to all youth development programs. A majority of BGCA Clubs (approximately 70%)

participated in the 2017 National Youth Outcome Initiative surveys that were examined in this study. This is a large percentage of the over 4,000 BGCA Clubs, suggesting that the results of this study may generalize to all BGCA members. However, it is possible that different profiles of program quality perceptions, including profiles with more variability across the indicators of quality, could emerge if youth data were examined at other youth programs or at the BGCA Clubs that did not participate in the 2017 National Youth Outcome Initiative surveys. Similarly, the cross-sectional and longitudinal results also may not generalize to youth programs that were not examined in this study.

Finally, BGCA's National Youth Outcomes Initiative was not designed as a longitudinal study. The current study took advantage of the fact that a sizeable proportion of BGCA participants who completed the annual BGCA survey in 2017 also took the survey in 2018. The participants that completed both surveys were younger on average and were more likely to attend the Club two or more times per week. Although the magnitudes of these differences were small, the longitudinal analyses may not be generalizable to the full age range of club participants or to participants that attend less regularly, and may be biased in other ways that were not examined. However, when comparing participants who completed both surveys to those who only completed the 2017 survey, the comparisons did not reach the cut-off for a small effect size for gender, free/reduced-price lunch status, years of program membership, program quality indicators, and social competence variables. Thus, the longitudinal data appear to be at least somewhat representative of the full sample. Moreover, as a strategy to reduce potential bias, sample weights were utilized to approximate the longitudinal results to the full sample that completed the 2017 survey (Seaman & White, 2013). The longitudinal findings were similar when sample weights were and were not used, suggesting minimal bias for the longitudinal

sample. However, the findings of this study should be replicated by longitudinal studies in order to increase confidence in the findings.

4.7 Conclusion

Youth program participants tended to rate various components of program quality in a similar way, and approximately 70% of young people rated all aspects of program quality above average. Participants that rated multiple aspects of program quality above average often reported higher levels of social competence, and sometimes reported higher levels of social competence over the course of a year. Thus, it is important for program leaders to prioritize strategies for improving program quality, especially for the subgroup of young people that report both low perceptions of program quality and low social competence.

REFERENCES

Arbreton, A., Bradshaw, M., Metz, R., Sheldon, J., & Pepper, S. (2008). *More Time For Teens: Understanding Teen Participation—Frequency, Intensity, and Duration—In Boys & Girls Clubs*. Public/Private Ventures and Boys & Girls Clubs of America. https://www.issuelab.org/resource/more-time-for-teens-understanding-teen-participation-frequency-intensity-and-duration-in-boys-girls-clubs.html

Asparouhov, T. (2017). *LCA with Covariates, Distal, and Imputation*. Mplus Discussion. http://www.statmodel.com/discussion/messages/13/24103.html?1542839703

Asparouhov, T., & Muthén, B. (2015). Residual associations in latent class and latent transition analysis. *Structural Equation Modeling*, *22*(2), 169–177. https://doi.org/10.1080/10705511.2014.935844

Asparouhov, T., & Muthén, B. (2018). *Auxiliary variables in mixture modeling: Using the BCH method in Mplus to estimate a distal outcome model and an arbitrary second model.* Mplus Web Notes. http://www.statmodel.com/examples/webnotes/webnote21.pdf

Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling*, *23*(1), 20–31. https://doi.org/10.1080/10705511.2014.955104

Baldwin, C. K., Stromwall, K., & Wilder, Q. (2015). Afterschool youth program design and structural quality: Implications for quality improvement. *Child & Youth Services*, *36*(3), 226–247. https://doi.org/10.1080/0145935X.2015.1046592

Bandura, A. (1977). *Social learning theory*. Prentice Hall.

Bartko, W. T. (2005). The ABCs of engagement in out-of-school-time programs. *New Directions for Youth Development*, *2005*(105), 109–120.

Bean, C., & Forneris, T. (2016). Re-examining the youth program quality survey as a tool to assess quality within youth programming. *Cogent Psychology*, *3*(1). http://dx.doi.org.ezproxy.gsu.edu/10.1080/23311908.2016.1149265

Berkowitz, L. (1972). Social Norms, Feelings, and Other Factors Affecting Helping and Altruism. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 6, pp. 63–108). Academic Press. https://doi.org/10.1016/S0065-2601(08)60025-8

Bohnert, A. M., Aikins, J. W., & Edidin, J. (2007). The Role of Organized Activities in Facilitating Social Adaptation across the Transition to College. *Journal of Adolescent Research*, *22*(2), 189–208. https://doi.org/10.1177/0743558406297940

Bolck, A., Croon, M., & Hagenaars, J. (2004). Estimating Latent Structure Models with Categorical Variables: One-Step Versus Three-Step Estimators. *Political Analysis*, *12*(1), 3–27. https://doi.org/10.1093/pan/mph001

Bornstein, M. H., Hahn, C.-S., & Haynes, O. M. (2010). Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence: Developmental cascades. *Development and Psychopathology*, *22*(4), 717–735. https://doi.org/10.1017/S0954579410000416

Boys & Girls Clubs of America. (2017). *Measuring the Impact of Boys & Girls Clubs: 2016 National Outcomes Report*. Boys & Girls Clubs of America. https://issuu.com/bgca/docs/2016_bgca_national__outcomes_report

Boys & Girls Clubs of America. (2018a). *2018 NYOI Member Survey Training and Administration Guide*. Boys & Girls Clubs of America.

Boys & Girls Clubs of America. (2018b). *Annual Report*. https://www.bgca.org/about-us/annual-report

Boys & Girls Clubs of America. (2018c). *Measuring the Impact of Boys & Girls Clubs: 2017 National Outcomes Report*. Boys & Girls Clubs of America. https://www.bgca.org/-/media/Documents/AboutUs/2017_NYOI_Report_LR.pdf

Boys & Girls Clubs of America. (2018d). *NYOI Guide to Measures: A Desk Reference for Boys & Girls Clubs Participating in the National Youth Outcomes Initiative (NYOI)*. Boys & Girls Clubs of America.

Boys & Girls Clubs of America. (2018e). *Our Mission & Story*. https://www.bgca.org/about-us/our-mission-story

Boys & Girls Clubs of America. (2018f). *Programs*. https://www.bgca.org/programs

Boys & Girls Clubs of America. (2020). *Teens and The CLUB*. https://www.bgca.org/i-am-a/teen

Brackett, M. A., Elbertson, N. A., & Rivers, S. E. (2015). Applying theory to the development of approaches to SEL. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice.* (pp. 20–32). Guilford Press.

Breslau, N., Reboussin, B. A., Anthony, J. C., & Storr, C. L. (2005). The Structure of Posttraumatic Stress Disorder: Latent Class Analysis in 2 Community Samples. *Archives of General Psychiatry*, *62*(12), 1343–1351. https://doi.org/10.1001/archpsyc.62.12.1343

Burt, K. B., Obradović, J., Long, J. D., & Masten, A. S. (2008). The interplay of social competence and psychopathology over 20 years: Testing transactional and cascade models. *Child Development*, *79*(2), 359–374. https://doi.org/10.1111/j.1467-8624.2007.01130.x

Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2019). Malleability, plasticity, and individuality: How children learn and develop in context. *Applied Developmental Science*, *23*(4), 307–337. https://doi.org/10.1080/10888691.2017.1398649

Caprara, G. V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science*, *11*(4), 302–306. https://doi.org/10.1111/1467-9280.00260

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2004). Positive Youth Development in the United States: Research Findings on Evaluations of Positive Youth Development Programs. *Annals of the American Academy of Political and Social Science*, *591*, 98–124.

Catalano, R. F., & Hawkins, J. D. (1996). The social development model: A theory of antisocial behavior. In J. D. Hawkins (Ed.), *Delinquency and crime: Current theories.* (pp. 149–197). Cambridge University Press.

Chapman, C. M., Deane, K. L., Harré, N., Courtney, M. G. R., & Moore, J. (2017). Engagement and mentor support as drivers of social development in the Project K youth development program. *Journal of Youth and Adolescence*, *46*(3), 644–655. https://doi.org/10.1007/s10964-017-0640-5

Chen, D. W. (2003). Preventing Violence by Promoting the Development of Competent Conflict Resolution Skills: Exploring Roles and Responsibilities. *Early Childhood Education Journal*, *30*(4), 203–208. https://doi.org/10.1023/A:1023379306124

Chen, H., Cohen, P., & Chen, S. (2010). How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Communications in Statistics—Simulation and Computation*. https://doi.org/10.1080/03610911003650383

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Wiley.

Corrales, T., Waterford, M., Goodwin-Smith, I., Wood, L., Yourell, T., & Ho, C. (2016). Childhood adversity, sense of belonging and psychosocial outcomes in emerging adulthood: A test of mediated pathways. *Children and Youth Services Review, 63*, 110–119. https://doi.org/10.1016/j.childyouth.2016.02.021

Denault, A.-S., & Poulin, F. (2016). What adolescents experience in organized activities: Profiles of individual and social experiences. *Journal of Applied Developmental Psychology, 42*(Supplement C), 40–48. https://doi.org/10.1016/j.appdev.2015.11.004

Denham, S. A., & Brown, C. (2010). "Plays Nice With Others": Social–Emotional Learning and Academic Success. *Early Education and Development*, *21*(5), 652–680. https://doi.org/10.1080/10409289.2010.497450

Deschenes, S. N., Arbreton, A., Little, P. M., Herrera, C., Grossman, J. B., & Weiss, H. B. (2010). *Engaging Older Youth: Program and City-Level Strategies to Support Sustained Participation in Out-of-School Time*. Harvard Family Research Project, Public/Private Ventures, and The Wallace Foundation. http://www.wallacefoundation.org/knowledge-center/Documents/Engaging-Older-Youth-City-Level-Strategies-Support-Sustained-Participation-Out-of-School-Time.pdf

Deshpande, S. K., Hasegawa, R. B., Weiss, J., & Small, D. S. (2020). The association between adolescent football participation and early adulthood depression. *PLOS ONE*, *15*(3), e0229978. https://doi.org/10.1371/journal.pone.0229978

Deutsch, N. L., Blyth, D. A., Kelley, J., Tolan, P. H., & Lerner, R. M. (2017). Let's talk after-school: The promises and challenges of positive youth development for after-school research, policy, and practice. In N. L. Deutsch (Ed.), *After-school programs to promote positive youth development: Integrating research into practice and policy.* (pp. 45–68). Springer Science + Business Media. https://doi.org/10.1007/978-3-319-59132-2_4

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017). Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development, 88*(2), 408–416. https://doi.org/10.1111/cdev.12739

Duerden, M. D., & Witt, P. A. (2010). The impact of socialization on youth program outcomes: A social development model perspective. *Leisure Sciences, 32*(4), 299–317. https://doi.org/10.1080/01490400.2010.488189

Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American Journal of Community Psychology, 45*(3–4), 294–309. https://doi.org/10.1007/s10464-010-9300-6

Dworkin, J. B., Larson, R., & Hansen, D. (2003). Adolescents' accounts of growth experiences in youth activities. *Journal of Youth and Adolescence, 32*(1), 17–26. https://doi.org/10.1023/A:1021076222321

Eccles, J. S., Barber, B. L., Stone, M., & Hunt, J. (2003). Extracurricular Activities and Adolescent Development. *Journal of Social Issues, 59*(4), 865. https://doi.org/10.1046/j.0022-4537.2003.00095.x

Eisenberg, N. (2014). Empathy-Related Responding and Its Relations to Socioemotional Development. *Asia-Pacific Journal of Research in Early Childhood Education*, *8*(1), 1–17.

Elias, M. J., & Gordon, J. S. (2009). Promoting social and emotional development in childhood and early adolescence. In T. P. Gullotta, M. Bloom, C. F. Gullotta, & J. C. Messina (Eds.), *A blueprint for promoting academic and social competence in after-school programs.* (pp. 63–77). Springer Publishing Co. https://doi.org/10.1007/978-0-387-79920-9_4

Ferguson, S. L., G. Moore, E. W., & Hull, D. M. (2019). Finding latent groups in observed data: A primer on latent profile analysis in Mplus for applied researchers. *International Journal of Behavioral Development*, 0165025419881721. https://doi.org/10.1177/0165025419881721

Flett, M. R., Gould, D., & Lauer, L. (2012). A study of an underserved youth sports program using the Youth Program Quality Assessment. *Journal of Applied Sport Psychology*, *24*(3), 275–289. https://doi.org/10.1080/10413200.2011.641061

Flunger, B., Trautwein, U., Nagengast, B., Lüdtke, O., Niggli, A., & Schnyder, I. (2019). Using Multilevel Mixture Models in Educational Research: An Illustration With Homework Research. *The Journal of Experimental Education*, 1–28. https://doi.org/10.1080/00220973.2019.1652137

Fredricks, J. A., Alfeld-Liro, C. J., Hruda, L. Z., Eccles, J. S., Patrick, H., & Ryan, A. M. (2002). A qualitative exploration of adolescents' commitment to athletics and the arts. *Journal of Adolescent Research*, *17*(1), 68–97. https://doi.org/10.1177/0743558402171004

Fredricks, J. A., Naftzger, N., Smith, C., & Riley, A. (2017). Measuring youth participation, program quality, and social and emotional skills in after-school programs. In N. L. Deutsch (Ed.), *After-school programs to promote positive youth development: Integrating research into practice and policy.* (pp. 23–43). Springer Science + Business Media. https://doi.org/10.1007/978-3-319-59132-2_3

Gershoff, E. T., Aber, J. L., Raver, C. C., & Lennon, M. C. (2007). Income Is Not Enough: Incorporating Material Hardship Into Models of Income Associations With Parenting and Child Development. *Child Development*, *78*(1), 70–95. https://doi.org/10.1111/j.1467-8624.2007.00986.x

Gravetter, F. J., & Wallnau, L. B. (2013). *Statistics for the Behavioral Sciences* (9th ed.). Wadsworth Cengage Learning.

