



8-2017

A Systematic Reveiw of Student Self-Report Instruments That Assess Student-Teacher Relationships

Kate L. Phillippo

Loyola University Chicago, kphillippo@luc.edu

Jerusha Conner

Villanova University

Shannon Davidson

Stanford University

Denise Pope

Stanford University

Author Manuscript

This is a pre-publication author manuscript of the final, published article.

Recommended Citation

Phillippo, Kate L.; Conner, Jerusha; Davidson, Shannon; and Pope, Denise. A Systematic Reveiw of Student Self-Report Instruments That Assess Student-Teacher Relationships. *Teachers College Record*, 119, 8: 1-42, 2017. Retrieved from Loyola eCommons, School of Education: Faculty Publications and Other Works,

This Article is brought to you for free and open access by the Faculty Publications at Loyola eCommons. It has been accepted for inclusion in School of Education: Faculty Publications and Other Works by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](https://creativecommons.org/licenses/by-nc-nd/3.0/).

© Teachers College, Columbia University 2017

A Systematic Review of Student Self-Report Instruments That Assess Student-Teacher Relationships

KATE L. PHILLIPPO

Loyola University Chicago

JERUSHA CONNER

Villanova University

SHANNON DAVIDSON

Education Northwest

DENISE POPE

Stanford Graduate School of Education

<ARTICLE DESCRIPTION>This article systematically analyzes self-report student survey instruments that assess student-teacher relationships. From its results, implications are drawn for conceptual specification and future survey design.

<BIO>KATE L. PHILLIPPO is an associate professor of Cultural and Educational Policy Studies at Loyola University Chicago's School of Education. Her research interests include student-teacher relationships; student support practice and policy in K–12 school settings; organizational, professional, sociocultural, political, and demographic influences on policy and practice enactment; and student, teacher, and parent experiences of education policy. Her recent publications include *Advisory in Urban High Schools: A Study of Expanded Teacher Roles* (Palgrave Macmillan Series on Urban Education) and “You’re trying to know me”: Students from nondominant groups respond to teacher personalism,” in *The Urban Review*.

JERUSHA CONNER is an associate professor at Villanova University's Department of Education and Counseling. Her research interests include education policy, student voice, and youth organizing. Her recent publications include “Speak Up and Speak Out: Student Voice in American Education Policy” in the *Teachers College Record*, *National Society for the Study of Education Yearbook*, and *Contemporary Youth Activism* (Praeger).

SHANNON DAVIDSON is a researcher at Education Northwest and a Ph.D. candidate in International and Comparative Education at the Stanford Graduate School of Education. Her research interests include teacher-student relationships and adolescent psychological well-being in the context of schooling.

DENISE POPE is a senior lecturer at the Stanford Graduate School of Education. She specializes in student engagement, curriculum studies, qualitative research methods, parent education, and service learning. She is cofounder of Challenge Success, a research and intervention project that provides schools and families the tools they need to increase student health and well-being, engagement with learning, and integrity. Recent publications include *Overloaded and Underprepared: Strategies for Stronger Schools and Healthy, Successful Kids* (Wiley).

Structured Abstract

Background: A large body of survey-based research asserts that the quality and strength of student-teacher relationships (STRs) predict a host of academic and nonacademic outcomes; however, advances in survey design research have led some to question existing survey instruments' psychometric soundness. Concurrently, qualitative research on STRs has identified important developmental and sociocultural variation in the ways students define, understand, and react to relationships with their teachers. The questions raised by survey methodologists, together with the conceptual elaboration of STRs, suggest that survey instruments used to assess STRs are due for a systematic review.

Purpose/Research Questions: This review of survey instruments examines the strengths and shortcomings of existing measures of STRs. Specifically, we ask: *How do student self-report survey instruments assess STRs?* We examined the extent to which these instruments reflect current survey design principles and existing knowledge about how STRs work, particularly for adolescents.

Research Design, Data Collection, and Analysis: A systematic search of peer-reviewed journal articles that (a) focused on North American middle- or high-school students, (b) linked STRs to student outcomes, and (c) used a student-report measure of STRs yielded 66 studies for which we could obtain the full instrument. Instruments were analyzed using a literature-informed protocol and an iterative process that resulted in strong inter-rater agreement. We used tables and matrices to examine patterns, themes, and outliers in our coded data.

Findings: The 66 studies varied considerably with respect to how they operationalized STRs and how they addressed the validity of their instruments. Similar survey items were used to measure different constructs, and constructs with the same names were measured inconsistently across studies. Many instruments were limited by (a) items that included words with ambiguous meanings, (b) inconsistent identification of instruments' focal students and teachers across instruments, and (c) the use of negatively worded items to measure STRs' strength.

Conclusions and Recommendations: If STR research is to meet its promise to guide and inform teachers' efforts to develop and sustain effective relationships with their students, the field needs to properly identify those behaviors that make a difference for different students and those that do not. The next generation of student-report STR survey instruments requires more stringent attention to construct specification and validity, as well as to item generation (specifically, language use), in order to most effectively measure and identify aspects of STRs that affect student performance and well-being.

Executive Summary

BACKGROUND AND OBJECTIVE

Empirical research on student-teacher relationships (STRs) consistently emphasizes their benefit to students. Qualitative and quantitative literature over the last four decades has identified associations between strong STRs and improved academic, social, and developmental outcomes; however, new advances in survey design research have led some to question existing survey instruments' psychometric soundness. Concurrently, qualitative research on STRs has identified important developmental and sociocultural variation in the ways students define, understand, and respond to relationships with their teachers. Together, these developments suggest that self-report survey instruments used to assess STRs are due for a thorough review.

This article's primary objective is to advance STR research by systematically reviewing a comprehensive, multidisciplinary set of survey instruments administered to middle- and high-school students. We posed our research question—*How do student self-report survey instruments assess STRs?*—with three foci in mind: We set out to learn the extent to which STR instruments: aligned with current survey research methodology principles; reflected common conceptualizations of how STRs work (that STRs involve both student and teacher, serve as a source of social capital or instrumental benefit, and are shaped by status, power and sociocultural differences between teachers and students); and aligned with adolescents' developmental characteristics and needs.

METHODS

Using electronic search engines and snowball sampling techniques, we conducted a systematic search for peer-reviewed journal articles, applying the following search parameters: survey-based studies that focused on North American middle- or high-school students, linked STRs to student outcomes, and utilized a student-report measure of STRs. Our initial sample of 128 articles was winnowed down to 66 studies for which we could obtain the full instrument, either because it was published with the related journal article, because it was otherwise accessible, or because the author provided it upon our request. Because some of these 66 studies used the same survey instruments, our analysis encompassed 49 unique instruments. We coded these instruments, using a rigorous and iterative process that resulted in strong inter-rater consensus. Following coding, we used tables and matrices to examine patterns, themes, and outliers.

FINDINGS

Descriptive analyses revealed that the 66 included studies varied considerably with respect to their samples, how they addressed the validity of their instruments, and how they operationalized STRs. We found that authors used different names to refer to the STR constructs they studied. Author-named constructs (a term we adopted to indicate constructs as named by our 66 articles' authors in their respective publications) included teacher support (62%), connectedness or connections between students and teachers (20%), teacher caring for students (6%), the general relationship between students and teachers (20%), and a smattering of other constructs that appeared less frequently, such as teachers' respect and liking for students, students' attachment to teachers, and teachers' support for student autonomy (8% in total). We found that 45% did not address the validity of the student-teacher instruments or items used. Of the studies that did address validity, only 36% provided a thorough explanation of the validity measures taken.

Our analysis of the 49 instruments used in these studies to assess STRs and the individual survey items that comprise them generated two sets of emergent findings. First, we found that instruments did not fully assess STRs according to survey methodology principles or common conceptualizations of how STRs

work for adolescents, often resulting in problems with construct specification and validity among the instruments we analyzed. Specifically, we found that similar survey items were used to measure different author-named constructs and that constructs with the same names were measured inconsistently across studies. Second, we found that survey-item generation was hindered by ineffective wording choices, as illustrated by (a) the inclusion of words with ambiguous meanings, (b) inconsistent identification of focal students and teachers, and (c) the use of negatively worded items, a practice that goes against survey-research methodologists' recommendations, particularly for adolescent respondents.

IMPLICATIONS

Our review highlights common areas assessed across the particular phenomena studied (e.g., the ubiquity of items coded as teacher caring or teacher support), suggesting characteristics of a multidimensional STR construct. We also note the variability of phenomena measured across instruments. This diversity illustrates how scholars have studied STRs using an array of approaches, which paints a multifaceted portrait of STRs. However, the heterogeneity of STR survey instruments' content and design—particularly where these qualities stray from survey-research methodologists' recommendations—creates a risk of inaccurate measurement of STRs and may therefore limit researchers' ability to further specify how to promote these critically important relationships. Therefore, we conclude that the next generation of student-report STR survey instruments requires more stringent attention to construct specification and validity and item generation (specifically, language use) in order to most fully measure aspects of STRs that matter to students.

Fuller, more explicit specification of an STR construct is necessary to move research in this field forward. We see this study's results as a step toward that end. First, the consistent identification of STR characteristics such as teacher caring, respect, and fairness across instruments suggests the presence of key characteristics of STRs and also suggests that STRs are multidimensional, consisting of multiple phenomena that represent the actions and perceptions of both students and teachers. These findings stand to extend common conceptualizations of STRs that we found in the broad body of STR research literature—that STRs occur between students and teachers; represent a source of social capital that benefits students; and are shaped by status, power, and sociocultural differences between students and teachers. While pursuing the development of a multidimensional STR construct, though, it will be important to narrow down and specify the phenomena considered to contribute to strong STRs. The establishment of convergent and discriminant validity of an STR construct, phenomena that the construct both includes and excludes, is critical to researchers' efforts to more clearly define and understand the STR itself. From the 40 separate survey item types we identified, we believe that it is possible and critical to identify key phenomena to assess so that future STR survey researchers can more precisely gauge students' understanding, experiences and perceptions of STRs.

The time is right for a second generation of student self-report STR survey instruments that are more developmentally and socioculturally attuned, specific, consistent, and actionable. Such instruments would enable educators and policymakers to better understand teachers' contributions to strong STRs. The knowledge these instruments produce could, in turn, inform the design of preservice learning experiences, professional development opportunities, and even performance evaluation criteria for teachers. Ultimately, a more robust understanding of STRs requires a critical interrogation of what it is that we think we know about STRs and greater precision in the future measurement of this complex phenomenon. While this charge may seem steep, our findings reveal a clear pathway by which research can help teachers to become even more informed, supportive, and effective and to promote classrooms in which students thrive.

<MAIN TEXT STARTS HERE>

<TXT>Empirical research on student-teacher relationships (STRs) consistently emphasizes their benefit to students. Over the last four decades, qualitative and quantitative literature has identified associations between strong STRs and improved academic, social, and developmental outcomes, including deeper academic engagement (e.g., Fredricks, Blumenfeld, & Paris, 2004; Murray & Zvoch, 2011), increased academic achievement (e.g., Lewis et al., 2012), resiliency in the face of trying life circumstances (e.g., Werner & Smith, 1982, 2001), and lower incidences of health-risk behaviors such as smoking and alcohol use (e.g., Darwich, Hymel, & Waterhouse, 2012; Resnick et al., 1997). Students who report having a strong, supportive relationship with a teacher are also less likely to drop out of school (Tuck, 2012). Taken together, this body of research strongly suggests that teachers can contribute to positive student outcomes, depending on the quality and types of relationships they have with their students.

