Soc Psychol Educ (2008) 11:133–147
DOI 10.1007/s11218-007-9046-7

Behavioral versus cognitive classroom friendship networks

Do teacher perceptions agree with student reports?

Matthew Pittinsky · Brian V. Carolan

Received: 5 July 2007 / Accepted: 29 October 2007 / Published online: 13 December 2007

Abstract Researchers of social networks commonly distinguish between "behavioral" and "cognitive" social structure. In a school context, for example, a teacher's perceptions of student friendship ties, not necessarily actual friendship relations, may influence teacher behavior. Revisiting early work in the field of sociometry, this study assesses the level of agreement between teacher perceptions and student reports of within-classroom friendship ties. Using data from one middle school teacher and four classes of students, the study explores new ground by assessing agreement over time and across classroom social contexts, with the teacher-perceiver held constant. While the teacher's perceptions and students' reports were statistically similar, 11–29% of possible ties did not match. In particular, students reported significantly more reciprocated friendship ties than the teacher perceived. Interestingly, the observed level of agreement varied across classes and generally increased over time. This study further demonstrates that significant error can be introduced by conflating teacher perceptions and student reports. Findings reinforce the importance of treating behavioral and cognitive classroom friendship networks as distinct, and analyzing social structure data that are carefully aligned with the social process hypothesized.

Keywords Social networks · Classroom sociometry · Cognitive social structure · Classroom friendships

An earlier version of this manuscript was presented at the American Education Research Association's Annual Meeting, April, 2007, Chicago, USA.

M. Pittinsky (✉)
Institute for Social and Economic Research and Policy (ISERP), Columbia University, 420 W. 118th Street, Room 270, New York, NY 10027, USA
e-mail: mp2055@columbia.edu

B. V. Carolan
College of Staten Island, CUNY, 2800 Victory Blvd., Building 3S-224, Staten Island, NY 10314, USA
e-mail: carolan@mail.csi.cuny.edu

1 Introduction

Researchers of social networks commonly distinguish between "behavioral" social structure and "cognitive" social structure (e.g., Krackhardt 1987). Cognitive social structure differs from behavioral social structure by its focus on social relations as *perceived*, distinct from whether such relations actually exist in any behavioral form. Everyday life demonstrates this distinction when two people considered to be friends turn out to be complete strangers. In such a case, the social structure one cognitively processes differs from the behavioral social structure that one's mind seeks to reflect.

In education, researchers such as Hallinan and Smith (1989) have argued that classroom friendship ties (e.g., friendship cliques) influence teacher pedagogical decisions. Indeed, a large and growing body of research on classroom social networks has demonstrated that behavioral forms of classroom friendship relations can have significant effects on learning related outcomes (e.g., Durland and Fredericks 2005; McFarland 2001). In many such studies, the social processes hypothesized are based on classroom friendship networks *as students experience them*. For example, McFarland (2001) found that a student's position in his or her classroom's friendship network served to embolden or suppress the student's misbehavior. In other cases, however, the social processes hypothesized are based on classroom friendship networks *as a teacher perceives them* (e.g., seating assignments). Who the teacher thinks is friends with whom, not necessarily actual friendships, is thought to influence teacher behavior and ultimately affect educational outcomes.

This distinction between behavioral and cognitive social structure carries with it an important practical consideration. To the extent research designs conflate cognitive and behavioral data—using student reports of friendships to test a hypothesis that is really based on teacher perception—researchers misalign theory and data. As Kilduff and Krackhardt (1994) have demonstrated, for phenomena based on an actor processing social structure—doing something because of the patterns of relations perceived—the proper unit of analysis is cognitive, not necessarily the state of affairs that may exist unperceived. For example, in their study of worker performance reputations in a hi-tech company, Kilduff and Krackhardt (1994) found that actually having a prominent friend did not enhance a worker's reputation, but *perceived* friendship ties exerted a significant "spillover" influence.

How might such a misalignment of theory and data affect empirical findings about educational phenomena? Stated another way, do student reports agree with teacher perceptions when it comes to classroom friendship ties and with what consequences for commonly used social network measures? This study addresses the first part of the question by examining the level of agreement between one teacher's perceptions of student friendships and the reports generated by her students within four middle-school Science classrooms, in both the autumn and spring of an academic year. Agreement is tested at the dyad level, meaning at the level of any two students being perceived as friends (or not friends) by the teacher consistent with what the students themselves reported.