Gresham, F. M. (2016). Social skills assessment and intervention for children and youth. *Cambridge Journal of Education*, *46*(3), 319–332. https://doi.org/10.1080/0305764X.2016.1195788

Guest, A. M. (2008). Reconsidering teamwork: Popular and local meanings for a common ideal associated with positive youth development. *Youth & Society*, *39*(3), 340–361. https://doi.org/10.1177/0044118X06297073

Hansen, D. M., Larson, R. W., & Dworkin, J. B. (2003). What Adolescents Learn in Organized Youth Activities: A Survey of Self-Reported Developmental Experiences. *Journal of Research on Adolescence*, *13*(1), 25–55. https://doi.org/10.1111/1532-7795.1301006

Hawkins, J. D., & Weis, J. G. (1985). The social development model: An integrated approach to delinquency prevention. *Journal of Primary Prevention*, *6*(2), 73–97. https://doi.org/10.1007/BF01325432

Hill, C. (2019, January 31). DIVING IN to Continuous Quality Improvement: A CQI Primer. *Club Experience Blog: A Youth Development Resource From Boys & Girls Clubs of America*. https://clubexperience.blog/2019/01/31/diving-in-to-continuous-quality-improvement-a-cqi-primer/

Hirsch, B. J., Deutsch, N. L., & DuBois, D. L. (2011). *After-school centers and youth development: Case studies of success and failure*. Cambridge University Press. https://doi.org/10.1017/CBO9780511994913

Huber, L., Plötner, M., & Schmitz, J. (2018). Social competence and psychopathology in early childhood: A systematic review. *European Child & Adolescent Psychiatry*, 1–17. https://doi.org/10.1007/s00787-018-1152-x

Kataoka, S., & Vandell, D. L. (2013). Quality of afterschool activities and relative change in adolescent functioning over two years. *Applied Developmental Science, 17*(3), 123–134. https://doi.org/10.1080/10888691.2013.804375

Khanlou, N., & Wray, R. (2014). A whole community approach toward child and youth resilience promotion: A review of resilience literature. *International Journal of Mental Health and Addiction, 12*(1), 64–79. https://doi.org/10.1007/s11469-013-9470-1

Kievit, R., Frankenhuis, W. E., Waldorp, L., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00513

Kim, H.-Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative Dentistry & Endodontics, 42*(2), 152–155. https://doi.org/10.5395/rde.2017.42.2.152

Kiser, L. J., & Black, M. M. (2005). Family processes in the midst of urban poverty: What does the trauma literature tell us? *Aggression and Violent Behavior*, *10*(6), 715–750. https://doi.org/10.1016/j.avb.2005.02.003

Klein, M. R., Lengua, L. J., Thompson, S. F., Moran, L., Ruberry, E. J., Kiff, C., & Zalewski, M. (2018). Bidirectional Relations Between Temperament and Parenting Predicting Preschool-Age Children's Adjustment. *Journal of Clinical Child & Adolescent Psychology*, *47*(sup1), S113–S126. https://doi.org/10.1080/15374416.2016.1169537

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford.

Kuperminc, G., Seitz, S., Joseph, H., Khatib, N., Wilson, C., Collins, K., & Guessous, O. (2019). Enhancing Program Quality in a National Sample of After-School Settings: The Role of Youth-Staff Interactions and Staff/Organizational Functioning. *American Journal of Community Psychology*, *63*, 391–404. https://doi.org/10.1002/ajcp.12329

Larson, R. W., & Walker, K. C. (2010). Dilemmas of practice: Challenges to program quality encountered by youth program leaders. *American Journal of Community Psychology*, *45*(3–4), 338–349. https://doi.org/10.1007/s10464-010-9307-z

Laursen, B., & Hoff, E. (2006). Person-Centered and Variable-Centered Approaches to Longitudinal Data. *Merrill-Palmer Quarterly*, *52*(3), 377–389. https://doi.org/10.1353/mpq.2006.0029

Lengua, L. J. (2012). Poverty, the development of effortful control, and children's academic, social, and emotional adjustment. In V. Maholmes & R. B. King (Eds.), *The Oxford handbook of poverty and child development.* (pp. 491–511). Oxford University Press.

Lerner, J., Bowers, E. P., Minor, K., Boyd, M. J., Mueller, M. K., Schmid, K. L., Napolitano, C. M., Lewin-Bizan, S., & Lerner, R. M. (2013). Positive youth development: Processes,

philosophies, and programs. In R. M. Lerner, M. A. Easterbrooks, J. Mistry, & I. B. Weiner (Eds.), *Handbook of psychology: Developmental psychology* (2nd ed., Vol. 6, pp. 365–392). John Wiley & Sons, Inc.

Mahoney, J. L., Parente, M. E., & Zigler, E. F. (2010). After-school program participation and children's development. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 379–397). Routledge.

Mäkikangas, A., Tolvanen, A., Aunola, K., Feldt, T., Mauno, S., & Kinnunen, U. (2018). Multilevel Latent Profile Analysis with Covariates: Identifying Job Characteristics Profiles in Hierarchical Data as an Example. *Organizational Research Methods*, *21*(4), 931–954. https://doi.org/10.1177/1094428118760690

Martí, M., Bonillo, A., Jané, M. C., Fisher, E. M., & Duch, H. (2016). Cumulative risk, the mother–child relationship, and social-emotional competence in Latino Head Start children. *Early Education and Development*, *27*(5), 590–622. https://doi.org/10.1080/10409289.2016.1106202

McDonough, M. H., Ullrich-French, S., Anderson-Butcher, D., Amorose, A. J., & Riley, A. (2013). Social responsibility among low-income youth in physical activity-based positive youth development programs: Scale development and associations with social relationships. *Journal of Applied Sport Psychology*, *25*(4), 431–447. https://doi.org/10.1080/10413200.2012.751563

McKeown, S., & Taylor, L. K. (2018). Perceived peer and school norm effects on youth antisocial and prosocial behaviours through intergroup contact in Northern Ireland. *British Journal of Social Psychology*, *57*(3), 652–665. https://doi.org/10.1111/bjso.12257

Meschke, L. L., Peter, C. R., & Bartholomae, S. (2012). Developmentally appropriate practice to promote healthy adolescent development: Integrating research and practice. *Child & Youth Care Forum*, *41*(1), 89–108. https://doi.org/10.1007/s10566-011-9153-7

Milligan, K., Sibalis, A., Morgan, A., & Phillips, M. (2017). Social competence: Consideration of behavioral, cognitive, and emotional factors. In J. L. Matson (Ed.), *Handbook of social behavior and skills in children.* (pp. 63–82). Springer International Publishing. https://doi.org/10.1007/978-3-319-64592-6_5

Moore, K. A., & Hamilton, K. (2010). *How out-of-school time program quality is related to adolescent outcomes* (No. 2010–19). Child Trends. https://www.childtrends.org/wp-content/uploads/2010/08/Child_Trends-2010_08_02_RB_OSTProgramQuality.pdf

Morgan, G. B., Hodge, K. J., & Baggett, A. R. (2016). Latent profile analysis with nonnormal mixtures: A Monte Carlo examination of model selection using fit indices. *Computational Statistics & Data Analysis*, *93*, 146–161. https://doi.org/10.1016/j.csda.2015.02.019

Muthén, B. (2017). *Local Independence Assumption*. Mplus Discussion. http://www.statmodel.com/discussion/messages/13/3895.html?1536115968

Muthén, B. (2018). *LCA with Covariates, Missing Data, and Plots*. Mplus Discussion. http://www.statmodel.com/discussion/messages/13/626.html?1540427590

Muthén, L., & Muthén, B. (2017). *Mplus User's Guide.* (8th ed.). Muthén & Muthén. https://www.statmodel.com/download/usersguide/Mplus%20Users%20Guide%20v6.pdf

National Research Council, & Institute of Medicine. (2002). *Community Programs to Promote Youth Development* (J. Eccles & J. Gootman, Eds.). The National Academies Press.

Navarro-González, D., Lorenzo-Seva, U., & Vigil-Colet, A. (2016). How response bias affects the factorial structure of personality self-reports. *Psicothema*, *28*(4), 465–470.

Nylund, K. L., Asparouhov, T., & Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*(4), 535–569. https://doi.org/10.1080/10705510701575396

Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science*, *4*(4), 440–461. https://doi.org/10.1037/tps0000176

Oh, Y., Osgood, D. W., & Smith, E. P. (2015). Measuring afterschool program quality using setting-level observational approaches. *The Journal of Early Adolescence*, *35*(5–6), 681–713.

Oshri, A., Topple, T. A., & Carlson, M. W. (2017). Positive youth development and resilience: Growth patterns of social skills among youth investigated for maltreatment. *Child Development*, *88*(4), 1087–1099. https://doi.org/10.1111/cdev.12865

Overstreet, S., & Mathews, T. (2011). Challenges associated with exposure to chronic trauma: Using a public health framework to foster resilient outcomes among youth. *Psychology in the Schools*, *48*(7), 738–754. https://doi.org/10.1002/pits.20584

Pierce, K. M., Bolt, D. M., & Vandell, D. L. (2010). Specific features of after-school program quality: Associations with children's functioning in middle childhood. *American Journal of Community Psychology*, *45*(3–4), 381–393. https://doi.org/10.1007/s10464-010-9304-2

Pittman, K. J. (2017). Positive youth development as a strategy for addressing readiness and equity: A commentary. *Child Development*, *88*(4), 1172–1174. https://doi.org/10.1111/cdev.12872

Ramey, H. L., Lawford, H. L., Rose-Krasnor, L., Freeman, J., & Lanctot, J. (2018). Engaging diverse Canadian youth in youth development programs: Program quality and community engagement. *Children and Youth Services Review*, *94*, 20–26. https://doi.org/10.1016/j.childyouth.2018.09.023

Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment*, *25*(4), 1137–1145. https://doi.org/10.1037/a0033323

Redmond, S., & Dolan, P. (2016). Towards a conceptual model of youth leadership development. *Child & Family Social Work*, *21*(3), 261–271. https://doi.org/10.1111/cfs.12146

Repetti, R. L., Taylor, S. E., & Seeman, T. E. (2002). Risky families: Family social environments and the mental and physical health of offspring. *Psychological Bulletin*, *128*(2), 330–366. https://doi.org/10.1037/0033-2909.128.2.330

Rosch, D., Ogolsky, B., & Stephens, C. M. (2017). Trajectories of Student Leadership Development through Training: An Analysis by Gender, Race, and Prior Exposure. *Journal of College Student Development*, *58*(8), 1184–1200. https://doi.org/10.1353/csd.2017.0093

Rose-Krasnor, L. (1997). The nature of social competence: A theoretical review. *Social Development*, *6*(1), 111–135. https://doi.org/10.1111/j.1467-9507.1997.tb00097.x

Roth, J. L., & Brooks-Gunn, J. (2003). What exactly is a youth development program? Answers from research and practice. *Applied Developmental Science, 7*(2), 94–111. https://doi.org/10.1207/S1532480XADS0702_6

Roth, J. L., & Brooks-Gunn, J. (2016). Evaluating youth development programs: Progress and promise. *Applied Developmental Science*, *20*(3), 188–202. https://doi.org/10.1080/10888691.2015.1113879

Roth, J. L., Malone, L. M., & Brooks-Gunn, J. (2010). Does the amount of participation in afterschool programs relate to developmental outcomes? A review of the literature. *American Journal of Community Psychology*, *45*(3–4), 310–324. https://doi.org/10.1007/s10464-010-9303-3

Ruiz, M. A., & Dorritie, M. T. (2020). Clinical Utility of the Minnesota Multiphasic Personality Inventory–2–Restructured Form (MMPI-2-RF) in a Residential Treatment Program for Homeless Individuals. *Assessment*, 107319111989948. https://doi.org/10.1177/1073191119899481

Sandy, S. V. (2014). The development of conflict resolution skills: Preschool to adulthood. In P. T. Coleman, M. Deutsch, & E. C. Marcus (Eds.), *The handbook of conflict resolution: Theory and practice.* (pp. 430–463). Jossey-Bass.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science, 18*(5), 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x

Seaman, S. R., & White, I. R. (2013). Review of inverse probability weighting for dealing with missing data. *Statistical Methods in Medical Research, 22*(3), 278–295.