Given findings that pair STRs with a wide range of desirable student outcomes, it is unsurprising that STR research has flourished. Survey-based research on STRs, a prominent form of STR research to date, was initially fueled by the inclusion of STR-relevant survey items in two major national longitudinal surveys: the National Longitudinal Study of Adolescent Health (AddHealth), first administered in 1994, and the National Education Longitudinal Study (NELS), first administered in 1988. At the same time, qualitative research and conceptual elaboration upon the nature and contextual meanings of STRs (e.g., Stanton-Salazar & Dornbusch, 1995; Valenzuela, 1999) have enriched the field's understanding of STRs. Considerable work has examined what students want from their teachers (Cushman, 2003; Poplin & Weeres, 1994; Wilson & Corbett, 2001; Zenkov, 2009), whereas other researchers have explored how students themselves define a supportive or caring teacher (Jones & Yonezawa, 2002; Rubin, 2003; Suldo et al., 2009; Theoharis, 2010). This expansion of STR research has occurred at the same time that internet survey technology development has contributed to a proliferation of survey research in general (Dillman, Smyth, & Christian, 2009), although some have questioned recent survey instruments' psychometric soundness. Alwin (2010), for example, cautioned that "we often assume too much about survey data" (p. 405) and what it can say about phenomena under investigation. The National Council on Measurement in Education (2015) calls upon school-based researchers to "take appropriate steps to minimize potential sources of invalidity in the research," reinforcing the importance of paying attention to survey-instrument design. The concurrent expansion of STR research and the tensions that have surfaced in survey-methodology research together suggest that survey instruments used to assess STRs are due for a systematic review.

In particular, student-report survey instruments administered to adolescents merit further analysis. Although STR research spans the P-12 range, self-report instruments are frequently used with middle- and high-school students. STR research with adolescents suggests that their perspectives and desires for relationships with their teachers are often misunderstood or muted (Phillippo, 2012; Rolón-Dow, 2005). These findings dovetail with other research on young people's experiences of schooling and school reform (Alonso, Anderson, Su, & Theoharis, 2009; Tuck, 2012), which indicates a lack of youth voice in educational research and policymaking. Gable and Wolf (1993) noted the importance of instruments' attunement to adolescents' cognitive capacities; we extend this notion to call for instruments' attunement to adolescents' broader developmental needs and qualities.

This paper's primary objective, then, is to advance STR research by systematically reviewing a comprehensive, multidisciplinary set of survey instruments administered to middle- and high-school students. We begin with a review of relevant principles of survey research methodology. We continue by providing an overview of extant knowledge about STRs, identifying common conceptualizations of how STRs work and how STRs match up with adolescent developmental traits and needs. We pose our research question—*How do student self-report survey instruments assess STRs?*—with these foci in mind: survey research methodology principles, common conceptualizations of how STRs work, and STRs' developmental match with adolescents. We then describe the procedures used to collect and analyze the content of student self-report STR surveys. We next discuss our findings, focusing on matters of STR measures' construct validity and survey-item generation. Finally, we consider these findings'

implications for subsequent STR research, including the refinement of student-report instruments used to measure these relationships and the potential for STR surveys to most effectively inform educator practice.

<A> PRINCIPLES OF SURVEY RESEARCH METHODOLOGY

<TXT>We begin with a consideration of key principles for the design and use of self-report survey instruments. Construct specification serves as the foundation for the development of sound survey instruments, which begins with the precise specification of the construct to be studied via thorough review of existing knowledge of the construct and its related domains (Bohrstedt, 2010; Gehlbach & Brinkworth, 2011). A key component of construct specification is establishing whether a construct is unidimensional or multidimensional, comprised of multiple facets (McKenzie, Wood, Kotecki, Clark, & Brey, 1999; Netemeyer, Bearden, & Sharma, 2003). The specification of key dimensions determines what, specifically, survey instruments will gauge. From this basis, survey instrument developers are positioned to ensure construct validity, which Netemeyer et al. (2003) described as “the extent to which an operational measure truly reflects the concept being investigated” (p. 71). The consistent measurement of a construct and the distinction of one construct from others are also critical. Scholars note the importance of establishing convergent validity—when multiple independent measures of the construct are highly correlated—and discriminant validity, in which “the novel measure does not correlate too highly with measures from which it is supposed to differ” (McKenzie et al., 1999, p. 77; see also Campbell & Fiske, 1959).

Another critical aspect of survey development is the importance of the survey items’ comprehensibility, as Tourangeau and Bradburn (2010, p. 318) wrote: “To answer a question, respondents must first understand what they are being asked.” Tourangeau and Bradburn also called attention to the importance of survey items’ comprehensibility, another critical aspect of survey item development (p. 318). Survey items’ clarity determines their ability to generate accurate, valid, and reliable data. People may need to make extensive inferences to understand specific meanings of survey questions, but room for too much inference on the part of respondents can be problematic (Schwarz, 1999). An item that relies too heavily on respondent inference may lead to answers based on an inaccurate interpretation of that item’s intended meaning. Krosnick and Presser (2010) described as “conventional wisdom” that survey designers should “avoid words with ambiguous meaning” (p.264) by choosing language that all respondents are likely to interpret in the same way; otherwise, the need to draw complex inferences about the language used, as Tourangeau, Rips, and Rasinski (2000) noted, may lead to discrepant interpretations of the survey item and accordingly discrepant responses. This is particularly the case when language is grammatically or logically complex or prone to varied interpretations by respondents (Tourangeau & Bradburn, 2010). In particular, reverse-scored items, which use inverse portrayals of the phenomenon being measured (e.g., measuring student trust of a teacher by asking about students’ mistrust of a teacher), are discouraged (Krosnick & Presser, 2010). Gable and Wolf (1993) noted that the practice of negative item wording is specifically problematic for adolescent respondents, who may find such items cognitively challenging and therefore misunderstand the response options. In addition, Schwarz (1999) called attention to the potentially problematic nature of self-report survey items, stating that “minor changes in question wording, question format or question context can result in major changes in the obtained results” (p. 93). Finally, the breadth of survey items used must be considered. Survey-research scholars note that survey items with too few response options, such as two, could produce information with limited meaning by reducing variability between response options (Gable & Wolf, 1993; Krosnick, 1999).

Survey-research methodologists present instrument validation processes as another important means of strengthening surveys’ ultimate results. One option is the use of psychometrically sound items from existing instruments (McKenzie et al., 1999), but absent this option, methodologists encourage a variety of means to test items’ validity and highlight areas where adjustments are needed. Some scholars (e.g., DeMaio & Rothgeb, 1996; Gehlbach & Brinkworth, 2011) encourage the use of human input, such as expert review of items and various approaches to item pretesting (often called cognitive pretesting), which Groves et al. (2009) noted can take the form of respondent think-alouds about items’ meanings (concurrent with or following survey administration) or respondent interviews that ask for paraphrasing of

items' meaning or definition of key terms used in survey items. Another means of establishing instrument validity is psychometric analysis—such as factor analysis of survey results—that informs the inclusion or exclusion of items in subsequent administration of an instrument or in final data analysis (Bohrstedt, 2010). In keeping with this research from survey methodologists, which calls for a deep knowledge of the concepts that survey instruments measure, we continue our review with a discussion of the literature pertaining to STRs.

<A> EXTANT KNOWLEDGE ABOUT STUDENT-TEACHER RELATIONSHIPS

<TXT> Knowledge about STRs provides additional vital information that frames our systematic review. We consider two bodies of knowledge that pertain to the measurement of relationships between students and teachers, in order to draw connections between the broad, diverse body of knowledge about STRs and the survey instruments that strive to assess and describe them. We begin by synthesizing common conceptualizations about how these relationships work. Next, we review and summarize knowledge about adolescents' particular needs for and experiences of STRs.

 COMMON CONCEPTUALIZATIONS OF HOW STUDENT-TEACHER RELATIONSHIPS WORK

<TXT>The literature on STRs engages a variety of theoretical perspectives, including self-system processes (Klem and Connell, 2004), ecological models (Woolley, Kol, & Bowen, 2009), and risk and resiliency (Bowen, Richman, Brewster, & Brown, 1998). This literature, however, does not draw from a consistent, overarching theory of STRs, nor does it identify a central, widely understood STR construct. Across this body of research, however, we note the presence of three general conceptualizations of how STRs operate: that STRs occur between students and teachers, that STRs represent a source of social capital for students, and that STRs are shaped by status, power, and sociocultural differences between students and teachers.

<C> *STRs Occur Between Students and Teachers*

<TXT>In most descriptions of what STRs entail, students and teachers interact to form a working relationship that involves both parties' perceptions of each other and subsequent actions based on those perceptions (Brinkworth, McIntyre, Harris, & Gehlbach, 2015). STRs are not, then, a "deliverable" that teachers unilaterally execute. Descriptions of the nature and impact of these mutual perceptions vary substantially. Pianta (1999) defined relationships between students and teachers as characterized by degrees of closeness, conflict, and dependence, and Murray (2009) framed them as characterized by warmth, trust, involvement, and expectations. A number of studies have identified particular qualities of STRs, such as student-teacher attachment (Hallinan, 2008; Learner & Kruger, 1997), teacher acceptance of students (Khan, Haynes, Armstrong, & Rohner, 2010), and teacher understanding of students (Cooper & Miness, 2014). Noddings (2005) popularized the notion of teachers engaging in caring relationships with their students. While Noddings emphasized the reciprocal nature of caring between teachers and students, others (e.g., Blustein, 1991; Mayeroff, 1971) have asserted that caring relationships between two parties are not necessarily reciprocal. The matter of reciprocity is clearly an important factor in STRs, even as students' and teachers' contributions are neither uniform nor precisely understood.

<C> *STRs Represent a Source of Social Capital for Students*

<TXT>Scholars concerned with STRs often describe them as a form of social capital that can convert into academic and social benefits for students, such as school engagement and persistence to graduation (e.g., Croninger & Lee, 2001; Woolley et al., 2009). Exemplifying this characterization, Crosnoe, Johnson, and Elder (2004) referred to "social capital that flows through affective bonds with teachers" (p. 62) in their analysis of the relationship between STRs and student achievement across racial and ethnic groups and in different school settings. Similarly, Stanton-Salazar (2011) and Stanton-Salazar et al. (1995) framed strong STRs as a vehicle for institutional support, in which teachers and members of teachers' social networks can potentially help students navigate unfamiliar bureaucratic systems and cultural practices, connect to learning and work opportunities, and acquire social and academic guidance.

While not all descriptions of STRs refer explicitly to social capital, many stress these relationships' instrumental value for students. Teachers' support of students—whether academic or affective—is often associated with desirable outcomes such as academic achievement, persistence and resiliency in the face of trying life circumstances such as poverty, family disruption, mental-health difficulties, and school closure (e.g., Conner, Miles, & Pope, 2014; Davis & Lease, 2007; Erickson, McDonald, & Elder, 2009; Green, Rhodes, Hirsch, Suarez-Orozco, & Camic, 2008; Gwynne & de La Torre, 2009; Hamre & Pianta, 2005; Lee & Smith, 1999; Masten & Tellegen, 2012; Olsson, 2009). Across these characterizations, students are portrayed as the beneficiaries of capital provided by their teachers when STRs are strong and positive. The conceptualization of STRs as social capital for students highlights an area of uncertainty in STR research: the very tension about STRs' reciprocity described above. Although Noddings described STRs as, ideally, mutual and reciprocal, STR research tends to focus on the benefits that flow from teachers to students.