Given whatever level of agreement is found, this study then addresses the second part of the question by exploring the magnitude of error that might be introduced when using student-reported friendship data to test a cognitive process instead of directly measuring the teacher's perceptions. The study assesses such error by using each classroom's teacher-perceived and student-reported friendship data, at each time point, to separately calculate two common ego social network measures. The results of these network measures are statistically compared to determine the degree to which using behavioral data to test a cognitive process might generate significantly and substantively different results.

Despite a large number of studies in the 1940s and 1950s that examined sociometric agreement between student and teacher reports of friendship and the implications of this (dis)agreement, surprisingly few studies have been conducted in recent years. Examples of recent studies include Gest (2006) and Pearl et al. (2007). Gest observed the paucity of studies as well, stating "perhaps surprisingly, there is little published data on the degree to which teacher reports of children's peer affiliation patterns correspond to peer reports or to direct observations" (2006, p. 249).

The present study's design has several distinct advantages. First, though the approach taken does not permit teacher variability to be measured, it explores new ground by assessing variation in agreement by the same teacher–perceiver across different classroom social contexts. Second, it tests for agreement at two time points, allowing for teacher–student agreement to vary from the autumn to spring. Third, it assesses agreement not only at the level of absolute correlation between two data sets (teacher-perceived and student-reported), but also in terms of how degrees of correlation ultimately affect measures commonly calculated using these data sets. Fourth, it analyzes data collected from an adolescent student population, in contrast to earlier studies, which analyzed data collected in elementary schools.[1] Finally, it benefits from advances in the field of social network analysis—e.g., software tools and measures—to more accurately assess agreement using network data that often violate standard statistical assumptions.

1.1 Teacher reports of student friendships: accuracy and perception

Two distinct research concerns have produced studies of teacher–student sociometric agreement. The first and earlier literature treated sociometric agreement as a question of teacher "accuracy," emphasizing that a teacher's ability to perceive student friendships is a skill; one that can enhance effective teaching (e.g., Bonney 1943, 1947; Gage et al. 1955; Gronlund 1951, 1955, 1957, 1959).[2] As argued by Gronlund (1955, p. 277), "a teacher's behavior in the classroom is guided to a large extent by how he *perceives* the needs and behavior of his students..." (emphasis added). If those perceptions are inaccurate, Gronlund argued, a teacher's effectiveness would suffer as a consequence.

Gronlund (1955,p. 277) defined sociometric accuracy as, "the degree to which an individual's judgments of the social acceptability of others agree with their actual social acceptability, as measured by a sociometric test." Importantly, his focus was on accuracy in judging *sociometric status* (rank order of popularity), not perceiving particular friendship choices, cliques or network-level characteristics. Gronlund (1951) found, not surprisingly, that teachers vary in their level of accuracy. The mean correlation among students and teachers in 40 sixth-grade classrooms in his sample was .60, with a range of .28 to .84. He found no difference in accuracy based on student sex, age of teacher, teacher years of experience, class size, or openness of classroom task structure; only the teacher's preference for a student proved a significant influence (overestimating popularity of favored students). Gronlund (1955) later found that the complexity of a classroom's behavioral friendship network was also a significant mediating factor, although he failed to consistently replicate this effect (1957).

Gronlund and Whitney (1958) explored the relationship between accuracy in perceiving student sociometric status and accuracy in judging student intelligence. "Actual" student

[1] It should be noted that the American school system has not always been partitioned as it is today. At least one "elementary" school study of sociometric agreement has included adolescent students (Moreno 1934), although most have not.

[2] Interestingly, Moreno (1934), often credited as the founding figure of sociometry, addressed this question in his first application of a sociometric instrument.

intelligence was measured using an intelligence test. They found high degrees of teacher accuracy in perceiving student sociometric status, with no difference for male and female students. As well, they found significant correlations between teacher sociometric accuracy and teacher judgments of student intelligence, although the correlations were stronger for girls than boys (.59 vs. .42). Consistent with Gronlund's earlier conclusions, they found significant variation among teachers in their degree of accuracy; most were accurate but a minority was highly inaccurate.