Serido, J., Borden, L. M., & Perkins, D. F. (2011). Moving Beyond Youth Voice. *Youth & Society*, *43*(1), 44–63. https://doi.org/10.1177/0044118X09351280

Sherif, M. (1936). *The psychology of social norms*. Harper & Brothers.

Shernoff, D. J. (2010). Engagement in after-school programs as a predictor of social competence and academic performance. *American Journal of Community Psychology*, *45*(3–4), 325–337. https://doi.org/10.1007/s10464-010-9314-0

Shonkoff, J. P., Garner, A. S., The Committee on Psychosocial Aspects of Child and Family Health, Committee on Early Childhood, Adoption, and Dependent Care, Section on Developmental and Behavioral Pediatrics, Siegel, B. S., Dobbins, M. I., Earls, M. F., Garner, A. S., McGuinn, L., Pascoe, J., & Wood, D. L. (2012). The Lifelong Effects of Early Childhood Adversity and Toxic Stress. *Pediatrics*, *129*(1), e232–e246. https://doi.org/10.1542/peds.2011-2663

Silliman, B., & Schumm, W. R. (2013). Youth program quality survey: Youth assessment of program quality. *Marriage & Family Review*, *49*(7), 647–670. https://doi.org/10.1080/01494929.2013.803010

Smith, C., Peck, S. C., Denault, A.-S., Blazevski, J., & Akiva, T. (2010). Quality at the point of service: Profiles of practice in after-school settings. *American Journal of Community Psychology*, *45*(3–4), 358–369. https://doi.org/10.1007/s10464-010-9315-z

Smith, E. P., Witherspoon, D. P., & Wayne Osgood, D. (2017). Positive youth development among diverse racial–ethnic children: Quality afterschool contexts as developmental assets. *Child Development*, *88*(4), 1063–1078. https://doi.org/10.1111/cdev.12870

Sullivan, P. J., & Larson, R. W. (2010). Connecting Youth to High-Resource Adults: Lessons from Effective Youth Programs. *Journal of Adolescent Research*, *25*(1), 99–123. https://doi.org/10.1177/0743558409350505

Terry, D. J., & Hogg, M. A. (1996). Group norms and the attitude–behavior relationship: A role for group identification. *Personality and Social Psychology Bulletin*, *22*(8), 776–793. https://doi.org/10.1177/0146167296228002

Urban, J. B. (2008). Components and characteristics of youth development programs: The voices of youth-serving policymakers, practitioners, researchers, and adolescents. *Applied Developmental Science*, *12*(3), 128–139. https://doi.org/10.1080/10888690802199400

Van Eck, K., Johnson, S. R., Bettencourt, A., & Johnson, S. L. (2017). How school climate relates to chronic absence: A multi–level latent profile analysis. *Journal of School Psychology*, *61*, 89–102. https://doi.org/10.1016/j.jsp.2016.10.001

Vandell, D. L., Shumow, L., & Posner, J. (2005). After-school programs for low-income children: Differences in program quality. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after-school and community programs.* (pp. 437–456). Lawrence Erlbaum Associates Publishers.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. *Applied Latent Class Analysis*, *11*, 89–106.

Villalta, L., Smith, P., Hickin, N., & Stringaris, A. (2018). Emotion regulation difficulties in traumatized youth: A meta-analysis and conceptual review. *European Child & Adolescent Psychiatry*, *27*(4), 527–544. https://doi.org/10.1007/s00787-018-1105-4

Walker, K., & Arbreton, A. (2004). *After-school pursuits: An examination of outcomes in the San Francisco Beacon initiative*. Public/Private Ventures and Stanford University School of Education. https://www.issuelab.org/resource/after-school-pursuits-an-examination-of-outcomes-in-the-san-francisco-beacon-initiative.html

Waters, E., & Sroufe, L. A. (1983). Social competence as a developmental construct. *Developmental Review*, *3*(1), 79–97. https://doi.org/10.1016/0273-2297(83)90010-2

Wilson, C. Y.-A., Seitz, S., Broomfield-Massey, K., Whitehead, L., Mangia, J., Pridgeon, K. M., & Kuperminc, G. (2017). "It's more like we want to come to this": Program Engagement in a Sexual Health Youth Leadership Council. *Journal of Youth Development*, *12*(3), 37–49.

Wright, R., John, L., Alaggia, R., & Sheel, J. (2006). Community-based arts program for youth in low-income communities: A multi-method evaluation. *Child & Adolescent Social Work Journal*, *23*(5–6), 635–652. https://doi.org/10.1007/s10560-006-0079-0

Yohalem, N., & Wilson-Ahlstrom, A. (2010). Inside the black box: Assessing and improving quality in youth programs. *American Journal of Community Psychology*, *45*(3–4), 350–357. https://doi.org/10.1007/s10464-010-9311-3

Yohalem, N., Wilson-Ahlstrom, A., Fischer, S., & Shinn, M. (2009). *Measuring Youth Program Quality: A Guide to Assessment Tools, Second Edition*. The Forum for Youth Investment. http://forumfyi.org/files/MeasuringYouthProgramQuality_2ndEd.pdf

Yoshikawa, H., Aber, J. L., & Beardslee, W. R. (2012). The effects of poverty on the mental, emotional, and behavioral health of children and youth: Implications for prevention. *American Psychologist*, *67*(4), 272–284. https://doi.org/10.1037/a0028015

APPENDICES

Appendix A: Study Measures

Appendix A.1: Program Quality/Club Experience Measures

Sense of Belonging. How much do you agree or disagree with the following statements?

1. I feel like I belong here.
2. I feel like my ideas count here.
3. People listen to me here.

Answer choices: Strongly Agree, Agree, Disagree, Strongly Disagree.

Emotional Safety. How much do you agree or disagree with the following statements?

1. People make sure rules about how we treat each other are followed.
2. I feel respected by staff at the Boys & Girls Club.
3. I feel respected by other kids at the Boys & Girls Club.
4. This Boys & Girls Club has rules for how people are supposed to treat each other.

Answer choices: Strongly Agree, Agree, Disagree, Strongly Disagree.

Physical Safety. How often do these things happen at the Boys & Girls Club?

1. I feel safe when I am at the Boys & Girls Club.
2. If someone wanted to hurt me or beat me up here, someone at the Boys & Girls Club would stop them.

Answer choices: All of the Time, Most of the Time, Sometimes, Never.

Fun. At the Club…

1. I have a good time.
2. I enjoy coming to the Boys & Girls Club.
3. I have more fun at the Boys & Girls Club than other places I spend time.

Answer choices: Not At All True, Not Very True, Sort of True, Very True.

Adult Caring Relationships. About how many staff at the Boys & Girls Club…

1. Pay attention to what's going on in your life?
2. Would say something to you if something in your life wasn't going right?
3. Say something nice to you when you do something good?
4. Could you talk to if you are upset or mad about something?
5. Could you go to for help in a crisis?
6. Could you go to if you need advice about personal problems?

Answer choices: None, One, Two or Three, More than Three.

Staff Expectations. At the Club…

1. There is an adult who believes that I will be a success.
2. There is an adult who expects me to follow the rules.
3. There is an adult who always wants me to do my best.

Answer choices: Not At All True, Not Very True, Sort of True, Very True.

Staff Recognition. At the Club…

1. Staff reward me when I do a good job.
2. Staff let others know when I do a good job.
3. Staff notice when I try hard.

Answer choices: Not At All True, Not Very True, Sort of True, Very True.

Appendix A.2: Social Competence Measures

Leadership.

1. Once I know what needs to be done, I am good at planning how to do it.
2. I am pretty good at organizing a team of kids to do a project.

3. If I'm the leader of a group, I make sure that everyone in the group feels important.
4. I feel like I can stand up for what I think is right, even if my friends disagree.

Answer choices: Strongly Agree, Agree, Disagree, and Strongly Disagree.

Conflict Resolution.

1. When I have problems with other people my age, I talk to an adult about it.
2. When I have problems with other people my age, I push or hit the other person so that it doesn't happen again.
3. When I have problems with other people my age, I yell at them.
4. When I have problems with other people my age, I talk things over with them.
5. When other people my age try to hit or push me around, I fight back.

Answer choices: Very True, Sort of True, Not Very True, Not True at All.

Teamwork.

1. I listen to what other people say.
2. I am willing to do whatever the group needs me to do.

Answer choices: Very True, Sort of True, Not Very True, Not True at All.

Concern for Others.

1. I try to help when I see people in need.
2. When I make a decision, I try to think about how other people will be affected.
3. I want to help when I see someone having a problem.

Answer choices: Strongly Agree, Agree, Disagree, and Strongly Disagree.

Appendix B: Figures of All Latent Profile Analysis Solutions for Youth and Teens

Solutions are not displayed if the maximum log likelihood value was not replicated.

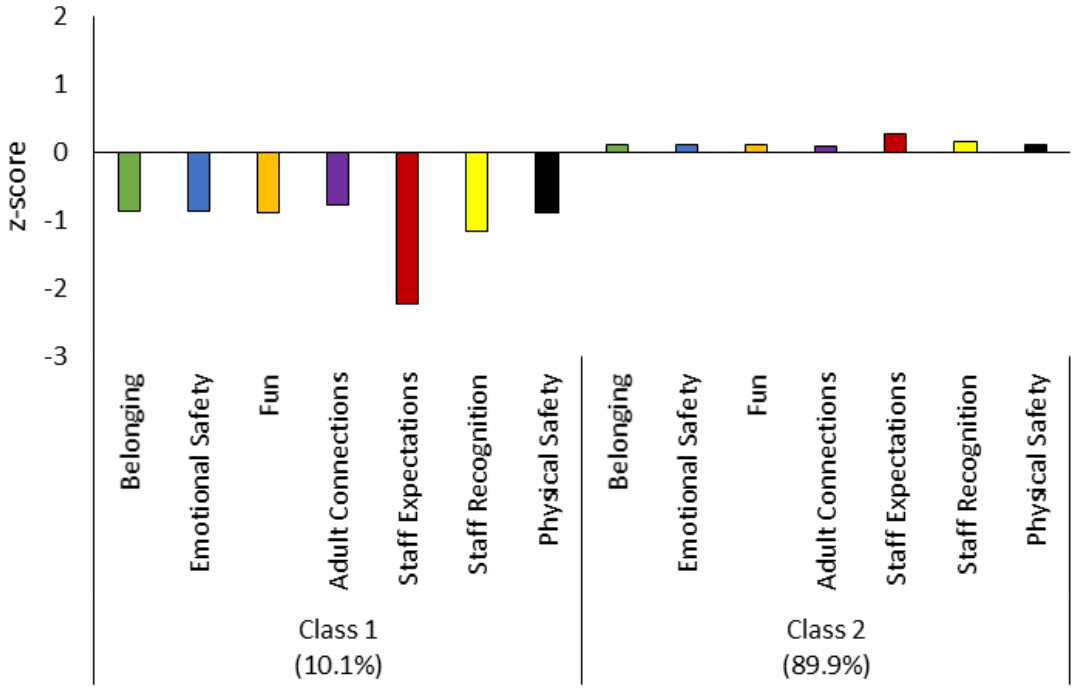

Figure B.1 Two-profile solution for indicators of program quality among youth (9-12 years old); n = 120,643.

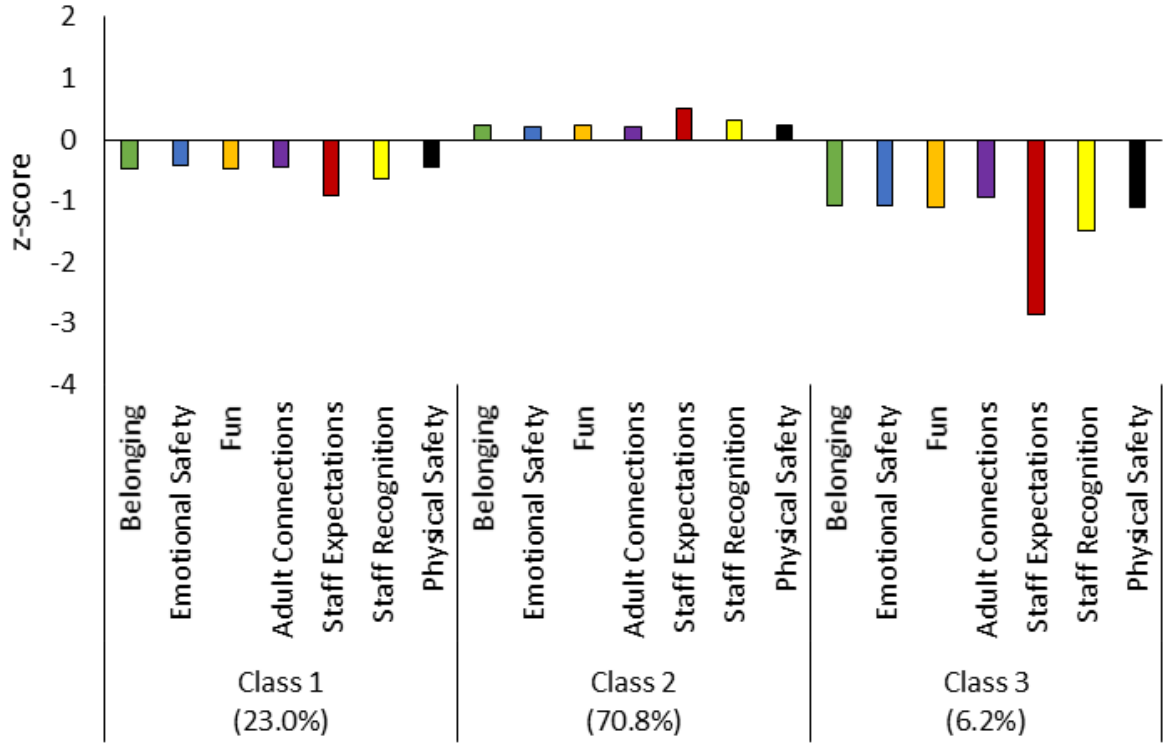

Figure B.2 Three-profile solution for indicators of program quality among youth (9-12 years old); n = 120,643.