<C> *STRs are Shaped by Status, Power, and Sociocultural Differences Between Students and Teachers*

<TXT> Situated in schools, STRs are inevitably shaped by within-school status differences between students and teachers. Additionally, the increasing diversity of the U.S. student population, paired with the persistent predominance of White women among the U.S. teacher workforce, has contributed to differences in race, ethnicity, immigration status, and gender between students and teachers. Although a notable swath of STR scholarship does not delve deeply into these differences, some recent research does so and consistently calls attention to how they shape STRs. Scholars concerned with teachers' cultural competence have presented STRs as necessary for teachers to bridge these gaps between students, on one hand, and their teachers and schools, on the other (e.g., Gay, 2010; Ladson-Billings, 1995; Nieto, 2010; Ware, 2006). Scholars have also explored how STRs interact with students' and teachers' sociocultural identities. This research critiques color-, culture-, and power-blind theories of teacher caring, raising questions about teachers' efforts to provide care for students when those efforts are informed by deficit-based views of students or by culturally biased judgments about what kinds of care students need (Antrop-González & De Jesús, 2006; Barber, 2002; Cooper & Minness, 2014; Garza, 2009; McIntyre, 1997; Noblit, 1993; Pennington, Brock, & Ndura, 2012; Toshalis 2012). Rolón-Dow (2005), for example, described educators' assumptions that students come from dysfunctional communities or families who do not care about their children's education as normalized racism. This body of research also calls attention to sociocultural variation in how individuals define care, as well as what kind of care students seek at school. Valenzuela (1999) contrasted authentic teacher caring—which involves teachers' demonstration of interest in students, efforts to develop truly reciprocal relationships with them, and a deliberate focus on issues of race, difference, and power—with aesthetic caring, in which students' conformity with majority-culture norms is understood to demonstrate their caring about their own education.

Given the apparent absence of either an overarching theory of STRs or an established STR construct, we feel that the three common themes addressed above—STRs occur between students and teachers, represent social capital that benefits students, and are shaped by status, power, and sociocultural differences between students and teachers—represent some of the best available ways to describe how STRs work for students. While this state signals a lack of explicit conceptual specification, it also serves as a starting point for further conceptual development.

<C> *Adolescents' Experiences of Student-Teacher Relationships*

<TXT> Given the differences between students and teachers that shape STR dynamics, it is essential to consider the unique developmental characteristics and needs that adolescents bring to relationships with their teachers. Research on STRs in preschool and primary-grade settings often uses teacher-report instruments or classroom observation to gauge the quality of STRs (e.g., Burchinal et al., 2008; Hamre and Pianta, 2005; Hughes, Luo, Kwok, & Lloyd, 2008). In contrast, STR research with adolescents tends to solicit student reports. Reflecting this perspective, Reddy, Rhodes, and Mulhall (2003) contended that student perception of the STR is more important than support experiences themselves. Underscoring the role of young people's perceptions of relationships, including STRs, Lynch and Cicchetti (1997) claimed that these perceptions are "important organizers of psychosocial development" (p. 86). STR research with

adolescents has likewise emphasized student perception of teachers' day-to-day behavior toward them (Phillippo, 2012; Valenzuela, 1999) and of teachers' intentional efforts to build STRs (Rolón-Dow, 2005) as keys to whether students ultimately engage in those relationships.

These perspectives on adolescent--teacher relationships dovetail with knowledge about adolescent development. Adolescents experience significant cognitive, psychological, social, and physical changes that reinforce the importance of supportive, developmentally attuned relationships with adults. Their larger social world extends beyond their families (Crosnoe et al., 2004). In U.S. schools, this world tends to include multiple classes led by different teachers. Adolescents navigate this expanded range of relationships with greater autonomy than younger children, in a sense trading dependence on parents for reliance on or closeness with nonfamilial others (Steinberg & Silverberg, 1986). During this time of identity development and separation-individuation from parents, adolescents negotiate their social worlds as individuals and maintain relationships in which individuality and attachment can coexist (Koepke & Denissen, 2012). Adults other than parents can be an important source of support for adolescents at this stage, because they can offer feedback or validation for an adolescent's forming self-concept that may seem more legitimate than either the criticism or the perceived obligatory support parents offer. Indeed, Colarossi and Eccles (2003) contended that adolescents "seek outsiders to confirm a current sense of self in the task of identity formation" (p. 28).

In research that emphasizes student voice and students' perspectives on their school experience, youth have called considerable attention to the importance of their relationships with teachers. For example, Wilson and Corbett (2001) found that urban middle-school students wanted teachers who would push them, discipline them, help them, teach them, and respect them, and that they equated these six actions with "teacher caring" (p. 88). Similarly, teachers' use of a range of techniques to help students understand the material has been associated with increased student perception of teacher support (Anderman, Andrzejewski, & Allen, 2011; Suldo et al., 2009). A recent youth participatory-action study (Tuck, 2012), conducted with older adolescents who have earned or are pursuing a GED in lieu of a high-school diploma, highlighted the salience of STRs in the "pushout" process. Tuck and her student co-researchers found that large numbers of youth who left high school felt disrespected, unwelcome, or unnoticed by their teachers. When students at the Institute for Collaborative Education (a small, alternative public school in New York) were empowered to design, implement, and lead an improvement program for their school, they chose to focus on improving STRs (Student Voice Collaborative, 2011). Youth-led research and reform projects like these show how deeply students care about their relationships with teachers and how these relationships affect them.

We conclude that STRs appear to benefit students and that students have valuable and necessary information to share about how those relationships can be cultivated and how they work. We have highlighted common conceptualizations of STRs, and we recognize the lack of a precise construct specification, which merits further exploration of the extent to which survey instruments incorporate extant knowledge about STRs. We are also reminded of the importance of understanding students' discernment of STRs, particularly in light of current knowledge about both survey development for adolescent respondents and adolescents' relational preferences and needs. This overview illustrates how the measurement of STRs using survey instruments has helped to advance the field's understanding of these relationships. At the same time, it leads us to wonder whether self-report survey instruments might also limit what educators, policymakers, and researchers ultimately understand about relationships between teachers and adolescents. The opportunity to extend and enhance extant knowledge about the measurement of STRs drives this study's research question.

<A> METHOD

 RESEARCH QUESTION AND STUDY DESIGN

<TXT> The following research question guides our study: *How do student self-report survey instruments assess student-teacher relationships?* To explore this question, we analyzed the content of student self-report STR survey instruments (Krippendorff, 2013), focusing on instruments' consistency with survey design principles as well as their reflection of existing knowledge about how STRs work, particularly for

adolescents. To begin, we searched for studies that used student-reported data to link STRs to outcomes among middle- and high-school students in North America. Informed by our overview of extant literature, we developed the following list of search terms: student-teacher relationships, teacher-student relationships, teacher support, teacher care, teacher caring, teacher connectedness, school connectedness, student-teacher interaction, and teacher liking. We then used these terms in our search for studies in electronic databases (ERIC, EBSCOhost, Google Scholar, JSTOR, and PsycInfo) and in the reference lists of all studies that we identified. Our search parameters called for peer-reviewed journal articles that focused on North American middle- or high-school students, linked STRs to student outcomes, and utilized a student-report measure of STRs. Because we used these criteria, we did not include studies based solely on qualitative data, nor did we include non-peer-reviewed reports, book chapters, literature reviews, conceptual pieces, unpublished dissertations, or master's theses. We initially identified 128 articles by article keywords, titles, and abstracts. Upon full review of each article, we eliminated articles that did not meet our criteria for inclusion. Further, we only included articles in our sample for which we could evaluate the entire instrument used, including every individual survey item to which students responded and each item's response options. Complete student-report survey instruments were included in some articles. In other cases, we located the instruments from references provided in the articles. In 22 cases, we contacted the lead authors directly to request instruments. Thirteen of these authors shared their instruments, while nine did not reply to multiple attempts to contact them. Our final sample consists of 66 peer-reviewed articles, with a total of 49 separate instruments, as six of the instruments were used multiple times across different articles. The years of publication for the articles in our sample ranged from 1989 to 2013.

DATA ANALYSIS

<TXT> To understand our sample's characteristics, we first created a table that presents key characteristics for each study, including a description of the study's sample and data sources, participants' demographic characteristics, constructs that authors measured, the instrument(s) used, a brief description of survey items, strategies the researchers used to promote validity, and study outcomes (table available from first author). We then compiled descriptive statistics for the sample (reported in Table 1).

Table 1. Descriptive Statistics of Articles in Sample

[<<insert Table 1 here>>](#)

Next, we analyzed each instrument using a six-section protocol, based on our review of survey methodology principles, conceptualizations of how STRs work, and STR-related adolescent development. Specifically, at the instrument level, we coded for (a) survey question structure (e.g., "how many teachers at your school . . . , " "how often does your teacher. . . "), with eight types of structure identified and some instruments containing more than one type; (b) the respondent identified in the survey instrument (oneself or other students at the participant's school); and (c) the teacher referent named by the instrument (e.g., "math teacher," "all my teachers"), with seven different varieties identified. At the item level, we coded for (d) student perceptions of or feelings about the STR identified (e.g., student trust in teacher, student's ease with talking to teacher) and (e) teacher actions named by the instrument (e.g., teacher respects student, teacher cares about student). Finally, for each instrument we tabulated (f) the number of items that did not exclusively pertain to STRs (e.g., "How close do you feel to students at this school?", "Students have to watch what they say in class").

With this protocol, we coded each instrument, with many instruments and individual items coded multiple times. For example, some individual survey items assessed both teacher caring and high expectations, such as the item used by Muller (2001): "My teachers care about me and expect me to succeed in school." This item was coded twice in section E of our protocol—once for the teacher action of caring, and once for the teacher action of holding positive expectations. Other items met criteria for codes in both protocol sections D and E, such as the item, "I have adults at school I can talk to, who care about my feelings and what happens to me" (Daly, Shin, Thakral, Selders, & Vera, 2009). In this case, we coded the item according to protocol section D (student believes she has an adult to talk to) and twice as part of

protocol section E (the teacher actions of caring about the student's feelings and of caring about what happens to the student).

To strengthen the reliability of our analytic process, our research team addressed intercoder agreement through a multistage process that promoted the development of shared understandings and subsequent analysis across all four authors (Saldaña, 2013). First, teams of two separately coded and then compared coding results for a small, randomly chosen subset of five instruments per pair. After discussing and resolving discrepancies in those pairs (reaching complete consensus for the analysis of each instrument), the entire research team met to review all coding decisions and then revised the protocol, adding new codes and revising others. We collapsed several infrequently used codes into aggregate codes. Using the refined protocol, we coded a subset of 10 instruments following the same protocol, which included the initial subset of instruments plus five more. Once each team achieved full consensus in applying the codes, the two teams of two researchers divided and coded all remaining instruments. All final coding was reviewed by both members of each coding team, and all four authors then collaboratively reviewed the coded data to ensure consistency and agreement on final coding decisions. Throughout this phase of data analysis, the authors met regularly to discuss and develop themes emerging from the data and to form and evaluate propositions about findings.

After coding concluded, we constructed tables and matrices to identify patterns in our data. We identified the 10 most frequently appearing codes and also identified codes used only once, then considered these outliers' relationship to the broader base of STR literature. We then analyzed areas of convergence and divergence among instruments with regard to the phenomena that they measured. We also examined patterns in survey-item language, including how survey items identified students and teachers, and considered how our findings corresponded with standards for survey design and construction.

<A>RESULTS

<TXT>In response to our research question—*How do student self-report survey instruments assess student-teacher relationships?*—our findings relate to our study's three focal points: survey-research methodology principles, common conceptualizations of how STRs work, and STRs' match with adolescent characteristics and needs. We found that the field's incomplete specification of a common STR construct led to divergent approaches to survey research about those relationships. Additionally, the language used to generate items frequently failed to accurately and clearly describe the phenomena involved in the relationships being studied and emphasized teacher behavior rather than student behavior or contributions to the relationship. Below, we first discuss descriptive analysis results and then address our emergent findings.

DESCRIPTIVE ANALYSIS RESULTS

<TXT>The 66 studies included in this review varied substantially in scope and purpose (see Table 1). Slightly more than half (52%) limited the sample to middle-school students, and a slight majority (59%) examined the perspectives of students attending public schools. Nearly one third (32%) focused on urban schools, 12% focused on suburban students, and 8% focused on rural students. While most studies included various demographic control variables, the number of studies specifically examining group differences by demographic variation was much smaller. Forty-two percent of our sample's studies directly addressed differences in outcome by student gender, 15% considered differences by student race or ethnicity, and only three studies examined student socioeconomic status. Given the recent focus on sociocultural differences in student experiences and perceptions of STRs, we were surprised by this sample's limited emphasis on race, ethnicity, and SES.