Two additional findings are noteworthy from this early research. First, studies found that teachers tend to overstate popularity rather than unpopularity, often based on the assumption that students who are not disliked by classmates are therefore liked, when in reality they are often simply ignored (Bonney 1947). Second, teachers commonly rely on student personality, visibility of student socialization and student participation in school groups as sources of information in judging student sociometric status (Bonney 1947). In summary, the early research literature suggests that teachers are generally "accurate" in their perceptions, although some teachers are remarkably inaccurate. This accuracy, however, may be shaped by a classroom's social context, which makes accurate perception more difficult in some contexts than others (no studies tested the same teacher's accuracy across classrooms). Finally, the size of the class and a variety of teacher and student attributes do not seem to affect a teacher's accuracy.

It is important to remember, however, that nearly all of these studies focused on levels of agreement regarding student sociometric status (how popular a student is), *not* sociometric choices (who considers whom to be a friend). Of the early studies, only Gage et al. (1955) analyzed correlations in specific choices, finding a significant correlation, but with considerable variability across teachers. As well, most of the statistical analyses conducted in the early literature were based on simple correlations, lacking the more sophisticated (and appropriate) approaches that have been developed in the field of social network analysis since.

In contrast to the studies discussed so far, recent investigations have generally moved beyond the question of accuracy, even avoiding the term in favor of "agreement."[3] Much as Casciaro (1998) and Krackhardt (1987) have argued regarding the social sciences in general, this second group analyzes teacher perceptions not because they might be accurate or inaccurate, but because they are analytically valuable in their own right. From this perspective, teacher perceptions—i.e. cognitive social structures—are too often overlooked in favor of direct measures of social structure. As Krackhardt argued, "The preoccupation with the ... [respondent] accuracy problem is symptomatic of a bias towards behavioral patterns even though the theoretical base is frequently cognitive or psychological" (1987, p. 111). For example, in a study of student behavior and its relationship to teacher grading, Takei et al. (1998) found that the perception a teacher held of her student's behavior, not the student's actual behavior, influenced the student's final grades. The present study fits squarely in this second group.

A non education-setting example of this motivating interest is Kilduff and Krackhardt's (1994) study, mentioned earlier, which demonstrated that having a prominent friend did not enhance a worker's reputation; *perceived* friendship ties were what mattered. Krackhardt (1987) has discussed extensively the meaning of cognitive social structures, as well as the practical measurement techniques and issues involved in studying them. The perceiver could be one of the actors themselves or an outside observer. In notation form, if traditional network data represents a friendship tie 'X' between actor 'i' and actor 'j' as X_{ij}, then cognitive

[3] While most studies of sociometric agreement conducted 40–50 years ago were motivated by an interest in teacher effectiveness, some contemporary studies have adopted this frame as well (e.g., Pearl et al. 2007).

data takes the form of X_{ijk}, where the 'k' subscript represents the actor perceiving the tie. Depending on the substantive question being investigated, cognitive social structures can be measured based on a single perceiver (what Krackhardt calls a "slice," and the approach used in this paper), or through techniques that create a composite picture of the network based on data from all perceivers (what Krackhardt calls a "consensus structure," and what Kilduff and Krackhardt used in their paper).

In education, Cairns and colleagues developed a similar approach, called the composite Social Cognitive Map (SCM) procedure, as a way of using a subset of network actors to construct complete network measures (e.g., Cairns et al. 1995). Using SCM, Gest (2006) provides a nice education-setting example of how the use of cognitive and behavioral networks can generate meaningfully different results, just as they did in Kilduff and Krackhardt (1994). In addition to assessing agreement between teacher perceptions and student reports of classroom friendships, Gest explored the degree of similarity between two tied students in the teacher's judgment of their behavior and academic orientation. In effect, he asked: To what extent are students similar to their classmate friends in the teacher's judgments of them? One might imagine a theory of peer-influence that frames the question in such a way that actual student friendships are what matter. One might imagine a theory of "halo effects" that frames the question in such a way that *perceived* student friendships are what matter. The former would argue that students are alike in teacher judgments because they select or influence each other; the teacher is still judging each student on his or her own. The latter would argue that students are alike because teachers balance their judgments among students they perceive to share friendship relations; the teacher "spills-over" judgment, the same process hypothesized by Kilduff and Krackhardt (1994). The frame employed by the researcher determines the relevant dataset.