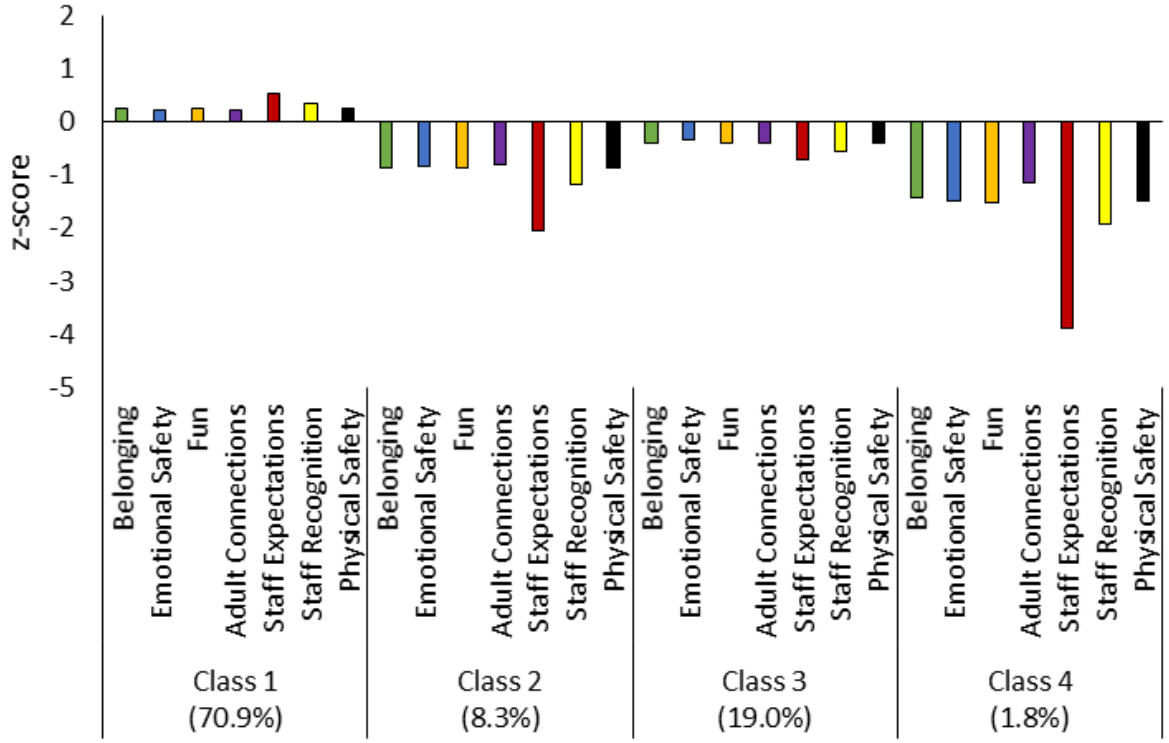

Figure B.3 Four-profile solution for indicators of program quality among youth (9-12 years old); n = 120,643.

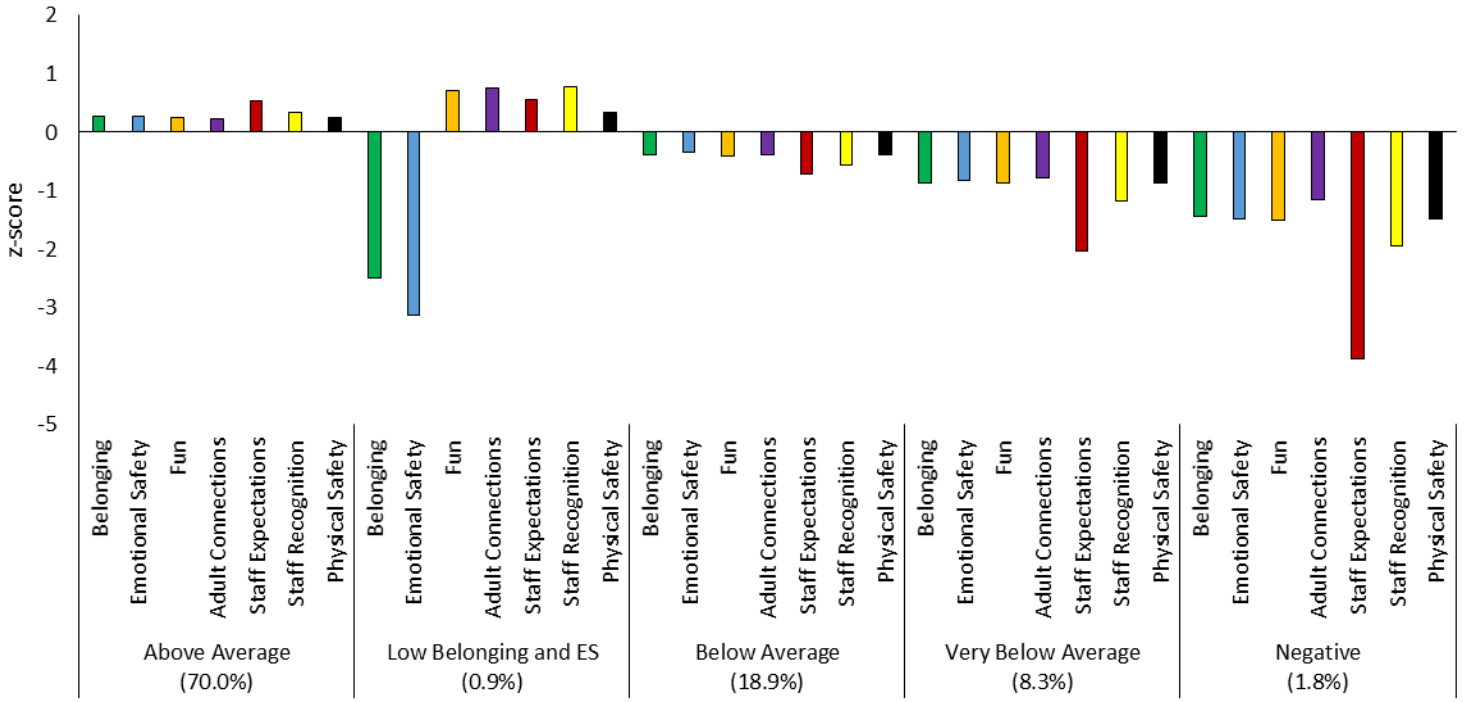

Figure B.4 Five-profile solution for indicators of program quality among youth (9-12 years old); ES = Emotional Safety; n = 120,643.

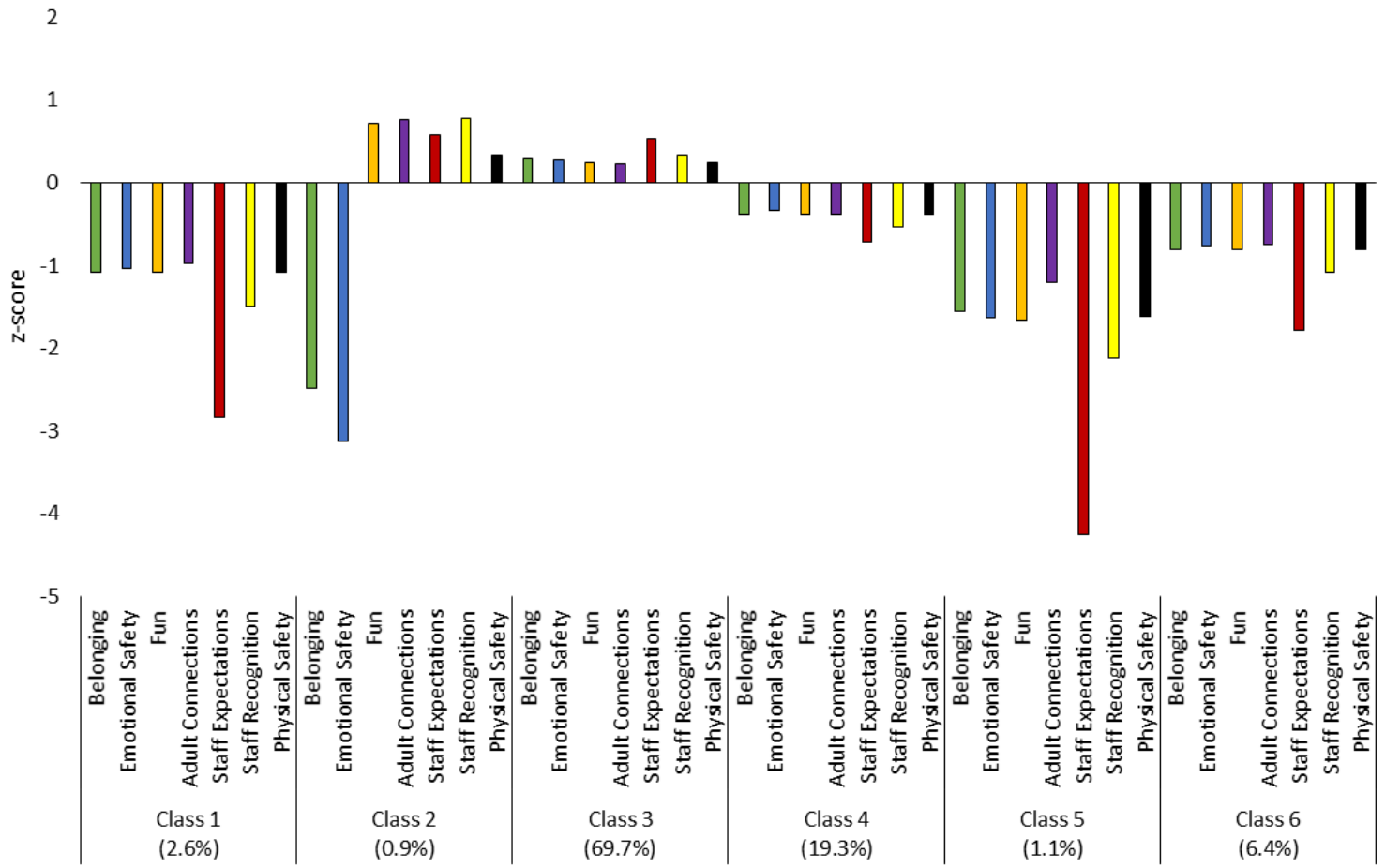

Figure B.5 Six-profile solution for indicators of program quality among youth (9-12 years old); n = 120,643.