Equally important as questions of generalizability are questions about instrument validity and the steps that authors took to assure their readers of their measures' psychometric soundness. The 66 studies varied considerably in this regard as well. Forty five percent did not address the validity of the student-teacher instruments or items used. Of the 36 studies that did mention validity, only 36% provided a thorough explanation of the validity measures taken. For example, Demaray and Malecki (2002) described their use of the Child and Adolescent Social Support Scale in terms of its convergent validity,

which was assessed using a sample derived from the database in the study and correlated with the validated instrument, the Social Support Scale for Children (Harter, 1985). Of the remaining 23 studies that mentioned validity, ten reported conducting analyses in the construction of the instrument to ensure its validity and reliability but did not describe those analyses or their results. Sixteen studies referenced a separate article that addressed the validity of the measure. One study simply mentioned the validity of the measure without providing any further support or explanation.

Limited attention to matters of validity occurred in both original and adapted instruments. Of our sample's 66 studies, 26% used original instruments, 68% selected items from existing instruments, and another 21% modified or adapted extant instruments. These categories are not mutually exclusive because some studies used more than one instrument, for example combining original instruments with existing instruments. Only 29% of the studies that adapted existing instruments explained how or why they were modified. Similarly, only 17% of the studies that used original instruments discussed how the instruments were developed and validated. In sum, we found wide variability in how, if at all, authors addressed the validity of the instruments they used.

Another key point of difference across the studies we reviewed was the terminology used. All 66 studies included at least one instrument designed to measure some aspect of STRs; however, these instruments have different names, depending on the construct that each article's author intended to measure. *Author-named constructs* (a term we adopted to indicate constructs as named by our 66 articles' authors in their respective publications) included teacher support (62%), connectedness or connections between students and teachers (20%), teacher caring for students (6%), the general relationship between students and teachers (20%), and a smattering of other constructs that appeared less frequently, such as teacher respect and liking for students, students' attachment to teachers, and teachers' support for student autonomy (8% in total). Twenty-three percent of studies measured more than one construct. Instruments' mean number of items was eight. Three instruments consisted of a single item, and four instruments contained more than 20.

EMERGENT FINDINGS

<TXT> Our analysis of these instruments and the individual survey items that comprise them generated two sets of emergent findings. First, we found that the instruments did not fully assess STRs according to survey-methodology principles or common conceptualizations of how STRs work for adolescents, often resulting in problems with construct specification and validity among the instruments we analyzed. Specifically, we found that similar survey items were used to measure different author-named constructs and that constructs with the same names were measured inconsistently across studies. Second, we found that survey item generation was hindered by ineffective wording choices, as illustrated by (a) the inclusion of words with ambiguous meanings; (b) inconsistent identification of focal students and teachers; and (c) the use of negatively worded items, a practice that goes against survey-research methodologists' recommendations, particularly for adolescent respondents.

<C>*Construct Specification and Validity Problems*

<TXT> Our sample's instruments measured STRs in ways that reflected inconsistent specification and measurement of STRs as a construct. We identified a constellation of related ideas that referred to STRs, but the authors in our sample used these ideas in highly varying ways that reflected the absence of an overarching STR construct that could have structured instrument and item construction. We noted authors' use of similar language and survey items to measure different author-named constructs, inconsistent measurement of constructs with the same names, and inadequate measurement of STRs according to survey-research methodology principles and common conceptualizations of STRs.

<C>*Similar Survey Items Measured Different Author-Named Constructs*

<TXT> Specific wording of certain items appeared consistently in our analysis of the items that made up different author-named constructs. Most prominently, individual survey items that gauged teachers' care for students appeared not only in all of the instruments used to measure the author-named construct of

teacher caring, but also in 53.6% of the instruments measuring teacher support, 80% of instruments measuring connectedness, 55.6% of instruments measuring general relationships between students and teachers, and 14.3% of the instruments measuring other phenomena (see Table 2).¹ After the items that measured teacher care, the most frequently appearing survey items across survey instruments concerned teachers listening to students (57.1%), treating students fairly (32.7%), and having time for students (32.7%). Slightly more than one quarter (28.6%) of the instruments included items pertaining to teachers treating students with respect or honesty, giving help or advice to students (26.5%), and having time for students (24.5%). These findings suggest that certain kinds of survey items are more central across our sample, which could inform the definition of an overarching STR construct and the establishment of that construct's convergent validity (indicated by similar phenomena being measured by multiple instruments). However, that work was not made explicit in our sample's articles. Of the 10 categories of survey items most frequently used within our sample (as described in Table 2), all were represented across at least four author-named construct types, and six were represented across all five types.²

Table 2. Categories of Survey Items Most Frequently Used to Assess Student-Teacher Relationships, Sorted by Author-Named Construct

<<insert Table 2 here>>

The representation of teachers' respect for students, measured across all five author-named construct types in Table 2, provides a clear example of the varied ways in which authors incorporated this phenomenon into their measurement of STRs. The School Success Profile (Bowen & Richman, 2008, used in six studies in our sample), measured teacher support with a multi-item scale including the item, "I am respected and appreciated by my teachers." Respect also surfaces in Eisenberg et al.'s (2007) measure of teacher caring: "How much do you feel teachers or other adults at school show respect for the students?" The item "In this school, teachers treat students with respect" was also part of Roeser and Urdu's (1996) instrument for gauging STRs. And finally, Nasir, Jones, and McLaughlin (2011) used the item "How much do the teachers respect you?" to gauge school connectedness. Because the items used to comprise these instruments overlapped substantially, the constructs themselves are conceptually entangled with one another. This entanglement suggests that convergent and discriminant validity, or a sense of what phenomena might be included in and excluded from an overarching STR construct, have yet to be established.

<C>Author-Named Constructs Were Measured Inconsistently Across Studies

<TXT>Even within studies that used the same author-named constructs, the items that made up those constructs varied. Each of the author-named constructs we identified was operationalized in a unique way by each instrument, with the exception of scholars using the same survey instruments and survey items across a series of studies. For example, Muller (2001) used a seven-item measure of student perceptions of *teacher caring* that included items such as "When I work hard, teachers praise my effort" and "The teaching is good." Meanwhile, Lewis et al. (2012), also studying teacher caring, used a three-item instrument that assessed whether each respondent's math teacher "takes a personal interest in students," "cares about how we feel," and "listens to what I have to say." These two studies both attempted to measure teacher caring, but they did so using very different survey items. Similarly, the author-named construct of *school connectedness* was measured in various ways across studies. Of the 12 studies that measured connectedness, only half included an item pertaining to how the student and teacher got along, and one third included an item that assessed whether teachers were interested in or paid attention to students. The meaning of these author-named constructs across studies, therefore, remains uncertain, and calls into question these instruments' convergent validity (where the instruments assess the same phenomena). The lack of consistency across instruments also renders it difficult to establish their discriminant validity (the degree to which they do not assess phenomena beyond STRs). Finally, this inconsistency of measurement undermines the instruments' face validity, or ability to measure what their titles suggest that they measure (Netemeyer et al., 2003).

<C>Instruments Did Not Fully Assess STRs According to Common Conceptualizations of How STRs Work for Adolescents

<TXT>Another construct-validity limitation present in our sample of instruments concerns the extent to which they measured STRs as they are currently understood. First, sample instruments often used limited items and response options that thereby created a risk of inadequate specification of STRs. Second, we found among these instruments an inconsistent representation of the common conceptualizations of STRs that we described above (that STRs occur between students and teachers, represent social capital that benefits students, and are shaped by status, power, and sociocultural differences between students and teachers). Instruments neglected common conceptualizations of STRs, included items with an unclear relationship to STRs, and inquired more about teachers than students. Together, these characteristics resulted in limited conceptual coverage.

Most instruments followed methodologists' recommendations to use multiple items and give multiple response options per item, but a sizeable minority did not. Eight instruments (12% of our sample) gauged STRs with three or fewer items, while three used only one item. Six instruments used a two-point response scale, all but one of which provided "true/false" response options. The use of two-point scales is criticized by survey methodologists as overly restrictive (Gable & Wolf, 1993) and as creating the possibility that instruments may "fail to discriminate between respondents with different underlying judgments" (Groves et al., 2009, p. 239).

Additionally, when analyzing material less frequently included in our sample, we identified some items that were relevant to extant STR literature as we have discussed it above, and therefore surprisingly absent from the majority of instruments. Items that reported on teachers' soliciting student voice ($N = 4$), such as "My opinion matters to my math teacher" (Niehaus, Rudasill, & Rakes, 2012), reflect research that suggests that student voice is critical to student engagement in schooling (Taines, 2012; Toshalis & Nakkula, 2012) as well as in STRs themselves (Phillippo, 2012), yet these items appeared quite rarely. Similarly, teacher attunement to students' sociocultural differences and to issues of equity in the classroom received surprisingly limited attention in our sample's instruments but evoked qualitative and practice-based research that suggests that issues of equitable treatment, culturally relevant instruction and culturally attuned STRs are critical to STRs' ultimate strength (e.g., Rolón-Dow, 2005; Valenzuela, 1999).

We also were surprised, given claims about the importance of pedagogical practice to STRs, by the relative absence of survey items that pertained to this dimension. Only two instruments included items pertaining to the quality of the teaching that students experienced, and only one item asked student respondents about their teachers' specific pedagogical practices ("Teachers at my school make schoolwork interesting," used by Voisin et al., 2005), despite studies that show a relationship between instructional support and students' perceived teacher support (Anderman, Andrzejewski, & Allen, 2011; Suldo et al., 2009). These items' rare appearance in STR instruments does not suggest that they are insignificant, but rather raises important questions about our sample instruments' connection to the broad, evolving base of knowledge about STRs.

Other infrequent items seemed to represent genuine outliers among our sample. Of the 407 total survey items that we analyzed, 20 remained uncoded because they did not fit our coding structure and appeared so infrequently. A few unusual survey items simply did not resemble others used in our sample or had a questionable relationship to STRs (e.g., "Some kids do/do not have a teacher who treats them like a person" and "I like to see how much I can get away with in class"), but other rare items seemed more thematically related. One cluster of three items concerned students' caring about their relationships with teachers and wanting a particular *quality* of relationship with them. These items—which included "Because I admired this teacher, I tried to act or do things the same way the teacher did," "I care what my teachers think of me," and "I want to be respected by my teachers"—focus on students' valuation of what their teachers think about them, which evokes Noddings's (2005) description of STRs as mutually constitutive rather than flowing one way, from teacher to student. Another pair of items concerned students' awareness of teachers' engagement with other students. Both items—"The teacher has favorites/likes other students better than me" and "I felt jealous when this teacher spent time with or helped other students"—get at what may be a critical issue: students' reactions to teachers' relationships with other students. Since secondary-school class sizes may often exceed 30 students, the matter of how or whether teachers balance their relationships with different students merits attention. In these cases,

the infrequent use of these survey items does not suggest those items' insignificance so much as it suggests areas for further attention.

Alongside the omission of important aspects of STRs from sample instruments, many instruments (27%) contained items that did not solely concern relationships between students and teachers. For example, Gutman and Midgley's (2000) teacher support scale included "Students have to watch what they say in this class." While it could be that student respondents must watch what they say because their teacher would react negatively, it could also be that fellow students in a particular class monitor and critique one another's use of language, which would make this item's relationship to STRs tenuous. Other items with an ambiguous relationship to STRs appeared in instruments used to measure students' school connectedness. Indeed, *school connectedness* presents a particular dilemma for STR measurement because the phenomenon of connectedness itself concerns more than relationships between students and teachers, including student relationships with peers, the school's social climate, and student investment in school (Libbey, 2004). Not surprisingly, then, some connectedness instruments (three of the seven) also included items that did not specifically measure STRs, such as "When I am at my school, I feel free to be who I am" (Niehaus et al., 2012) and "You feel close to people at your school" (Voisin & Nielands, 2010b). This quality does not call into question the validity of school connectedness instruments themselves, but rather the use of connectedness instruments to gauge STRs.