One of Gest's (2006) notable findings is that teacher judgments were roughly twice as large in terms of similarity when the average of *teacher-perceived* friends was used (10–15%) than *student-reported* friends (5–10%). And it was four to five times as large when the teacher's *general judgment* of "a student's friends" (unspecified) was used (25–40%) than the average of student-reported friends. In other words, consistent with the halo theory, there *may* be an effect on teacher judgments based on who a teacher believes is friends with whom and her esteem for those friends. Were only behavioral (student-reported) network data used to test the cognitive theory, the magnitude of similarity would have been significantly understated.

Whether motivated by the belief that a teacher's ability to perceive student friendship ties is an important skill, or a dimension of social structure with explanatory value in its own right, or both, the question of agreement is an important one. Most studies to date have analyzed sociometric agreement at the elementary school level using data collected at a single time point. As a result, student populations have generally been pre-adolescent, and differences in agreement by the same teacher across classrooms and over time have been understudied. The present study addresses these particular limitations. As well, it builds on Gest's (2006) recent findings by assessing the risk to empirical investigations of using behavioral data to test a cognitive process.

2 Methods

To examine agreement between a teacher's and her students' reports of friendship, sociometric and demographic data were collected in four 8th-grade Science classes taught by the same teacher in the 2005–2006 academic year. The four classes were essentially sections of the same teacher's course, holding many critical variables constant: they were taught by the same instructor, served students within the same track, used the same curriculum and assessments, and were delivered within the same school and community context, in the same year. In contrast to earlier studies that analyzed data from multiple teachers who typically taught only one class, the use of same-teacher provides a naturally semi-controlled setting for an exploratory investigation into variability in agreement across classroom learning environments and populations.

2.1 Setting

Data were collected from a middle school located in a suburb of New York City. According to 2003 statistics, the community's racial composition was approximately 50% White, 30% Black, 15% Hispanic, and 5% Asian/Pacific Islander. As well, families living in the district divided approximately into thirds in terms of those who earned less than \$50,000, between \$50,000–99,999, and more than \$100,000. At the time of data collection, the community was experiencing long-term changes in its racial composition. As a consequence, in the study year (2005–2006), the 8th grade was comprised of fewer White students (−6%) and more Black students (+8%) than the community, reflecting the community's changing population mix in favor of younger minority families with children.

The school's principal helped identify the participating teacher. This Hispanic, bi-lingual (English and Spanish) teacher was a tenured 33-year old with seven total years of classroom experience, 5 years in her current school. Students interviewed as part of the study, regardless of their opinion of Science or school in general, described the teacher in positive terms. The department chair frequently consulted her on curriculum and policy decisions. In short, she was an experienced teacher who generally enjoyed positive relations with her students.

Of her five classes, four were selected to participate in the study. The classes selected were periods 1 and 4, which occurred before lunch and gym, and Periods 7 and 9, which occurred after lunch and gym. The basis for selecting the four was pragmatic. All classes that the teacher was willing to include, and for which sufficient informed consents were provided, were included. Period 2 was the only class that provided both a low response rate of consents and garnered expressions of concern by the teacher that involvement in the study might negatively affect their performance. All four of her classes were taught in the same physical classroom. Students were assigned to one of five tables, with five to six students to a table. Students socialized most with classmates at their tables and shared information such as grades most with classmates at their table. The teacher noted that table assignments were used as a primary behavior management technique.

2.2 Participation rates

Participation rates in the study were high and data from nearly all participants were collected, thanks in part to the small sample and an ability to follow-up if and when students were absent. There was no significant difference in the profile of participant and non-participants. The teacher participated fully, completing all surveys and interviews and nearly all students participated as well. For the primary networks used in the study, teacher-reported data are complete, while student-reported data are missing for at most 7% of students across class periods and time periods. Naturally, the incidents of missing answers within surveys were higher, but not troublingly so (e.g., less than 4% of cells in a student-reported classroom friendship matrix). Missing data are addressed differently depending on the analysis conducted.