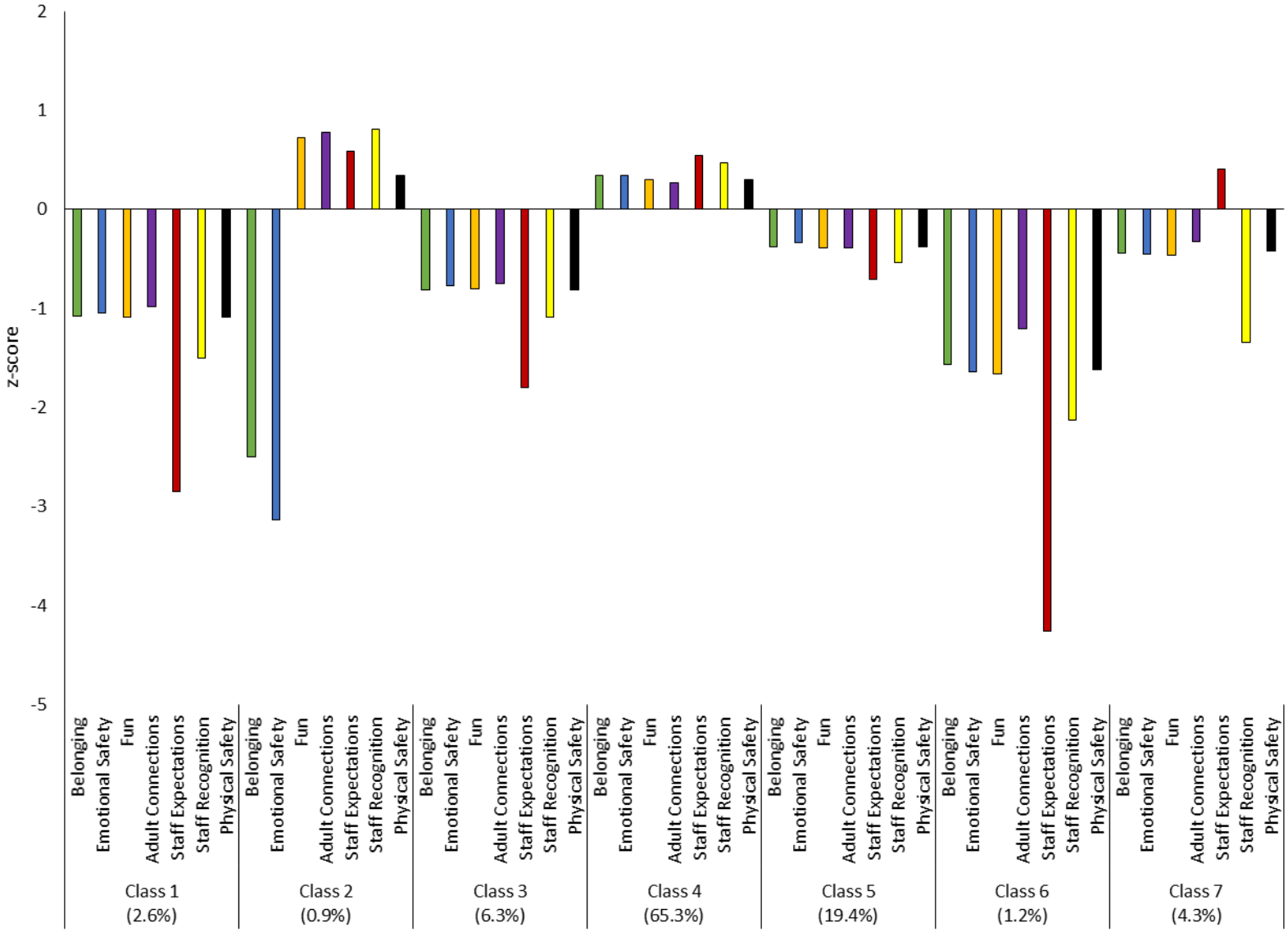

Figure B.6 Seven-profile solution for indicators of program quality among youth (9-12 years old); n = 120,643.

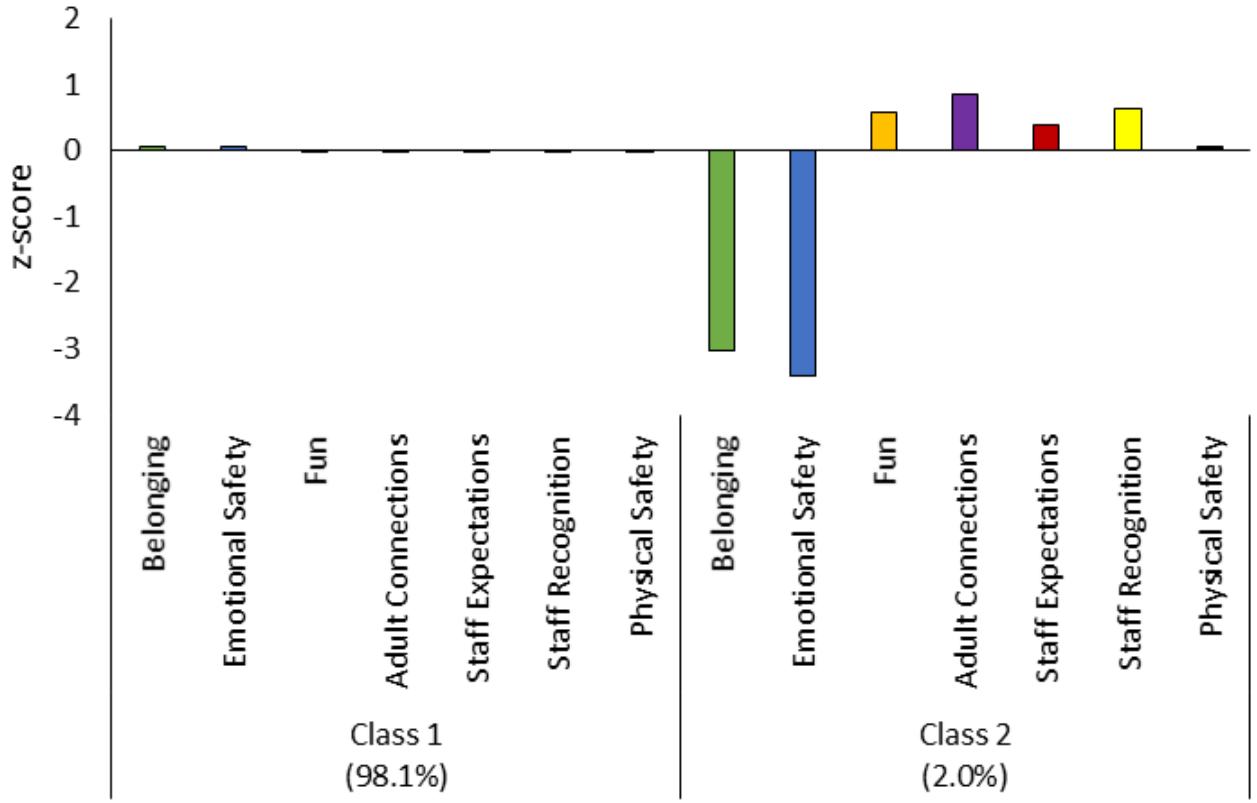

Figure B.7 Two-profile solution for indicators of program quality among teens (13-18 years old); n = 45,816.

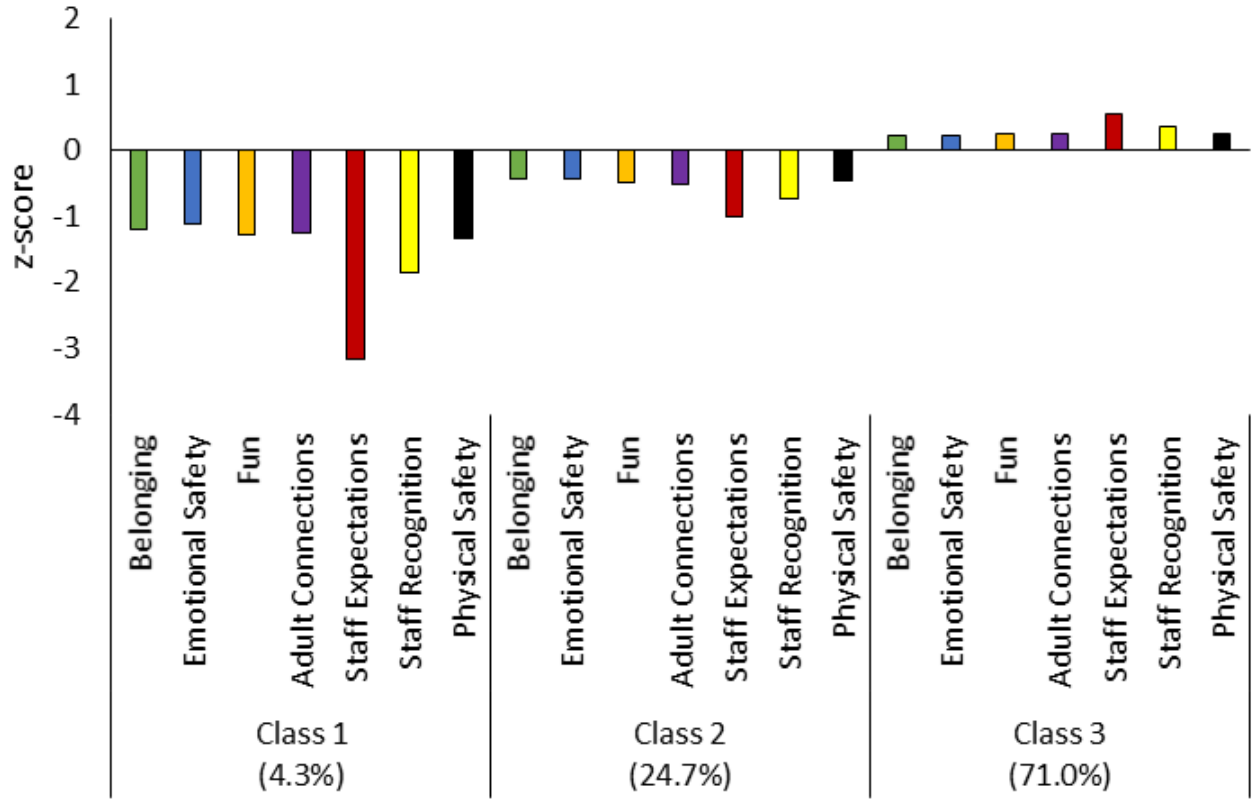

Figure B.8 Three-profile solution for indicators of program quality among teens (13-18 years old); n = 45,816.

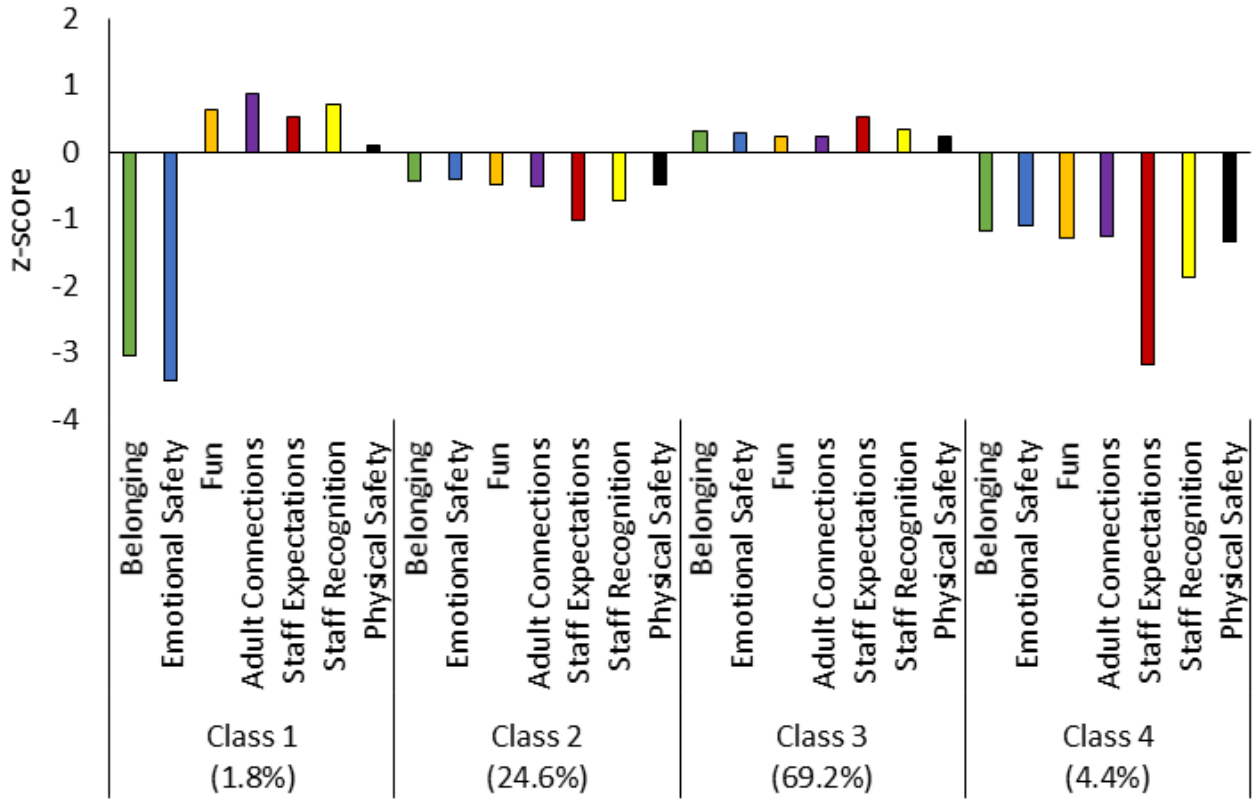

Figure B.9 Four-profile solution for indicators of program quality among teens (13-18 years old); n = 45,816.