Another area of insufficient conceptual coverage involved instruments' specification of the particular students and teachers involved in STRs. Our sample of self-report survey instruments focused on student perceptions of teacher actions far more than they gauged students' contributions to STRs and in so doing, portrayed teachers as STRs' lead agents. The overwhelming majority of coded items (300) invested the teacher with agency, while far fewer (60) endowed the student with such power. Fourteen of a total 49 instruments (29%) included at least one item in which the student was an agent. These items constituted between 3% and 66% of the instruments, with the exception of one one-item instrument (Tummala-Narra & Sathasivam-Rueckert, 2013). The remaining 71% of instruments consisted of only teacher-as-agent survey items. As a result, sample survey data may inform readers more about teachers' actions toward students than students' actions toward or perceptions of teachers. More rare were student-as-actor items such as "I enjoy talking to the teachers here" (Jiang, Huebner, & Siddall, 2012; Siddall, Huebner, & Jiang, 2013). Among this subset of items, most frequent were those that measured student respondents' perception that they could talk to their teachers ($N = 22$), with four referring specifically to discussions about school issues and six concerning conversations about nonschool issues. Items that included the phrase "get along" were the next most frequently noted student-as-agent items, appearing in 12 instruments. This finding reflects a tension present in extant STR literature, in which many scholars emphasize how STRs flow from teachers to students (e.g., Crosnoe et al., 2004), whereas other research notes the importance of student voice and discernment of teachers to students' experiences of STRs (e.g., Rolón-Dow, 2005). As a result, our findings raise questions about how STR survey instruments can best gauge the students' perceptions, experiences, and actions that ultimately shape these relationships.

<C>Survey-Item Generation: Problematic Language Usage

<TXT>In addition to the construct-related issues described above, we found evidence that survey-item generation was hindered by problematic use of language, which is a critical issue in survey research, particularly with adolescent respondents. Problematic survey language appeared in three areas: the inclusion of words with ambiguous meanings, the inconsistent identification of focal students and teachers, and the use of negatively worded items. These findings strongly suggest a potential for data to be clouded by respondents' divergent understandings of survey items.

<C>Inclusion of Words with Ambiguous Meanings

<TXT>The general language used to describe STRs in many items suggests potential problems with the consistency of respondents' interpretation of those items. Many items that measured the most frequently appearing STR qualities, such as *teacher care* and *respect*, tended toward general language. Items concerned with teacher care tended to refer to care in a general sense, such as De Wit, Karjoja, Rye, and

Shain's (2011) use of "My teachers really care about me." However, other items indicated specific foci of teacher care, such as "Our teachers care how we feel" (Lewis et al., 2012) or "My teachers care about how I do in school" (Murray, 2009). The teacher-care items on five instruments concerned caring about attendance or whether a student shows up at school, but in most cases, the respondent was left to interpret the specific meaning of teacher care. In light of research that highlights differences in students' understanding of and receptiveness to teacher care across different developmental stages and sociocultural groups (Garza, 2009; Phillippo, 2012), this finding raises questions about how student responses to survey items might reflect diverse understandings of those items' meaning. Similarly, items used to measure *respect* used broad language (e.g., "In this school, teachers treat students with respect"), even though this term, too, might be associated with different teacher behaviors or attitudes for different groups of students. Other broad terms from our sample's items included "Students get along well with most teachers," "Most of my teachers treat me fairly," and "How many of your teachers are interested in you as a person?" These terms may mean different things to different students and thus may not necessarily assess specific teacher behavior. More specific items, in contrast, included "My teachers are willing to work with me after school" (Woolley et al., 2009) and "This teacher knows something about me from outside of class (the activities I enjoy, the music I like, etc.)" (Hafen et al., 2011). While broad language may capture a variety of student experiences, its ambiguity leaves readers with less precise information about what really constitutes a caring or supportive relationship in different contexts and for different students.

<C>*Inconsistent Identification of Focal Students and Teacher*

<TXT>Given this study's focus on student self-report survey instruments and the clear power and positional differences between students and teachers, we sought to understand how instruments and items represented each. Many instruments' and items' wording obscured specific actors, with highly variable depictions of both students and teachers. In most cases, instruments asked students to speak for themselves, with items that included *you, I, me, or my* (377 items), such as "I can get extra help from adults in my school if I need it" (Darwich et al., 2012). However, in a substantial number of cases (56 items), instruments asked respondents to speak for other students, depicting student actors as "most students," "all students," or simply "students" (e.g., "Students get along well with most teachers," Konishi, Hymel, Zumbo, & Li, 2010). Twelve instruments included a mix of such items, with some items inquiring about the respondent's personal experience or perspective and other items asking the respondent to generalize to all or most students in their class or school. Three instruments limited themselves to measuring the respondent's perception of STRs for most or all students, not inquiring specifically about the respondent's own experiences.

Our sample also included a range of teacher and adult referents. These ranged widely: a particular teacher (e.g., a math teacher), a particular but underspecified teacher ("my teacher," in secondary schools where students presumably have multiple teachers), the respondent's teachers, most teachers, all teachers, unspecified adults in the school, and other adults in school (a designation that appeared to exclude teachers). Many instruments (18) included a mix of adult referents. For example, Wilson (2004) used a seven-item instrument that includes the following items: "My teachers listen when I have something to say," "I have a teacher who really cares about me," and "When someone breaks the rules, teachers and administrators always take appropriate action." The respondent is asked first to generalize across all of his or her teachers, then to think of a specific teacher, and finally to generalize across all teachers and administrators in the school. The most common phrasing for teacher-specific referents was "your teachers" or "my teachers," used in 31 instruments. Seventeen instruments included items that referenced "most teachers," "all teachers," or "the teachers." Fourteen instruments asked students to reply to the items with a specific teacher in mind, although only nine of these directed the respondent's attention to a particular teacher, such as a homeroom teacher or a language arts teacher. The math teacher was the most common of these, serving as the focus of five instruments. Ten instruments included items that asked about adults in the school, including but not limited to teachers, and three instruments included at least one item that focused on the students' interactions with or perceptions of other adults in the school who were not teachers. The inclusion of items that refer to other adults weakens instruments' ability to assess STRs.

<C>Negative Item Wording

<TXT>Finally, a substantial number of instruments failed to meet survey-research methodologists' recommendations because they included negatively worded items, a particular problem for adolescent respondents. About 80% of the sample's items were positively worded. No instruments used only negatively worded survey items. Some instruments contained only positively worded items (e.g., Wilson, 2004; Zullig, Huebner, & Patton, 2010), whereas others mixed positively and negatively worded items (e.g., Murray, 2009; Ryan & Patrick, 2001). Niehaus et al. (2012), for example, included items such as "In class I often feel 'put down' by my teachers" and "I feel cared about." Interestingly, McCarty, Rhew, Murowchick, McCauley, and Vander Stoep (2012) balanced each of their instruments' items between positive and negative wording (e.g., "Some kids don't have a have a teacher who is *fair* to them [response option provided], BUT other kids *do* have a teacher who cares if they feel bad [separate response option provided]" [emphasis in original]). Voisin and Neilands (2010a) similarly matched some, but not all, items in their teacher support scale by providing items that assessed the same phenomenon using first positive and then negative language (e.g., "My teacher has plenty of time for me," "My teacher doesn't seem to have enough time for me"). Roorda, Koomen, Spilt, and Oort (2011) noted that the absence of weakness or problems from STRs does not amount to a robust relationship, suggesting that the act of reverse-scoring items that measure negative teacher behaviors (e.g., putting down a student) or negative student perceptions of teachers (e.g., treating students differentially) may not provide valid data. Further, Gable and Wolf (1993) described negative item wording as a validity threat, particularly with adolescent respondents, because of the cognitive challenge these items present.

<A>DISCUSSION

<TXT>Our review of student-report STR instruments takes stock of 49 instruments (across 66 peer-reviewed studies) that assess relationships between adolescent students and their teachers. It highlights common areas assessed across the particular phenomena studied (e.g., teacher caring, teacher support), suggesting characteristics of a multidimensional STR construct. We also note the variability of phenomena measured across instruments. This diversity illustrates how scholars have studied STRs using an array of approaches, which paints a multifaceted portrait of STRs. However, the heterogeneity of STR survey instruments' content and design—particularly where these qualities stray from survey-research methodologists' recommendations—creates a risk of inaccurate measurement of STRs and may therefore limit researchers' ability to further specify how to promote these critically important relationships. As such, we feel that the next generation of student-report STR survey instruments requires more stringent attention to both construct specification and validity and item generation—specifically, language use—in order to most fully measure aspects of STRs that matter to students. Following our description of these implications, we make recommendations for increasing the validity of self-report instruments for adolescents, a critical step in advancing STR survey-research methodology.

CONSTRUCT SPECIFICATION AND VALIDITY

<TXT>Even before identifying and analyzing this study's sample instruments, we found that an explicit, overarching STR construct does not yet exist. Instead, STR scholars referred to STRs but used a range of survey instruments (and study results) to describe these relationships. Fuller, more explicit specification of an STR construct is necessary to move research forward. We see this study's results as a step toward that end. First, the consistent identification of STR characteristics (such as teacher caring, respect, and fairness) across instruments suggests the presence of key characteristics of STRs and also suggests that STRs are multidimensional, consisting of multiple phenomena that represent the actions and perceptions of both students and teachers. These findings stand to extend the common conceptualizations of STRs that we found in the broad body of STR research literature (that STRs occur between students and teachers, represent a source of social capital that benefits students, and are shaped by status, power, and sociocultural differences between students and teachers). While pursuing the development of a multidimensional STR construct, though, it will be important to narrow down and specify the phenomena considered to contribute to strong STRs. Establishing the convergent and discriminant validity of an STR construct—phenomena that the construct both includes and excludes—is critical to more clearly defining the STR itself. From the 40 separate survey-item types we identified, we believe it is possible and critical

to identify key phenomena to assess so that future STR survey researchers can more precisely gauge students' understanding, experiences, and perceptions of STRs.

Survey items that remain conceptually distinct from one another may best remain distinguished as measuring separate phenomena. The clearest example of this distinction is the set of items and instruments measuring student connectedness to school. While connectedness research has always concerned phenomena including *and also* exceeding STRs, its results do not solely represent evidence of STRs or of these relationships' potential impact on student learning and wellness. Our research suggests that conflation of connectedness and STRs contributes to conceptual confusion. Efforts to isolate and define connectedness-specific dimensions of STRs (as illustrated by a Venn diagram, identifying dimensions that the two constructs share along with dimensions that remain distinct) could help to resolve this confusion. (Gehlbach and Brinkworth, 2011, used Venn diagrams for this purpose.)

A second area for further STR conceptual specification involves the need for a fuller representation of the breadth of extant STR knowledge. The STR survey instruments we analyzed have not yet fully incorporated recent knowledge about the role of student sociocultural identity and youth agency in STRs. Even though issues of identity have emerged as central in contemporary STR research (e.g., Antrop-González & De Jesús, 2006; Barber, 2002; Cooper & Miness, 2014; Valenzuela, 1999), identity-related content is scarce among the instruments we studied. Ethnicity, race, and gender were generally limited to representation as predictor variables, and were almost entirely absent from survey items themselves. Although many studies compared students' perceptions of STRs on the basis of their gender or race, these instruments rarely asked student respondents about how their identities shaped or intersected with the relationships they had with their teachers. Furthermore, when surveys ask students to generalize across their teachers or even quantify how many of their teachers act in a particular way toward them, the resulting data cannot specify whether the teachers they indicate in their responses share their ethnic or racial identity or their gender. This group of survey instruments seems to presume a detachment between teacher behavior and teacher identity, and that students perceive or value the same teacher behaviors regardless of each party's race, ethnicity, gender, age, or discipline. Although STR research regarding the influence of teacher identity has been less central in relation to studies' primary research questions (e.g., Phillippo, 2013; Rolón-Dow, 2005; Warikoo, 2004), this is an area that merits further exploration, particularly as our nation's school-aged population becomes more diverse alongside a teacher population that remains predominantly female and white.