2.3 Sample characteristics

At the beginning of the study (autumn), each of the teacher's four classes enrolled 27 students (total $n = 108$). At 27 students per class, these classes were larger than the school's 8th grade average of 23. Participating students represented approximately 25% of the middle school's total 8th grade population. This sample of students was roughly representative of the gender composition of the 8th grade, but not its racial composition, with an over-representation of Black and Hispanic students and under-representation of White students. When asked to explain these differences, the school administrator responsible for scheduling attributed it to the greater number of White students in the school's open-enrollment honors Science classes. With a mean of 79.4, the sample earned final marks in 7th grade Science that were slightly higher than the prior achievement of their 8th grade peers (78.5).

The overall sample did not change significantly during the school year in terms of gender, race or prior achievement. Seven students left the sample entirely, one student joined the sample, and one student stayed in the sample but moved across the teacher's class periods. The result of these changes in the spring was to decrease the total sample to 104 students, and to create a distribution of class sizes from 24 students to 27 students. The effect of these changes is more apparent when classroom-level descriptive statistics are provided.

As Table 1 illustrates, Period 1, autumn, was 30% male, balanced in terms of Black and White students (both 41%), but had the smallest Hispanic population of any of the classes (11%). As well, it was the most varied in terms of past Science achievement (*SD* = 10.00). Period 4, autumn, was the only balanced class in terms of sex and the most racially balanced relative to the total population studied and the middle school overall. Its past achievement in Science was the highest of the class periods. In general, the class periods differed significantly in racial and gender composition, but not past achievement. These differences were surprisingly large given the study's design and illustrate how class-scheduling processes can create non-random distributions of students across even same-track, same-teacher classes.

In terms of changes in class-by-class composition during the school year, three of the four class periods experienced at least one addition or removal. These changes may have affected the behavioral and cognitive network structures of the student population. Overall,

Table 1 Sample characteristics

	Period 1		Period 4		Period 7		Period 9	
	Autumn	Spring	Autumn	Spring	Autumn	Spring	Autumn	Spring
Student reports, participation rate	.96	1.00	.96	.96	.93	1.00	1.00	1.00
Teacher reports, participation rate	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Male	.30	.27	.59	.59	.78	.75	.30	.30
Female	.70	.73	.41	.41	.22	.25	.70	.70
African-American	.41	.42	.41	.41	.30	.25	.63	.63
White	.41	.39	.37	.37	.30	.33	.19	.19
Hispanic	.11	.12	.22	.22	.37	.38	.15	.15
Other	.07	.07	.00	.00	.03	.04	.03	.03
Mean science achievement	79.00	79.00	82.50	83.00	79.00	78.30	77.50	77.50
(SD)	(10.00)	(10.20)	(8.00)	(7.80)	(7.80)	(8.10)	(8.20)	(8.20)
N	27	26	27	27	27	27	24	27

however, they did not significantly alter the distinctive descriptive compositional features of each class.

2.4 Sociometric survey instrument

To measure the level of agreement between student and teacher reports of classroom student friendships, sociometric data were collected from each student in all four classes and the teacher provided similar reports on all students. To collect student reports of friendships, students were given a class roster and asked to describe their relationship with each student in the class. Choices included best friend, friend, know-like, know, know-dislike, strongly dislike, and do not know. In the terminology of network analysis, these sociometric data are "valued" (degrees of friendship, not just yes or no) and "directed" (friendship nominations were not presumed to be reciprocal). Data were collected in the autumn and spring. All "best friend" and "friend" choices are coded as '1' (friend), while all other choices are coded as '0' (not friend). Reciprocated friendships were then identified; meaning both students in the dyad marked the other as a friend. This focus on reciprocated friendships is consistent with the literature (e.g., Gest 2006; Kilduff and Krackhardt 1994). In the case of 8th graders, reciprocated ties also cut through the noise of those students who might have generously nominated large numbers of classmates as friends, when in fact only a subset were mutual and on par with the meaning of "friend" used by the other students in the class.

The teacher's reports of students' friendships were generated in a similar manner. In both the autumn and spring, the teacher could nominate as many friends as she thought appropriate for each student and she had the ability to consider directed relationships (i.e. she could say student A is a friend of student B, without necessarily saying that student B is a friend of student A). Here, too, only reciprocated friendships are included in the analysis. As well, she had access to a roster to help stimulate her thinking.