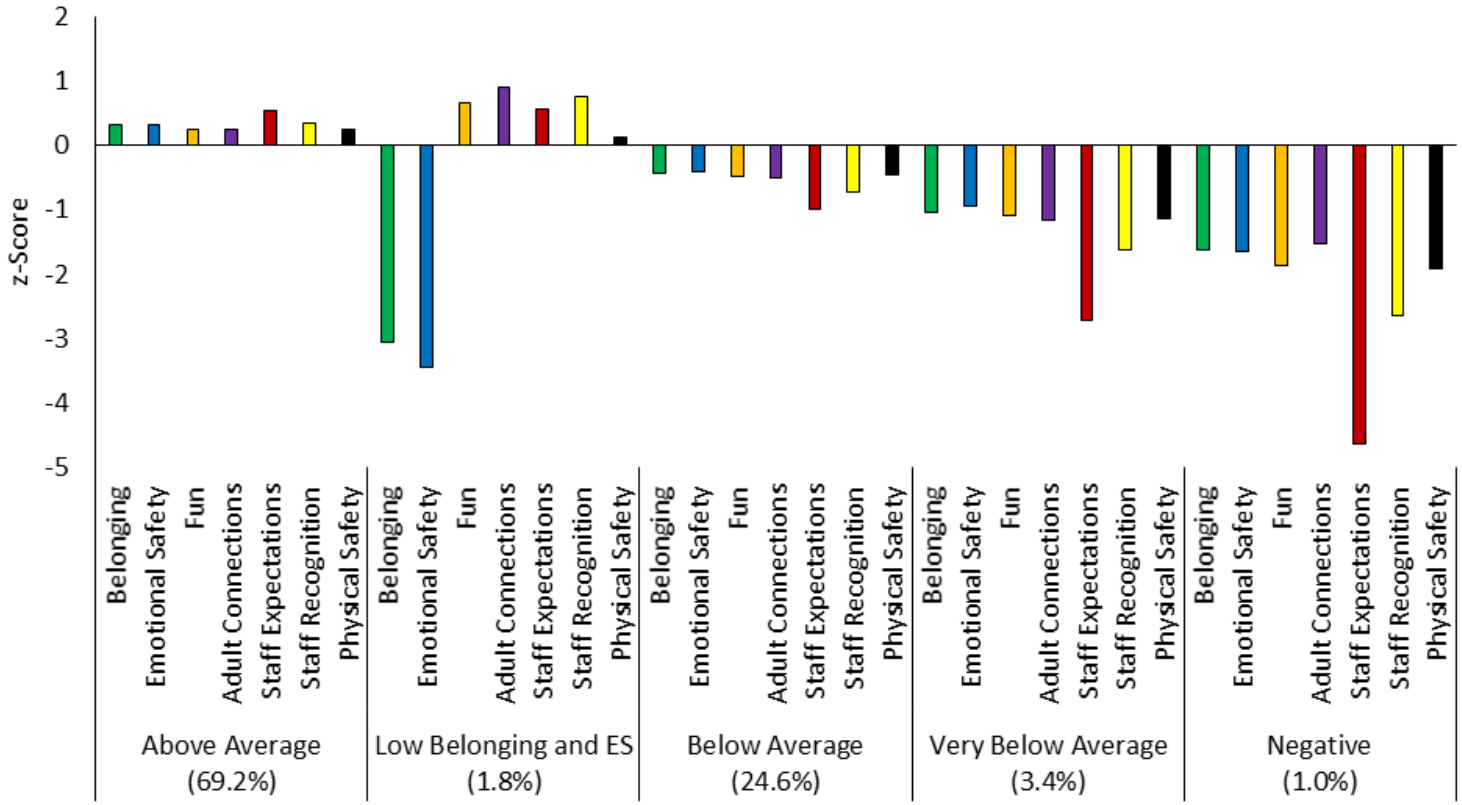

Figure B.10 Five-profile solution for indicators of program quality among teens (13-18 years old); ES = Emotional Safety; n = 45,816.

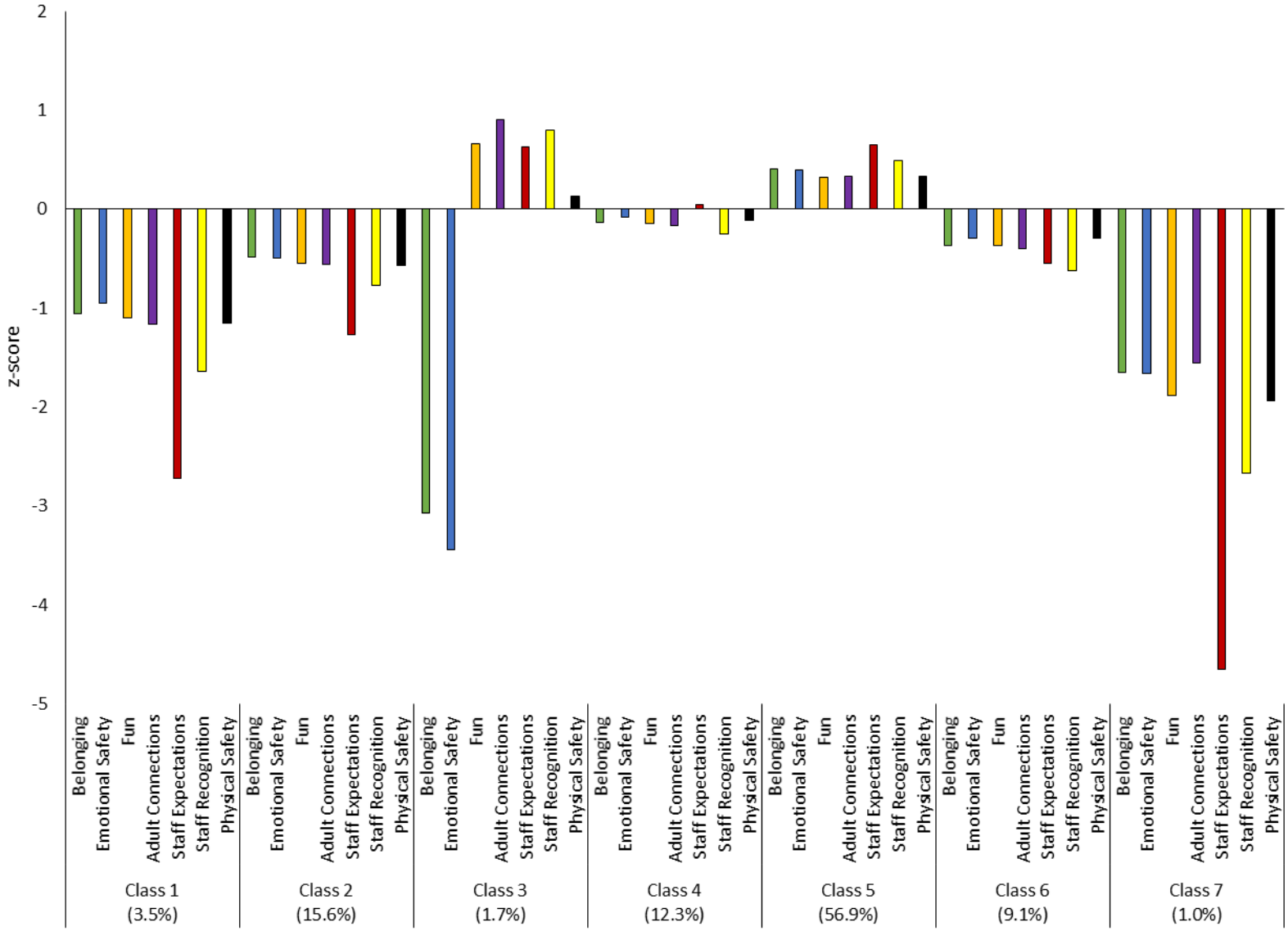

Figure B.11 Seven-profile solution for indicators of program quality among teens (13-18 years old); *n* = 45,816.

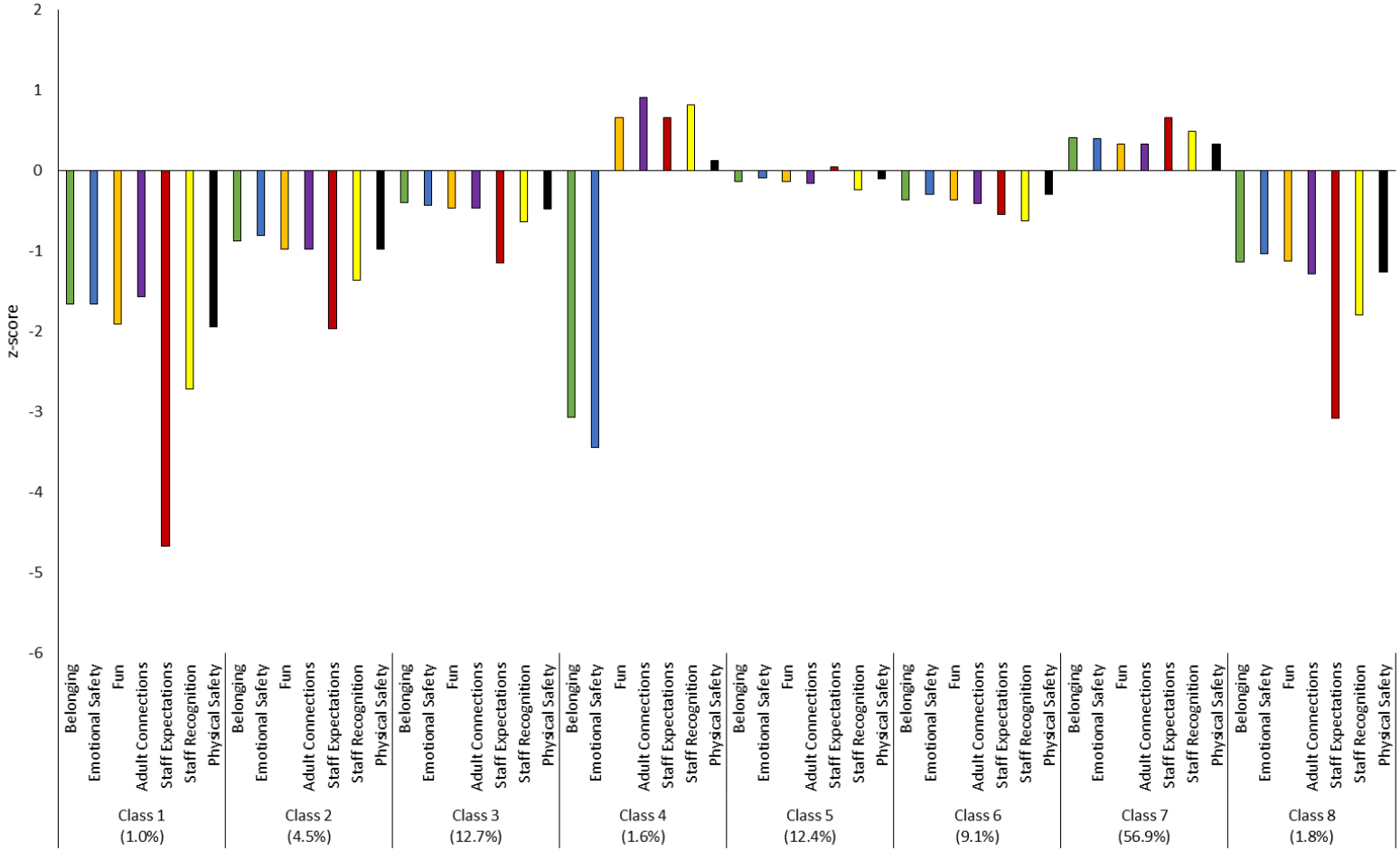

Figure B.12 Eight-profile solution for indicators of program quality among teens (13-18 years old); n = 45,816.

Appendix C: Additional Estimates for Cross-sectional Analyses in 2017

Table C.1 Additional Estimates for the Cross-Sectional Analysis in 2017 for Youth

	Leadership	Conflict Resolution	Teamwork
Model 1	Estimate (SE)	Estimate (SE)	Estimate (SE)
Age	0.00 (0.00)	-0.10 (0.00)*	-0.02 (0.00)*
Gender (Boy)	-0.04 (0.00)*	-0.24 (0.01)*	-0.04 (0.00)*
Attend (Ref.: <1x/week)			
1x/week	0.00 (0.01)	0.00 (0.01)	-0.01 (0.01)
2+x/week	0.00 (0.01)	-0.03 (0.01)	-0.04 (0.01)*
Years Member	0.00 (0.00)	-0.01 (0.00)*	-0.01 (0.00)*
Free/reduced-price lunch	0.00 (0.01)	-0.15 (0.01)*	-0.07 (0.01)*
Residual Variance	0.30 (0.00)*	0.48 (0.00)*	0.40 (0.00)*
Correlations			
Conflict Resolution	0.05 (0.00)*		
Teamwork	0.11 (0.00)*	0.12 (0.00)*	

	Concern for Others
Model 2	Estimate (SE)
Age	-0.02 (0.00)*
Gender (Boy)	-0.06 (0.01)*
Attend (Ref.: <1x/week)	
1x/week	0.00 (0.01)
2+x/week	-0.01 (0.01)
Years Member	-0.01 (0.00)
Free/reduced-price lunch	-0.01 (0.01)
Residual Variance	0.33 (0.00)*

Note: Unstandardized estimates are presented; N = 120,643 for Model 1 and N = 61,238 for Model 2; SE = Standard error; Ref. = Reference group. * $p < .05$ after Bonferroni correction (.05 / 24 = .002 for Model 1, and .05 / 7 = .007 for Model 2).