In addition, issues of youth agency and voice seemed doubly underrepresented among the instruments we analyzed. First, whereas all included survey instruments sought students' perspectives, only five contained at least one item pertaining to the role of student voice in the classroom or school. The underrepresentation of student-voice items in self-report STR survey instruments parallels other findings about the muting of student voice in school reform and educational policy conversations (Taines, 2012). Although there is increased recognition that students' insights into their educational experiences can help strengthen teachers' practice and education policy (Conner, Ebby-Rosin, & Brown, 2015; Cook-Sather, 2009), our results suggest that most STR survey instruments do not yet acknowledge student voice in the classroom as a salient dimension of STRs. Second, students' reports of their contributions to STRs are dwarfed by their accounts of their teachers' actions in the surveys we sampled. When survey questions primarily ask students to describe their teachers' behaviors, the results offer little to no information on how students initiate, engage in, or sustain relationships with their teachers. These measurement approaches overlook the active role students play in forming and maintaining relationships with their teachers (Cooper & Miness, 2014; Noddings, 2005).

ITEM GENERATION: MORE METICULOUS ATTENTION TO LANGUAGE NEEDED

<TXT>Problems with survey item language in our sample return us to survey research methodologists' recommendations to pay careful attention to language while generating survey items. Our thoughts on this matter fall into two categories: the importance of unambiguous language, and the importance of producing results that are actionable for educators.

<C>*Unambiguous Language*

<TXT>Survey instrument language—particularly word choice and construction—shapes the kinds of answers received, or, at worst, “skews answers” (Beam, 2012, p. 152). Given the focus of so many STR student self-report surveys on adolescents, alongside the diversity of the U.S. secondary-school population, survey designers’ use of unambiguous language is critical. Survey-research methodologists note that general language leaves survey items open to multiple interpretations (Converse & Presser, 2005; Tourangeau et al., 2000). Research on teacher caring, for example, suggests that the apparently simple measurement of a teacher’s caring for students may measure one student’s perception of a teacher’s caring about academic performance, another student’s perception of a teacher’s caring about his personal issues, and a third student’s perception of a teacher’s caring about how he navigates a challenging educational system (Garza, 2009; Stanton-Salazar, 2011). Further, instruments that use general language do not gauge students’ reactions to teachers’ attempts to demonstrate caring—possibly perceiving it as welcome, misguided, or invasive (Phillippo, 2012), reactions driven by both power differences between adolescent students and their teachers and also varying sociocultural norms regarding personal privacy and relationships with institutional representatives. Concepts such as respect, fairness, caring, or support are likely to vary for adolescents across developmental stages as well as sociocultural groups, making specific language all the more important. In addition, the use of negatively worded items (which researchers would then reverse-score) further complicates respondents’ understanding of and ultimate response to survey items, a problem that Gable and Wolf (1993) noted is of particular concern for adolescent respondents. Further, reverse-scored items (e.g., indicating that a teacher treats students unfairly) may not indicate strong STRs but rather the presence and extent of negative STRs. For this reason, we echo Roorda et al.’s (2011) call to researchers to develop instruments that measure both negative and positive dimensions of STRs, as both have unique bearing on student outcomes.

Survey items’ depiction of students and teachers also requires redoubled attention in subsequent STR surveys. Lax specification of STR surveys’ focal teachers (e.g., some teachers, all teachers, specific teachers, adults at respondents’ schools) leaves open the possibility of respondents’ inaccurate reporting on relationships between themselves and their teachers. Further, surveys that solicit information on other students’ interactions with teachers, which we noted in our sample, are vulnerable to respondent reports based on that which they observe between other students and teachers, which may represent incomplete information about those relationships. More precise survey wording can specifically identify focal students and teachers, while also focusing respondent attention on STRs about which they possess full knowledge.

<C>*Actionable Results for Teachers*

<TXT>Research on STRs illustrates that teachers do not always know *how* to best establish relationships with their students and may even attempt to develop STRs in ways that make students uncomfortable or alienate them (Phillippo, 2012; Rolón-Dow, 2005). Survey item wording, therefore, holds the potential to clearly specify what teachers who have strong STRs do to promote those relationships. Just as it might be difficult for a student to determine the meaning of survey item language about caring, respect, or fairness, teachers may similarly find these terms perplexing. What exactly should a teacher do to show that she cares for and respects her adolescent students? Without more specific item language to illustrate those behaviors that students see as promoting strong STRs, survey research may not contribute to improved educator practice. Even terms that seem fairly actionable and straightforward, such as “praise,” “treat like adults,” and “give advice,” are not as clear as they may seem, particularly with adolescent students who are striving for independence while still often wanting teacher support.

MEASURES TO PROMOTE INSTRUMENT VALIDITY

<TXT>Moving forward, STR researchers stand to develop and use more valid instruments, and to in turn produce more accurate and useful data, by taking steps already identified by survey research methodologists. First, methods that solicit youth input and feedback promise survey language that more specifically describes youth experiences of STRs as well as the full range of phenomena that contribute to strong STRs. Cognitive pretesting or other means of soliciting youth feedback on instrument content (both concepts and language) (DeMaio & Rothgeb, 1996; Groves et al., 2009) would strengthen STR

instruments' ability to fully capture students' experiences of and perspectives on STRs. (See Bowen and Richman, 2008, for an example of cognitive pretesting of a student self-report survey). Such steps could also identify where students exhibit different understandings of item language that might point to variation on the student side of STRs, such as different conceptualizations of teacher respect. A diverse group of student respondents would be needed to ensure researchers' understanding of how item content and language are perceived by students across age groups, geographic regions, and sociocultural groups, and how those perceptions might ultimately shape responses and the data they create. Clear language about what teachers do (or do not do) to promote STRs is necessary to inform policymakers' and educators' efforts in this area.

Second, efforts to establish an overarching, coherent STR construct that provides for clear, consistent survey instrument content will benefit from psychometric analysis of survey results. Actions such as exploratory factor analysis, initial item analysis, and initial estimates of validity (Netemeyer et al., 2003) will promote an understanding of which survey items resemble and diverge from one another and will inform researchers' efforts to include meaningful items while trimming unnecessary ones. Toward the end of generating valid, attuned STR survey instrumentation, we recommend an iterative process that alternates between cognitive pretesting and psychometric analysis.

LIMITATIONS

<TXT>Our study advances the field's knowledge about STRs and STR measurement by suggesting qualities of these relationships that are central as well as qualities that are more peripheral but possibly worth further consideration for inclusion in future survey instruments. It also calls attention to qualities of extant instruments that jeopardize the validity and usefulness of the instruments and the data that they generate. Still, this study has its limitations. In regard to the creation of our sample of student self-report STR survey instruments, it is possible that our method of identifying survey instruments (multiple search-engine search using key terms related to STRs for peer-reviewed articles, then subsequently searching each identified article's reference list) may have missed pieces of research that include relevant survey instruments. Further, we did not analyze instruments that were either unavailable in their full form in published articles or were not furnished upon our request by the authors who used them, which also shrank our sample. The thoroughness of our search, nonetheless, helped us to create a rich database.

Another possible limitation to our study concerns our analytic methods. We chose not to conduct a quantitative inter-rater reliability (IRR) analysis (Krippendorff, 2013). Because this study examined a number of STR instrument characteristics, we determined that quantitative IRR procedures would have produced data that were voluminous but thin in meaning, generating a Cohen's kappa figure for each of the 62 analytic codes that we used. Instead, we analyzed our data using what Kvale and Brinkmann (2009, p. 243) called "dialogical intersubjectivity," in which team members reviewed and critiqued one another's coding until reaching full consensus on the definition and application of all codes used. In this way, we were able to simultaneously identify emergent findings and remain attentive to our methods' rigor and interpretive attunement.

<A>CONCLUSION

<TXT>Politicians, media commentators, and researchers often repeat the argument that the teacher is the single most important school-level factor to student success. Still, we need to know more about what it is that teachers do to promote not only academic achievement, but also healthy youth development. Research has long pointed to the salience of STRs toward these ends. However, our deep dive into the first-generation survey instruments measuring STRs has revealed murkiness in how these instruments measure and define STRs.

The time is right for a second generation of student self-report STR survey instruments that are more developmentally and socioculturally attuned, specific, consistent, and actionable. Such instruments would enable educators and policymakers to better understand teachers' contributions to strong STRs. The knowledge these instruments produce could, in turn, inform the design of preservice learning experiences, professional development opportunities, and even performance evaluation criteria for

teachers. This review can guide the development of new instruments and the refinement of older instruments, helping to move the field forward by providing concrete evidence of current STR instruments' strengths, shortcomings, and areas for future development. Ultimately, a more robust understanding of STRs requires a critical interrogation of what it is that we think we know about STRs and greater precision in the future measurement of this complex phenomenon. While this charge may seem steep, our findings reveal a clear pathway by which research can help teachers to become even more informed, supportive, and effective, and to promote classrooms in which students thrive.

<NOTES>

Notes

1. For the sake of clarity, we distinguish the author-named construct of “teacher caring” from individual survey items that inquire whether teachers care about their students. The author-named construct (*teacher caring*) represents the phenomenon the instrument as a whole strives to measure. When we use the phrase *teacher care*, as opposed to *teacher caring*, we refer to the various survey items that specifically measure teacher care. These items appear in surveys measuring constructs of teacher caring but also in surveys measuring other constructs, such as connectedness.

2. As a note of caution when interpreting the results in Table 2, percentages are affected by the co-occurrence of multiple author-named constructs within each instrument.

3. References marked with an asterisk were included in this study’s analysis of self-report survey instruments. A full listing of included studies is available upon request.

<REF>

References³

- Alonso, G., Anderson, N., Su, C., & Theoharis, J. (Eds.). (2009). *Our schools suck: Students talk back to a segregated nation on the failures of urban education*. New York, NY: New York University Press.
- Alwin, D. F. (2010). How good is survey measurement? Assessing the reliability and validity of survey measures. In P. V. Mardsen & J. D. Wright (Eds.), *Handbook of Survey Research* (2nd ed., pp. 405–436). Bingley, U.K.: Emerald Group.
- Anderman, L., Andrzejewski, C., & Allen, J. (2011). How do teachers support student motivation and learning in their classrooms? *Teachers College Record*, 113, 969–1003.
- Antrop-González, R., & De Jesús, A. (2006). Toward a theory of critical care in urban small school reform: Examining structures and pedagogies of caring in two Latino community-based schools. *International Journal of Qualitative Studies in Education*, 19, 409–433. doi:10.1080/09518390600773148
- Barber, T. (2002). "A special duty of care": Exploring the narration and experience of teacher caring. *British Journal of Sociology of Education*, 23, 383–395. doi:10.1080/0142569022000015427
- Beam, G. (2012). *The problem with survey research*. New Brunswick, NJ: Transaction.
- Blustein, J. (1991). *Caring and commitment: Taking the personal point of view*. New York, NY: Oxford University Press.
- Bohrnstedt, G. W. (2010). Measurement models for survey research. In P. V. Mardsen & J. D. Wright (Eds.), *Handbook of Survey Research* (2nd ed., pp. 347–404). Bingley, U.K.: Emerald Group.
- *Bowen, G. L., Richman, J. M., Brewster, A., & Brown, N. (1998). Sense of school coherence, perceptions of danger at school, and teacher support among youth at risk of school failure. *Child and Adolescent Social Work Journal*, 15, 273–286. doi:10.1023/A:1025159811181
- Bowen, G. L., & Richman, J. M. (2008). *The school success profile* (5th ed.). Chapel Hill, NC: Jordan Institute for Families, School of Social Work, University of North Carolina at Chapel Hill.
- Brinkworth, M. E., McIntyre, J., Harris, A. D., & Gehlbach, H. (2015). Understanding teacher-student relationships and student outcomes: The positives and negatives of assessing both perspectives. Manuscript submitted for publication.
- Burchinal, M., Howes, C., Pianta, R. C., Bryant, D., Early, D., Clifford, R., & Barbarain, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science*, 12, 140–153. doi:10.1080/10888690802199418
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81.
- *Colarossi, L. G., & Eccles, J. S. (2003). Differential effects of support providers on adolescents' mental health. *Social Work Research*, 27(1), 19–30. doi:10.1093/swr/27.1.19
- Conner, J., Ebby-Rosin, R., & Brown, A. (2015). *Student voice in American educational policy*. New York, NY: Teachers College Press.