These student and teacher reports of friendships in both the autumn and spring were collapsed into non-directed (reciprocal), binary sociomatrices, in which a '1' indicates that any two students are friends. Given four class periods, two time points, and two data sources (teacher and student), the result is 16 sociomatricies providing reports of who is friends with whom, from the students' and teacher's perspectives, in both the autumn and spring.

3 Results

3.1 Analyzing agreement

The level of agreement between the teacher's and students' reports of friendship is first assessed at the dyad level. The dyad level is the level of any two students being tied (or not tied) as friends, and is distinct from other levels of analysis such as the group level (i.e. whether teacher-perceived friendship groups agree with student-reported friendship groups, when the participants were asked specifically about friendship *groups*). Findings provide evidence of the extent to which the teacher perceived the friendship dyads reported by the students. While one might assume the level of disagreement found is indicative of error, as will be demonstrated, this is not necessarily the case.

Before describing the first statistical technique used to test agreement, it is helpful to remember that the rows and columns of each sociomatrix are comprised of the class roster. Recall that the intersection of any student's row with any classmate's column was coded with a '1' if they were reciprocated friends and '0' if they were not. The first analysis of agreement focuses on whether there is a match in the value ('0' or '1') found in the same cell across the teacher and student generated matrices. Since reciprocated friendships mean that student A is a friend of B and B is a friend of A, the matrices are symmetrical. In a classroom of 27 students, there are 351 ((27×26)/2) cells to be tested for matches.

To assess dyad agreement, first, QAP (quadratic assignment procedure) correlations for each class's two matrices (teacher and student generated) were analyzed in the autumn and spring. A QAP correlation is used to calculate the degree of association between two sets of relations; it tests whether the probability of dyad overlap in the teacher matrix is correlated with the probably of dyad overlap in the student matrix. It does so by running a large number of simulations. These simulations generate random matrices with sizes and value distributions based on the original two matrices being tested. It then computes an average level of correlation between the matrices that would be expected at random. Similarly, it calculates the probability that the observed degree of correlation between two matrices would be as large or as small as that observed based on the range of correlations generated in the random permutations, with an associated significance statistic.

Based on 5,000 permutations, the results from the simple matching QAP correlations are presented in Table 2. To interpret the results, the values in the "Expected" and "Observed" columns can be considered as percentages. For example, in Period 1, autumn, there were 351 mutual dyad-level observations in each of the friendship matrices (teacher and student). Based on the size of the network and the distribution of nominations, we would expect 77% of those observations to match (both '1' or both '0'). The observed (actual) correlation value was 83.2%, a significantly greater-than-expected level of matching ($p < 0.001$). Or, put another way, the results of the procedure indicate that the observed level of matching exceeds chance levels when compared to thousands of randomly permuted matrices with the same number of nodes (students) and ties (friendships). With this example in mind, it is clear that across all class periods and times, the teacher's perceptions agreed with students' reports at a significant level. This agreement ranged from 6 points higher than expected in Periods 1, 4 and 7, autumn, to 14 points higher in Period 9, spring. Indeed, the teacher demonstrated an identical greater-than-average level of agreement across three of the four classrooms in the autumn, but was notably higher for Period 9 (9 points above average compared with 6 points for the other three classes). From autumn to spring, agreement increased meaningfully, remaining highest for Period 9, but with the interesting exception of Period 4. Whereas agreement improved by

Table 2 QAP correlation[a] between student and teacher reports of students' dyadic friendship ties

		No. of Observations	Actual	Expected	Difference
Period 1	Autumn	351	.832	.770	0.06*
	Spring	325	.892	.780	0.11*
Period 4	Autumn	351	.798	.736	0.06*
	Spring	351	.869	.794	0.08*
Period 7	Autumn	351	.846	.784	0.06*
	Spring	276	.707	.590	0.12*
Period 9	Autumn	325	.729	.641	0.09*
	Spring	351	.789	.651	0.14*

[a] The QAP procedure compares the observed correlation with a distribution of random correlations generated according to the null hypothesis of no relationship between the matrices

* $p < .05$, the proportion of random correlations that are as large or larger than the observed correlation

roughly 5 points between autumn and spring in the other three classes, period 4 evidences only a 2-point improvement.