Table C.2 Additional Estimates for the Cross-Sectional Analysis in 2017 for Teens

	Leadership	Conflict Resolution	Teamwork
Model 1	Estimate (SE)	Estimate (SE)	Estimate (SE)
Age	0.02 (0.00)*	0.04 (0.00)*	0.03 (0.00)*
Gender (Boy)	-0.03 (0.01)*	-0.17 (0.01)*	-0.01 (0.01)
Attend (Ref.: <1x/week)			
1x/week	-0.02 (0.01)	-0.02 (0.01)	-0.03 (0.01)
2+x/week	-0.03 (0.01)*	-0.02 (0.01)	-0.05 (0.01)*
Years Member	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Free/reduced-price lunch	0.00 (0.01)	-0.12 (0.01)*	-0.06 (0.01)*
Residual Variance	0.29 (0.00)*	0.38 (0.00)*	0.34 (0.00)*
Correlations			
Conflict Resolution	0.03 (0.00)*		
Teamwork	0.09 (0.00)*	0.08 (0.00)*	

	Concern for Others
Model 2	Estimate (SE)
Age	0.02 (0.00)*
Gender (Boy)	-0.03 (0.01)*
Attend (Ref.: <1x/week)	
1x/week	-0.02 (0.02)
2+x/week	-0.04 (0.02)*
Years Member	0.00 (0.00)
Free/reduced-price lunch	-0.02 (0.01)
Residual Variance	0.31 (0.01)*

Note: Unstandardized estimates are presented; N = 45,816 for Model 1 and N = 16,767 for Model 2; SE = Standard error; Ref. = Reference group. * $p < .05$ after Bonferroni correction (.05 / 24 = .002 for Model 1, and .05 / 7 = .007 for Model 2).

Appendix D: Longitudinal Analysis Results without Sample Weights

Table D.1 Results of the Longitudinal Analysis without Sample Weights for Youth

Variable	(1) Above Average Profile	(2) Low Belonging and ES Profile	(3) Below Average Profile	(4) Very Below Average Profile	(5) Negative Profile	Significant Profile Differences
Model 1						
Leadership	2.39 (0.04)	2.45 (0.05)	2.28 (0.04)	2.28 (0.04)	2.25 (0.05)	1 > 3,4,5 2 > 3,4,5
Conflict Resolution	2.00 (0.05)	2.01 (0.06)	1.97 (0.05)	1.95 (0.05)	1.93 (0.05)	1 > 4
Teamwork	2.76 (0.05)	2.79 (0.06)	2.67 (0.04)	2.61 (0.05)	2.60 (0.06)	1 > 3,4,5 2 > 3,4,5 3 > 4
n	25,991	339	6,407	2,670	511	
Model 2						
Concern For Others	2.49 (0.07)	2.57 (0.08)	2.39 (0.07)	2.38 (0.07)	2.38 (0.09)	1 > 3,4 2 > 3,4
n	10,582	137	2,506	960	182	

Note: Intercepts of the outcome variables (and standard errors in parentheses) are presented. For Model 1, significant differences are at $p < .05$ after Bonferroni correction (.05 / 30 = .002) and $N = 35{,}918$; For Model 2, significant differences are at $p < .05$ after Bonferroni correction (.05 / 10 = .01) and $N = 14{,}367$; ES = Emotional Safety.

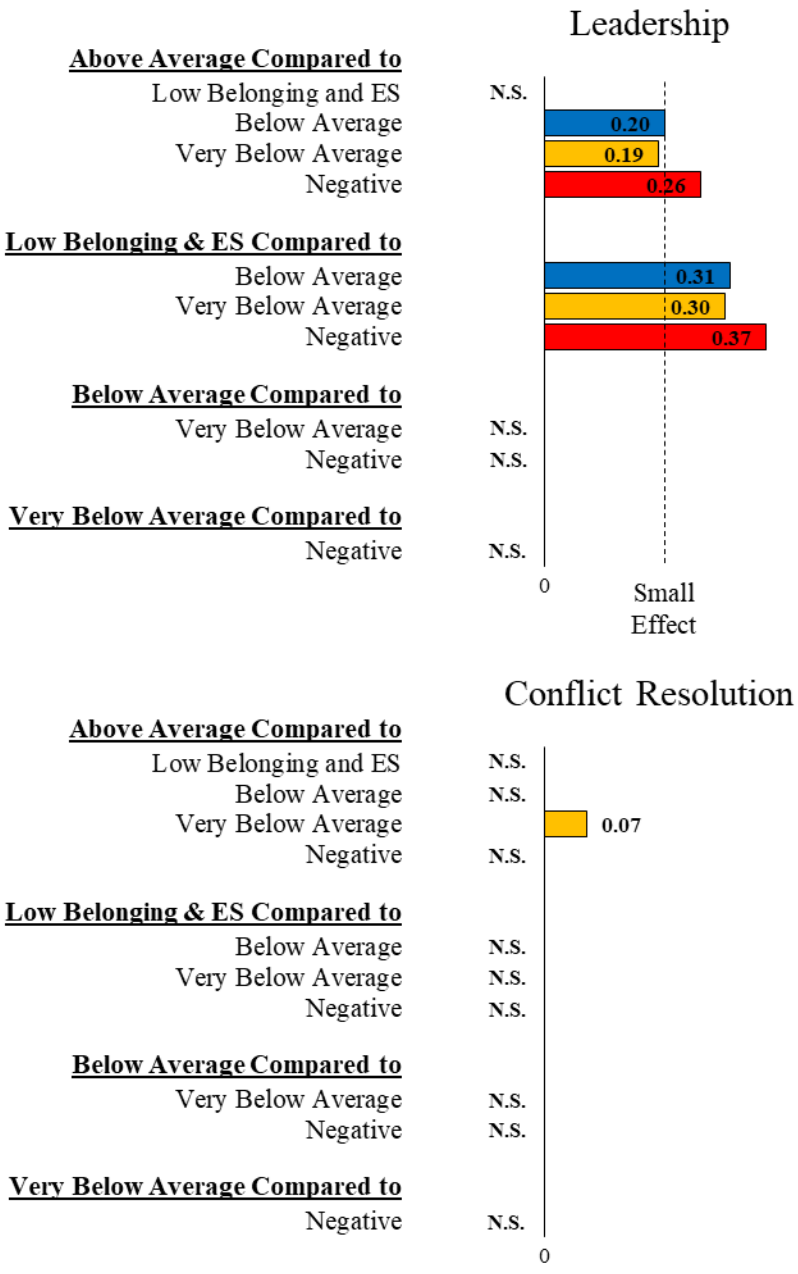

Figure D.1 Effect sizes for profile comparisons in longitudinal analyses without sample weights for youth, part 1. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

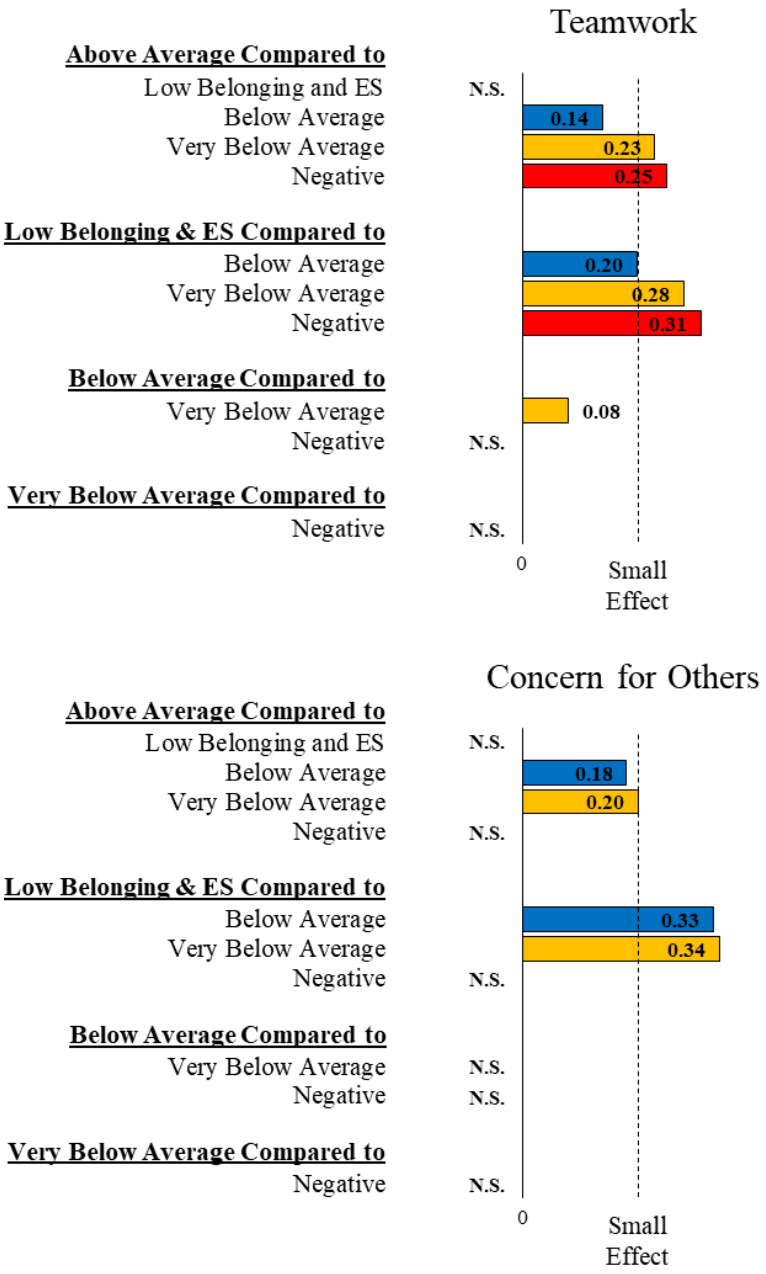

Figure D.2 Effect sizes for profile comparisons in longitudinal analyses without sample weights for youth, part 2. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

Table D.2 Results of the Longitudinal Analysis without Sample Weights for Teens

Variable	(1) Above Average Profile	(2) Low Belonging and ES Profile	(3) Below Average Profile	(4) Very Below Average Profile	(5) Negative Profile	Significant Profile Differences
Model 1						
Leadership	1.82 (0.07)	1.91 (0.08)	1.72 (0.07)	1.77 (0.08)	1.85 (0.12)	1 > 3 2 > 3
Conflict Resolution	0.91 (0.07)	0.87 (0.09)	0.86 (0.08)	0.79 (0.08)	0.78 (0.10)	1 > 3,4
Teamwork	2.34 (0.08)	2.38 (0.09)	2.23 (0.08)	2.15 (0.09)	2.32 (0.12)	1 > 3,4 2 > 4
n	6,855	155	2,164	289	62	
Model 2						
Concern For Others	1.44 (0.15)	1.45 (0.17)	1.33 (0.14)	1.27 (0.16)	1.48 (0.21)	1 > 3
n	1,844	40	528	77	21	

Note: Intercepts of the outcome variables (and standard errors in parentheses) are presented. For Model 1, significant differences are at $p < .05$ after Bonferroni correction (.05 / 30 = .002) and $N = 9{,}525$; For Model 2, significant differences are at $p < .05$ after Bonferroni correction (.05 / 10 = .01) and $N = 2{,}510$; ES = Emotional Safety.