- Conner, J., Miles, S., & Pope, D. (2014). How many teachers does it take to support a student? Examining the relationship between teacher support and adverse health outcomes in high-performing, pressure-cooker schools. *The High School Journal*, 98(1), 22–42. doi:10.1353/hcj.2014.0012
- Converse, J. M., & Presser, S. (2005). *Survey questions: Handcrafting the standardized questionnaire*. Thousand Oaks, CA: Sage.
- Cook-Sather, A. (2009). *Learning from the students' perspective*. Boulder, CO: Paradigm.
- Cooper, K., & Miness, A. (2014). The co-creation of caring student-teacher relationships: Does teacher understanding matter? *The High School Journal*, 97, 264–290. doi:10.1353/hcj.2014.0005
- Croninger, R. G., & Lee, V. E. (2001). Social capital and dropping out of high school: Benefits to at-risk students of teachers' support and guidance. *Teachers College Record*, 103, 548–581. doi:10.1111/0161-4681.00127
- *Crosnoe, R., Johnson, M. K., & Elder, G. H, Jr. (2004). Family dynamics, supportive relationships, and educational resilience during adolescence. *Journal of Family Issues*, 25, 571–602. doi:10.1177/0192513X03258307
- Cushman, K. (2003). *Fires in the bathroom*. Providence, RI: The New Press.
- *Daly, B. P., Shin, R. Q., Thakral, C., Selders, M., & Vera, E. (2009). School engagement among urban adolescents of color: Does perception of social support and neighborhood safety really matter? *Journal of Youth Adolescence*, 38, 63–74. doi:10.1007/s10964-008-9294-7
- *Darwich, L., Hymel, S., & Waterhouse, T. (2012). School avoidance and substance use among lesbian, gay, bisexual, and questioning youths: The impact of peer victimization and adult support. *Journal of Educational Psychology*, 104, 381–392. doi:10.1037/a0026684
- *Davis, H. A., & Lease, A. M. (2007). Perceived organizational structure for teacher liking: The role of peers' perceptions of teacher liking in teacher-student relationship quality, motivation, and achievement. *Social Psychology of Education*, 10, 403–427. doi:10.1007/s1128-007-9031-1
- *Demaray, M. K., & Malecki, C. K. (2002b). The relationship between perceived social support and maladjustment for students at risk. *Psychology in the Schools*, 39, 305–316. doi:10.1002/pits.10018
- DeMaio, T. J., & Rothgeb, J. M. (1996). Cognitive interviewing techniques: In the lab and in the field. In N. Schwarz & S. Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research* (pp. 177–196). San Francisco, CA: Jossey-Bass.
- *De Wit, D. J., Karioja, K., Rye, B. J., & Shain, M. (2011). Perceptions of declining classmate and teacher support following the transition to high school: Potential correlates of increasing student mental health difficulties. *Psychology in the Schools*, 48, 556–572. doi:10.1002/pits.20576
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method* (3rd ed.). Hoboken, NJ: J. Wiley.
- *Eisenberg, M. E., Ackard, D. M., & Resnick, M. D. (2007). Protective factors and suicide risk in adolescents with a history of sexual abuse. *The Journal of Pediatrics*, 151(5), 482–487. doi:10.1016/j.jpeds.2007.04.033
- Erickson, L. D., McDonald, S., & Elder, G. H. (2009). Informal mentors and education: Complementary or compensatory resources? *Sociology of Education*, 82, 344–367. doi:10.1177/003804070908200403

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59–109. doi:10.3102/00346543074001059

Gable, R. K. & Wolf, M. B. (1993). *Instrument development in the affective domain*. Boston, MA: Kluwer.

Garza, R.. (2009). Latino and white high school students' perceptions of caring behaviors. *Urban Education, 44*, 297–321. doi:10.1177/0042085908318714

Gay, G. (2010). *Culturally responsive teaching: Theory, research and practice* (2nd ed.). New York, NY: Teachers College Press.

Gehlbach, H., and Brinkworth, M. E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology, 15*(4), 380–387. doi:10.1037/a0025704

Green, G., Rhodes, J. E., Hirsch, A. H., Suarez-Orozco, C., & Camic, P. M. (2008). Supportive adult relationships and the academic engagement of Latin American immigrant youth. *Journal of School Psychology, 46*, 393–412. doi:10.1016/j.jsp.2007.07.001

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Hoboken, NJ: Wiley.

*Gutman, L. M., & Midgley, C. (2000). The role of protective factors in supporting the academic achievement of poor African American students during the middle school transition. *Journal of Youth and Adolescence, 29*, 223–248. doi:10.1023/A:1005108700243

Gwynne, J., & de La Torre, M.. (2009). *When schools close: Effects on displaced students in Chicago public schools*. Chicago, IL: Consortium on Chicago School Research.

*Hafen, C. A., Allen, J. P., Mikami, A. Y., Gregory, A., Hamre, B., & Pianta, R. C. (2011). The pivotal role of adolescent autonomy in secondary school classrooms. *Journal of Youth Adolescence, 41*, 245–255. doi:10.1007/s10964-011-9739-2

*Hallinan, M. T. (2008). Teacher influence on students' attachment to school. *Sociology of Education, 81*, 271–283. doi:10.1177/003804070808100303

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first–grade classroom make a difference for children at risk of school failure? *Child Development, 76*, 949 – 967. doi:10.1111/j.1467-8624.2005.00889.

Hughes, J. N., Luo, W., Kwok, O., & Lloyd, L. K. (2008). Teacher-student support, effortful engagement, and achievement: A 3-year longitudinal study. *Journal of Educational Psychology, 100*(1), 1–14. doi:10.1037/0022-0663.100.1.1

*Jiang, X., Huebner, E. S., & Siddall, J. (2012). *A short-term longitudinal study of differential sources of school-related social support and adolescents' school satisfaction*. Springer Science+Business Media.

Jones, M., & Yonezawa, S. (2002). Student voice, cultural change: Using inquiry in school reform. *Equity and Excellence in Education, 35*, 245–254. doi:10.1080/713845322

*Khan, S., Haynes, L., Armstrong, A., & Rohner, R. P. (2010). Perceived teacher acceptance, parental acceptance, academic achievement, and school conduct of middle school students in the Mississippi Delta region of the United States. *Cross-Cultural Research, 44*, 283–294. doi:10.1177/1069397110368030

- *Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health, 74*, 262–273. doi:10.1111/j.1746-1561.2004.tb08283.x
- Koepke, S., & Denissen, J. J. (2012). Dynamics of identity development and separation-individuation in parent-child relationships during adolescence and emerging adulthood: A conceptual integration. *Developmental Review, 32*(1), 67–88. doi:10.1016/j.dr.2012.01.001
- *Konishi, C., Hymel, S., Zumbo, B. D., & Li, Z. (2010). Do school bullying and student-teacher relationships matter for academic achievement? A multilevel analysis. *Canadian Journal of School Psychology, 25*(1), 19–39. doi:10.1177/0829573509357550
- Krippendorff, K. H. (2013). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage.
- Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (pp. 263–313). Bingley, U.K.: Emerald.
- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50*(1), 537–567.
- Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing* (2nd ed.). Thousand Oaks, CA: Sage.
- Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. *Theory Into Practice, 34*, 159–165. doi:10.1080/00405849509543675
- Lerner, D. G., & Kruger, L. J. (1997). Attachment, self-concept, and academic motivation in high-school students. *American Journal of Orthopsychiatry, 67*, 485–492. doi:10.1037/h0080249
- *Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school academic press. *American Educational Research Journal, 36*, 907–945. doi:10.3102/00028312036004907
- *Lewis, J. L., Ream, R. K., Bocian, K. M., Cardullo, R. A., Hammond, K. A., & Fast, L. A. (2012). Con cariño: teacher caring, math self-efficacy, and math achievement among Hispanic English learners. *Teachers College Record, 114*, 1–42.
- Libbey, H. P. (2004). Measuring student relationships to school: Attachment, bonding, connectedness and engagement. *Journal of School Health, 74*, 274–283. doi:10.1111/j.1746-1561.2004.tb08284.x
- Lynch, M., & Cicchetti, D. (1997). Children's relationships with adults and peers: An examination of elementary and junior high school students. *Journal of School Psychology, 35*(1), 81–99. doi:10.1016/S0022-4405(96)00031-3
- Masten, A. S., & Tellegen, A. (2012). Resilience in developmental psychopathology: Contributions of the Project Competence Longitudinal Study. *Development and psychopathology, 24*, 345–361. doi:10.1017/S095457941200003X
- Mayeroff, M. (1971). *On caring*. New York, NY: Harper and Row.
- * McCarty, C. A., Rhew, I. C., Murowchick, E., McCauley, E., & Vander Stoep, A. (2012). Emotional health predictors of substance use initiation during middle school. *Psychology of Addictive Behaviors, 26*, 351–357. doi:10.1037/a0025630
- McIntyre, A. (1997). Constructing an image of a white teacher. *Teachers College Record, 98*, 653–681.

McKenzie, J. F., Wood, M. L., Kotecki, J. E., Clark, J. K., & Brey, R. A. (1999). Establishing content validity: Using qualitative and quantitative steps. *American Journal of Health Behavior, 23*(4), 311–318.

*Muller, C. (2001). The role of caring in the teacher-student relationship for at-risk students. *Sociological Inquiry, 71*, 241–255. doi:10.1111/j.1475-682X.2001.tb01110.x

*Murray, C. (2009). Parent and teacher relationships as predictors of school engagement and functioning among low-income urban youth. *The Journal of Early Adolescence, 29*, 376–404. doi:10.1177/0272431608322940

Murray, C., & Zvoch, K. (2011). Teacher-student relationships among behaviorally at-risk African American youth from low-income backgrounds: Student perceptions, teacher perceptions, and socio-emotional adjustment correlates. *Journal of Emotional and Behavioral Disorders, 19*(1), 41–54. doi:10.1177/1063426609353607

*Nasir, N. S., Jones, A., & McLaughlin, M. (2011). School connectedness for students in low-income urban high schools. *Teachers College Record, 113*, 1755–1793.

National Council on Measurement in Education. (2015). *Code of professional responsibilities in educational measurement*. Retrieved from <http://ncme.org/resource-center/code-of-professional-responsibilities-in-educational-measurement/>

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues applications*. Thousand Oaks, CA: Sage.

*Niehaus, K., Rudasill, K. M., & Rakes, C. R. (2012). A longitudinal study of school connectedness and academic outcomes across sixth grade. *Journal of School Psychology, 50*, 443–460. doi:10.1016/j.jsp.2012.03.002

Nieto, S. (2010). *Language, culture and teaching: Critical perspectives*. New York, NY: Routledge.

Noblit, G. W. (1993). Power and caring. *American Educational Research Journal, 30*(1), 23–38. doi:10.3102/00028312030001023

Noddings, N. (2005). *The challenge to care in schools: An alternative approach to education* (2nd ed.). New York, NY: Teachers College Press.