Agreement was generally consistent across class periods and improved over time. Agreement occurred at levels statistically greater than chance; yet it is clearly premature to conclude that the two sources of data—teacher and student—agree. After all, in period 7, spring, as many as 29% of cells did not match. In any classroom friendship network, it is likely that most potential ties are '0,' untied. As a result, the high absolute level of agreement found in the QAP correlations may be generated by agreement on the absence of ties, but not necessarily the existence of ties. While this agreement was greater than chance, and the absence of ties can be a substantive phenomenon (e.g., White et al. 1976), it remains to be demonstrated whether the teacher perceived student classroom friendship ties (i.e., just the '1's).

While there are several approaches to testing agreement among ties only (e.g., Gest 2006), a convenient yet effective approach is to statistically compare the densities of the teacher-perceived and student-reported sociomatricies for each class and time period. Density is a social network measure that calculates the percentage of all possible dyads that could have been tied, which were indeed tied.[4] For example, in a classroom of 27 students, with 351 possible reciprocated friendship ties, if 50 friendship ties are reported then the density would be 50/351, or 14.2%. If the densities are found to be significantly different across paired student and teacher sociomatrices, then the number of friendships is significantly different.

Contrary to the QAP correlation results, the student-reported and teacher-perceived densities are significantly different across the classes and time periods. For example, in Period 7, spring, the density statistic generated by student-reported data is approximately 3 times the density statistic generated by teacher-perceived data. In Period 1, spring, the difference is only approximately 1.5 times. These differences are statistically significant, driven in every case by the teacher under-perceiving student-reported friendships. It is not possible with available data to assess whether the student reports were "inflated," overstating "actual" friendships, or behavioral and simply mis-perceived. Descriptively, it appears that the teacher's density measures moved in the same direction as the students' density measures, and the gap between teacher and student measures closed over time for most class periods. Period 7 is a notable exception in which the teacher's perceptions became more divergent, not less. In this case, the

Table 3 Comparison between student and teacher reports of density of classroom friendship ties

		Period 1		Period 4		Period 7		Period 9	
		Autumn	Spring	Autumn	Spring	Autumn	Spring	Autumn	Spring
Density[a]	Student-reported	.174	.157	.219	.160	.171	.384	.332	.279
	Teacher reported	.086*	.092*	.080*	.068*	.068*	.112*	.080*	.160*

[a] This procedure, based on the bootstrap approach, provides a more accurate estimate of standard error. Available in UCINET v.6.109 (Borgatti et al. 2002)

*The density of classroom-level friendship ties from the student reports is greater than those reported by the teacher, $p < .05$

[4] More precisely, density is a straightforward test of agreement about the number of ties given the size of each network. It is possible that the teacher and students' agree on density and still disagree on some or all particular ties, requiring additional analysis to confirm agreement. In the present analysis, density is sufficient given the significant difference in number of ties.

teacher's density statistic moved in the same direction as the students' (density increased); it simply did not keep pace with the dramatic change reported by the students (.171 autumn to .384 spring).

The data describe classes whose densities varied meaningfully by class period in the scale and direction of change over time, suggesting the teacher faced different social contexts as she moved from one class period to another. Based on these results, it would appear that the teacher perceived all four class periods similarly in the autumn, and differently as the school year progressed and she learned more about their unique social and contextual dynamics.

3.2 Analyzing error

Based on these findings, it would appear student reports produce sociomatricies statistically similar to those produced by teacher perceptions. At the same time, the teacher significantly under-perceives reciprocated friendship ties. It is not yet clear whether the nature of these differences would produce statistically and/or substantively different findings if the data sources were conflated. What might be the consequences of using student reports to test a social process that is based on teacher cognition?

Each student has his or her own "ego-network," or network of relations with classmates to whom she/he is connected (and who themselves may be connected with each other). To explore this second question, two common and meaningful measures of a student's position in their ego network are calculated. First, each student's degree is calculated, simply defined as the total number of reciprocated friendship ties that a student has received. This measure approximates a student's popularity. Second, each student's centrality is calculated using Freeman's (1977) betweenness measure. Specifically, betweenness centrality measures how central a given student is in his/her friendship network. Higher normalized betweenness scores, ranging from 0 to 1, suggest that a student is located on a greater number of pathways that connect any two of his/her friends. These two ego-level measures capture the average student's structural location in his/her friendship network.