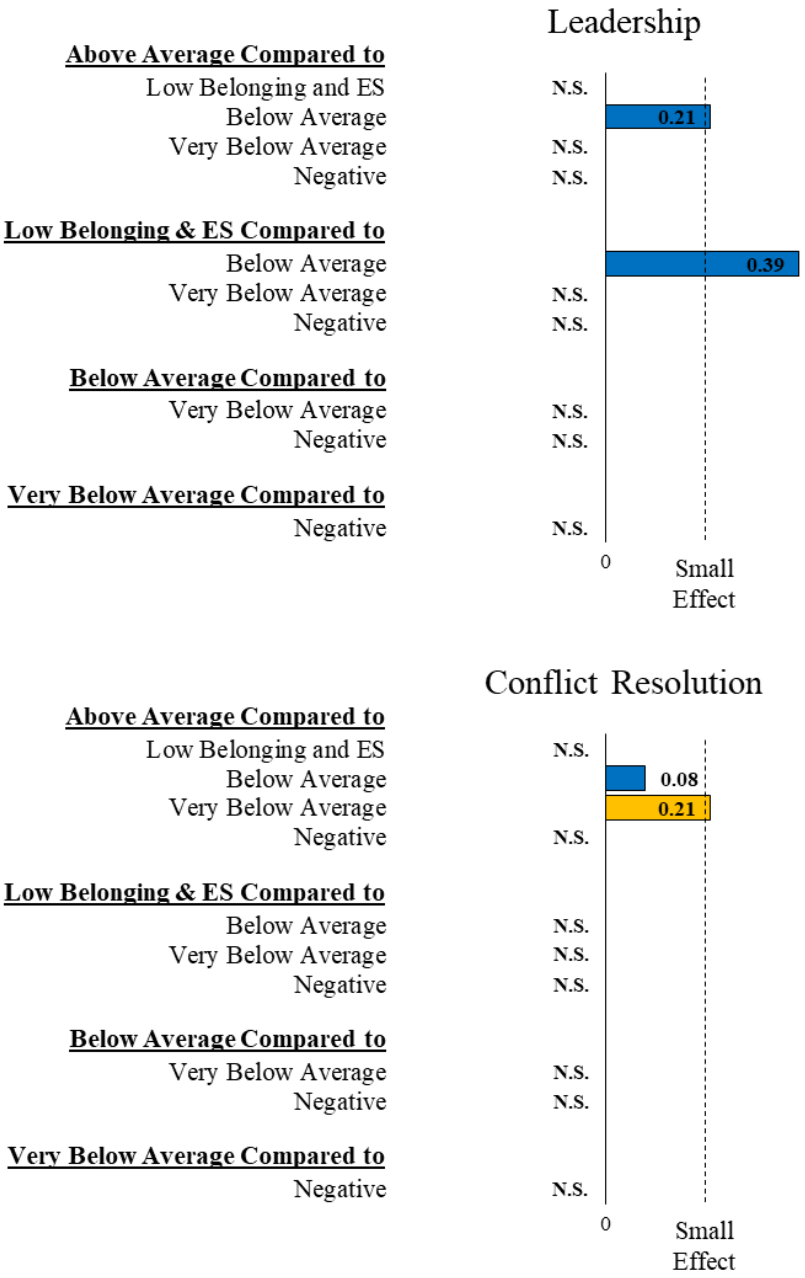

Figure D.3 Effect sizes for profile comparisons in longitudinal analyses without sample weights for teens, part 1. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

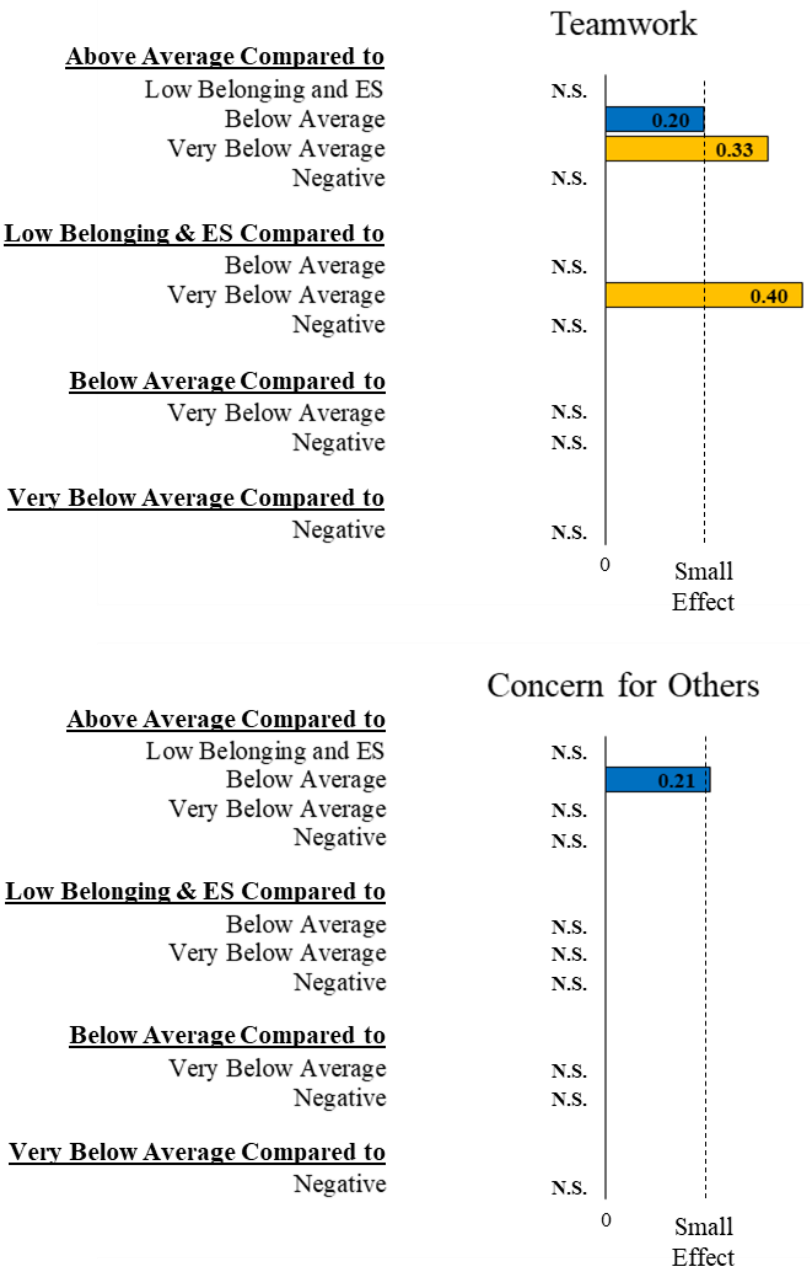

Figure D.4 Effect sizes for profile comparisons in longitudinal analyses without sample weights for teens, part 2. Effect sizes represent Cohen's d (small effect: 0.20 – 0.49). Effect sizes are not shown for comparisons that were not significant after Bonferroni correction. N.S. = Not Significant. ES = Emotional Safety.

Appendix E: Additional Estimates for Longitudinal Analyses

Table E.1 Additional Estimates for the Longitudinal Analysis with Sample Weights for Youth

	Leadership	Conflict Resolution	Teamwork
Model 1	Estimate (SE)	Estimate (SE)	Estimate (SE)
Dependent Var. in 2017	0.29 (0.01)*	0.41 (0.01)*	0.21 (0.01)*
Age	0.00 (0.00)	-0.03 (0.00)*	0.01 (0.00)
Gender (Boy)	-0.05 (0.01)*	-0.11 (0.01)*	-0.04 (0.01)*
Attend (Ref.: <1x/week)			
1x/week	0.01 (0.02)	-0.01 (0.02)	-0.01 (0.02)
2+x/week	-0.01 (0.02)	-0.04 (0.02)	-0.04 (0.02)
Years Member	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Free/reduced-price lunch	0.00 (0.01)	-0.08 (0.01)*	-0.05 (0.01)*
Residual Variance	0.30 (0.00)*	0.39 (0.00)*	0.40 (0.01)*
Correlations			
Conflict Resolution	0.06 (0.00)*		
Teamwork	0.12 (0.00)*	0.12 (0.00)*	

	Concern for Others
Model 2	Estimate (SE)
Dependent Var. in 2017	0.28 (0.01)*
Age	-0.01 (0.01)
Gender (Boy)	-0.06 (0.01)*
Attend (Ref.: <1x/week)	
1x/week	0.05 (0.03)
2+x/week	0.00 (0.03)
Years Member	0.00 (0.00)
Free/reduced-price lunch	-0.01 (0.02)
Residual Variance	0.32 (0.01)*

Note: Unstandardized estimates are presented; N = 35,918 for Model 1 and N = 14,367 for Model 2; SE = Standard error; Ref. = Reference group; Var. = Variable. * $p < .05$ after Bonferroni correction (.05 / 27 = .002 for Model 1, and .05 / 8 = .006 for Model 2).

Table E.2 Additional Estimates for the Longitudinal Analysis with Sample Weights for Teens

	Leadership	Conflict Resolution	Teamwork
Model 1	Estimate (SE)	Estimate (SE)	Estimate (SE)
Dependent Var. in 2017	0.35 (0.01)*	0.44 (0.01)*	0.21 (0.01)*
Age	0.03 (0.01)*	0.05 (0.01)*	0.04 (0.01)*
Gender (Boy)	-0.05 (0.01)*	-0.10 (0.01)*	-0.06 (0.01)*
Attend (Ref.: <1x/week)			
1x/week	0.02 (0.03)	0.02 (0.03)	-0.02 (0.03)
2+x/week	0.00 (0.02)	0.02 (0.02)	-0.02 (0.02)
Years Member	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Free/reduced-price lunch	0.02 (0.02)	-0.06 (0.02)	-0.06 (0.02)*
Residual Variance	0.25 (0.01)*	0.29 (0.01)*	0.32 (0.01)*
Correlations			
Conflict Resolution	0.03 (0.00)*		
Teamwork	0.10 (0.01)*	0.07 (0.00)*	

	Concern for Others
Model 2	Estimate (SE)
Dependent Var. in 2017	0.40 (0.03)*
Age	0.05 (0.01)*
Gender (Boy)	-0.09 (0.02)*
Attend (Ref.: <1x/week)	
1x/week	-0.01 (0.05)
2+x/week	-0.02 (0.04)
Years Member	-0.01 (0.01)
Free/reduced-price lunch	0.01 (0.03)
Residual Variance	0.26 (0.01)*

Note: Unstandardized estimates are presented; N = 9,525 for Model 1 and N = 2,510 for Model 2; SE = Standard error; Ref. = Reference group; Var. = Variable. * $p < .05$ after Bonferroni correction (.05 / 27 = .002 for Model 1, and .05 / 8 = .006 for Model 2).

Table E.3 Additional Estimates for the Longitudinal Analysis without Sample Weights for Youth

	Leadership	Conflict Resolution	Teamwork
Model 1	Estimate (SE)	Estimate (SE)	Estimate (SE)
Dependent Var. in 2017	0.29 (0.01)*	0.41 (0.01)*	0.20 (0.01)*
Age	0.00 (0.00)	-0.03 (0.00)*	0.01 (0.00)
Gender (Boy)	-0.05 (0.01)*	-0.12 (0.01)*	-0.04 (0.01)*
Attend (Ref.: <1x/week)			
1x/week	0.01 (0.02)	0.00 (0.02)	0.00 (0.02)
2+x/week	-0.01 (0.02)	-0.04 (0.02)	-0.03 (0.02)
Years Member	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)
Free/reduced-price lunch	0.00 (0.01)	-0.08 (0.01)*	-0.05 (0.01)*
Residual Variance	0.29 (0.00)*	0.39 (0.00)*	0.39 (0.01)*
Correlations			
Conflict Resolution	0.06 (0.00)*		
Teamwork	0.12 (0.00)*	0.12 (0.00)*	

	Concern for Others
Model 2	Estimate (SE)
Dependent Var. in 2017	0.29 (0.01)*
Age	-0.01 (0.01)
Gender (Boy)	-0.06 (0.01)*
Attend (Ref.: <1x/week)	
1x/week	0.05 (0.03)
2+x/week	0.01 (0.03)
Years Member	0.00 (0.00)
Free/reduced-price lunch	-0.01 (0.01)
Residual Variance	0.31 (0.01)*

Note: Unstandardized estimates are presented; N = 35,918 for Model 1 and N = 14,367 for Model 2; SE = Standard error; Ref. = Reference group; Var. = Variable. * $p < .05$ after Bonferroni correction (.05 / 27 = .002 for Model 1, and .05 / 8 = .006 for Model 2).

Table E.4 Additional Estimates for the Longitudinal Analysis without Sample Weights for Teens

	Leadership	Conflict Resolution	Teamwork
Model 1	Estimate (SE)	Estimate (SE)	Estimate (SE)
Dependent Var. in 2017	0.35 (0.01)*	0.45 (0.01)*	0.22 (0.01)*
Age	0.03 (0.00)*	0.05 (0.01)*	0.04 (0.00)*
Gender (Boy)	-0.05 (0.01)*	-0.10 (0.01)*	-0.05 (0.01)*
Attend (Ref.: <1x/week)			
1x/week	0.02 (0.03)	0.01 (0.03)	-0.01 (0.03)
2+x/week	0.01 (0.02)	0.02 (0.02)	-0.01 (0.03)
Years Member	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)
Free/reduced-price lunch	0.02 (0.01)	-0.06 (0.02)*	-0.06 (0.02)*
Residual Variance	0.25 (0.01)*	0.29 (0.01)*	0.32 (0.01)*
Correlations			
Conflict Resolution	0.03 (0.00)*		
Teamwork	0.10 (0.00)*	0.07 (0.00)*	

	Concern for Others
Model 2	Estimate (SE)
Dependent Var. in 2017	0.39 (0.02)*
Age	0.05 (0.01)*
Gender (Boy)	-0.08 (0.02)*
Attend (Ref.: <1x/week)	
1x/week	0.00 (0.05)
2+x/week	-0.01 (0.04)
Years Member	0.00 (0.00)
Free/reduced-price lunch	0.01 (0.03)
Residual Variance	0.27 (0.01)*

Note: Unstandardized estimates are presented; N = 9,525 for Model 1 and N = 2,510 for Model 2; SE = Standard error; Ref. = Reference group; Var. = Variable. * $p < .05$ after Bonferroni correction (.05 / 27 = .002 for Model 1, and .05 / 8 = .006 for Model 2).