Olsson, E. (2009). The role of relations: Do disadvantaged adolescents benefit more from high-quality social relations? *Acta Sociologica, 52*, 263–286. doi:10.1177/0001699309339802

Pennington, J. L., Brock, C. H., & Ndura, E. (2012). Unraveling the threads of white teachers' conceptions of caring: Repositioning white privilege. *Urban Education, 47*, 743–775. doi:10.1177/0042085912441186

Phillippo, K. L. (2012). "You're trying to know me": Students from nondominant groups respond to teacher personalism. *Urban Review, 44*, 441–467. doi:10.1007/s11256-011-0195-9

Phillippo, K. L. (2013). *Advisory in urban high schools: A study of expanded teacher roles*. New York, NY: Palgrave Macmillan.

Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. Washington, DC: American Psychological Association. doi:10.1037/10314-000

Poplin, M., & Weeres, J. (1994). *Voices from the inside: A report on schooling from inside the classroom*. Claremont, CA: The Institute for Education in Transformation at the Claremont Graduate School.

*Reddy, R., Rhodes, J. E., & Mulhall, P. (2003). The influence of teacher support on student adjustment in the middle school years: A latent growth curve study. *Development and Psychopathology*, 15, 119–138. doi:10.1017/S0954579403000075

Resnick, M. D., Bearman, P. S., Blum, R. W., Bauman, K. E., Harris, K. M., Jones, J., . . . Udry, J. R. (1997). Protecting adolescents from harm: Findings from the National Longitudinal Study on Adolescent Health. *Journal of the American Medical Association*, 278, 823–832. doi:10.1001/jama.1997.03550100049038

*Roeser, R. W., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology*, 88, 408–422. doi:10.1037/0022-0663.88.3.408

Rolón-Dow, R. (2005). Critical care: A color(full) analysis of care narratives in the schooling experiences of Puerto Rican girls. *American Educational Research Journal*, 42(1), 77–111. doi:10.3102/00028312042001077

Roorda, D. L., Koomen, H., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, 81, 493–529.

Rubin, B. (2003). "I'm not getting any F's": What "at-risk" students say about the support they need. In B. Rubin & E. Silva (Eds.), *Critical voices in school reform: Students living through change* (pp. 188–207). New York, NY: RoutledgeFalmer.

Rubin, B., & Silva, E. (Eds.). (2003). *Critical voices in school reform: Students living through change*. New York, NY: RoutledgeFalmer.

*Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal*, 38, 437–460. doi:10.3102/00028312038002437

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Thousand Oaks, CA: Sage.

Schwarz, N. (1999). Self reports: How the questions shape the answers. *American Psychologist*, 54(2), 93–105.

*Siddall, J., Huebner, E. S., & Jiang, X. (2013). A prospective study of differential sources of school-related social support and adolescent global life satisfaction. *American Journal of Orthopsychiatry*, 83(1), 107–114. doi:10.1111/ajop.12006

Stanton-Salazar, R. D., & Dornbusch, S. M. (1995). Social capital and the social reproduction of inequality: The formation of information networks among Mexican-origin high school students. *Sociology of Education*, 68, 116–135. doi:10.2307/2112778

Stanton-Salazar, R. D. (2011). A social capital framework for the study of institutional agents and their role in the empowerment of low-status students and youth. *Youth & Society*, 43, 1066–1109. doi:10.1177/0044118X10382877

Steinberg, L., & Silverberg, S. B. (1986). The vicissitudes of autonomy in early adolescence. *Child Development*, 57(4), 841–851. doi:10.2307/1130361

Student Voice Collaborative. (2011). *Student-led school improvement: Work, findings, next steps*. New York, NY: Author. Retrieved from

<http://nebula.wsimg.com/f80bc1acc7ec4614561788309dc6f665?AccessKeyId=477CBFF51112C7C068EE&disposition=0>

*Suldo, S. M., Friedrich, A. A., White, T., Farmer, J., Minch, D., & Michalowski, J. (2009). Teacher support and adolescents' subjective well-being: A mixed-methods investigation. *School Psychology Review, 38*(1), 67–85.

Taines, C. (2012). Intervening in alienation: The outcomes for urban youth of participating in youth activism. *American Educational Research Journal, 49*, 53–86. doi:10.3102/0002831211411079

Theoharis, J. (2010) "I hate it when people treat me like a Fxxx-up." In G. Alonso, N. Anderson, C. Su, & J. Theoharis (Eds.), *Our schools suck: Students talk back to a segregated nation on the failures of urban education* (pp. 69–111). New York, NY: New York University Press.

Toshalis, E. (2012). The rhetoric of care: Preservice teacher discourses that depoliticize, deflect and deceive. *Urban Review, 44*(1), 1–35. doi:10.1007/s11256-011-0177-y

Toshalis, E., & Nakkula, M. (2012). *Motivation, engagement, and student voice*. Boston, MA: Jobs for the Future.

Tourangeau, R., & Bradburn, N. M. (2010). The psychology of survey response. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (2nd ed., pp. 315–346). Bingley, U.K.: Emerald Group.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York, NY: Cambridge University Press.

Tuck, E. (2012). *Urban youth and school pushout*. New York, NY: Routledge.

*Tummala-Narra, P., & Sathasivam-Rueckert, N. (2013). Perceived support from adults, interactions with police, and adolescents' depressive symptomology: An examination of sex, race, and social class. *Journal of Adolescence, 36*, 209–219. doi:10.1016/j.adolescence.2012.11.004

Valenzuela, A. (1999). *Subtractive schooling: U.S. Mexican youth and the politics of caring*. Albany, NY: State University of New York Press.

*Voisin, D. R., DiClemente, R. J., Salazar, L. F., Crosby, R. A., & Yarber, W. L. (2006). Ecological factors associated with STD risk behaviors among detained female adolescents. *Social Work, 51*(1), 71–79. doi:10.1093/sw/51.1.71

Voisin, D. R., & Neilands, T. B. (2010a). Low school engagement and sexual behaviors among African American youth: Examining the influences of gender, peer norms, and gang involvement. *Children and Youth Services Review, 32*, 51–57. doi:10.1016/j.childyouth.2009.06.016

Voisin, D. R., & Neilands, T. B. (2010b). Community violence and health risk factors among adolescents on Chicago's Southside: Does gender matter? *Journal of Adolescent Health, 46*, 600–602. doi:10.1016/j.jadohealth.2009.11.213

Wilson, B. L., & Corbett, H. D. (2001). *Listening to urban kids: School reform and the teachers they want*. Albany, NY: State University of New York Press.

Ware, F. (2006). Warm demander pedagogy: Culturally responsive teaching that supports a culture of achievement for African American students. *Urban Education, 41*, 427–456. doi:10.1177/0042085906289710

Warikoo, N. (2004). Race and the student-teacher relationship: Interpersonal connections between West Indian students and their teachers in a New York City high school. *Race, Ethnicity and Education*, 7, 135–147. doi:10.1080/1361332042000234268

Werner, E. E., & Smith, R. S. (1982). *Vulnerable but invincible: A longitudinal study of resilient children and youth*. New York, NY: McGraw-Hill.

Werner, E. E., & Smith, R. S. (2001). *Journeys from childhood to midlife: Risk, resilience and recovery*. Ithaca, NY: Cornell University Press.

*Wilson, D. (2004). The interface of school climate and school connectedness and relationships with aggression and victimization. *Journal of School Health*, 74, 293–299. doi:10.1111/j.1746-1561.2004.tb08286.x

Wilson, B. L., & Corbett, H. D. (2001). *Listening to urban kids: School reform and the teachers they want*. Albany, NY: State University of New York Press.

*Woolley, M. E., Kol, K. L., & Bowen, G. L. (2009). The social context of school success for Latino middle school students: Direct and indirect influences of teachers, family, and friends. *The Journal of Early Adolescence*, 29(1), 43–70. doi:10.1177/0272431608324478

Zenkov, K. (2009). Seeing the pedagogies, practices, and programs urban students want. *Theory Into Practice*, 48, 168–175. doi:10.1080/00405840902997253

*Zullig, K. J., Huebner, E. S., & Patton, J. M. (2010). Relationships among school climate domains and school satisfaction. *Psychology in the Schools*, 48(2), 133–145. doi:10.1080/00405840902997253

Table 1. Descriptive Statistics of Articles in Sample

Number of articles (<i>N</i>)	66	
Mean sample size	4,391	Range: 62–83,731; median: 636
Articles analyzing group differences (%) [†]		
Gender	42	
Race/ethnicity	15	
Age or grade	14	
Other	17	SES, sexual orientation, language, disability
Articles that address validity (%)	55	
With detailed explanation	20	
With a general overview	15	
With a citation	24	
Origin of measures (%)		
Original	26	17% of which describe how developed
Existing measure(s)	68	29% of which explain how or why adapted
Modified or adapted existing measure(s)	21	
Author-named construct measured (%) [†]		
Teacher support	62%	Includes “teacher support” and “Support from adults in school”, “school support”, “teacher affective support” “social support from teachers”, “support from adult outside the home”
Connectedness	20	Includes “teacher connectedness,” “teacher connection”, “connection to school”, “conventional and unconventional connectedness to teacher”, “student-teacher connectedness”
General relationship between student and teacher	20	Includes “teacher relationship,” “domain-specific relationship,” “student-teacher relationship”
Teacher caring	6	“Teacher caring” or “teacher care”
Other	8	Includes teacher liking, beliefs about teachers, teacher monitoring, teacher respect, autonomy support, teacher acceptance, attachment to a teacher, teacher structure
More than one construct	23	
Mean number of items in instruments	8	Range: 1–36; median: 7
Instruments including at least one item that does not directly pertain to STRs (%)	27	Examples include items that refer simply to “adults” or “someone in school,” or that measure general qualities of school, such as fairness of rules or school safety.

Note. Categories are not mutually exclusive.

Table 2. Categories of Survey Items Most Frequently Used to Assess Student-Teacher Relationships, Sorted by Author-Named Construct

Item category	Instruments with item in category (%)	Instruments with item in category (%), sorted by author-named construct					Sample items
		Teacher caring	Teacher support	Connectedness	General relationship	Other	
Teachers care about student(s)	59.2	100	53.6	80.0	55.6	14.3	"My teacher cares about what I think and likes me to talk about it," "Teacher lets me know (s)he cares about me"
Teachers listen to student(s)	57.1	50	35.7	30.0	33.3	0	"My teachers really listen to what I have to say," "Teachers will listen if I want to talk about a problem."
Teachers treat student(s) fairly	32.7	0	28.6	40.0	55.6	42.9	"Teachers always try to be fair," "This teacher treated me unfairly"
Teachers show interest in or pay attention to student(s)	32.7	75.0	25.0	40.0	22.2	28.6	"[The teacher] takes a personal interest," "[The teacher] is interested in you as a person"
Teachers treat student(s) with respect or honesty	28.6	25.0	39.3	10.0	11.1	14.3	"I am respected and appreciated by my teachers," "Teachers here respect me"
Teachers give help or advice	26.5	0	35.7	20.0	22.2	28.6	"Teachers go out of their way to help," "This teacher never gave me help when I needed it."
Teachers have time for student(s)	24.5	0	32.1	20.0	33.3	28.6	"If students want to talk about something this teacher will find time to do it," "[The teacher] spends very little time just talking with students."
Student(s) and teachers get along	24.5	25.0	14.3	50.0	22.2	14.3	"I get along with my teachers," "This teacher and I always seem to be struggling with each other"
Student(s) can talk to teachers	22.4	25.0	17.9	20.0	22.2	42.9	"I enjoy talking to the teachers here," "Do you have a hard time talking to your teachers?"
Teachers criticize student(s) or make them feel bad	20.4	25	14.3	20.0	22.2	28.6	"This teacher was always criticizing and punishing me," "[The teacher] sees me as a big nuisance"

Note. Author-named constructs are not mutually exclusive and co-occur within items and instruments.