To assess agreement across classrooms, permutation-based t-tests were computed and scores across classrooms for each time period were calculated. Table 3 reports the mean scores of both of these measures. These mean scores are the classroom averages of individual-level measures.

It is evident from Table 4 that cognitive and behavioral data produce significantly different individual-level results for each class, autumn and spring. Take for example Period 1, spring, where on average students report having approximately 4 friends (mean = 3.92, $SD = 2.84$). Data based on the teacher's perceptions result in students having on average approximately 2 friends (mean = 2.31, $SD = 1.56$). This difference in means of 1.62, is statistically significant ($p < .01$). Across all four classes, both autumn and spring, a similar pattern regarding the reports on degree is evident: the teacher data indicate a fewer number of friends per student, on average, and this difference between student and teacher reports is statistically significant. The extent of differences varies notably across class periods at each time point, but does decline from autumn to spring.

A similar pattern is evident for the betweenness scores. Examining Period 1, spring, once again, student-reported data generate an average student betweenness score of .342. This statistic, the average of students' normalized betweenness scores in a given class, can be interpreted as a percentage that indicates the frequency with which a given student is located on a path that connects any two students in her/her own ego-network. On average, student data in Period 1, spring, indicate that the average student is "between" roughly one-third (.342, $SD = .246$) of all relations constituting in their ego-network. Contrast this value with

Table 4 Comparison between mean student and teacher reports of individual-level attributes: degree and betweenness

Ego-level attribute		Period 1 Autumn	Period 1 Spring	Period 4 Autumn	Period 4 Spring	Period 7 Autumn	Period 7 Spring	Period 9 Autumn	Period 9 Spring
Degree	Student-reported	4.52(3.34)	3.92(2.84)	5.70(3.51)	4.15(3.00)	4.44(3.22)	8.83(3.52)	8.31(4.00)	7.26(4.26)
	Teacher reported	2.22(1.80)	2.31(1.56)	2.07(1.46)	1.78(1.66)	1.78(1.77)	2.58(1.35)	2.00(1.36)	4.15(2.50)
	Difference in means[a]	2.30*	1.62*	3.63*	2.37*	2.67*	6.25*	6.31*	3.11*
Betweenness	Student-reported	.329(.243)	.342(.328)	.287(.239)	.284(.266)	.389(.304)	.166(.085)	.222(.193)	.264(.254)
	Teacher reported	.175(.327)	.259(.404)	.285(.394)	.091(.232)	.139(.318)	.058(.155)	.165(.300)	.394(.367)
	Difference in means	.154*	.083*	.002	.193*	.250*	.108*	.057	−.130

[a] The t-test used to establish significance uses a permutation test to generate the significance level so that standard assumptions on independence and random sampling are not required (Snijders and Borgatti 1999). Conventional t-statistics are not generated by this procedure, available in UCINET v.6.109 (Borgatti et al. 2002)
* $p < .05$, one-tailed

that of the average betweenness score generated by the teacher's perceptions for the same class, same semester. Here, the average betweenness score is .175 ($SD = .327$), suggesting that the teacher's data, on average, would result in students seeming less central in their own ego-networks.

With a few notable exceptions (Period 1, autumn, and Period 9, autumn and spring), this difference between means is statistically significant and provides further evidence of disagreement between teacher perceptions and student reports. As with the degree statistic, agreement varies across class periods at the same time, but unlike the degree statistic, agreement does not generally decline over time. This pattern of disagreement suggests that conflating the two data sources can generate significantly and substantively different empirical findings.

4 Conclusion

A premise of this study is that cognitive social structures matter. For example, teachers manage classrooms and may rely on their perceptions of student friendship patterns to make consequential decisions (e.g., student assignments to cooperative learning groups). Given that teacher perceptions likely matter, they deserve direct measurement. Relying on student-reported behavioral friendship data can introduce significant error when testing teacher cognitive social processes; so too can relying on teacher perceptions as a substitute for collecting data directly from students. Surprisingly, the degree to which teacher perceptions and student reports of classroom friendship relations agree has been an understudied topic since the late 1950s.

In contrast with early studies that examined the level of agreement between student and teacher reports of sociometric status (rank order of popularity), the analyses reported here focused on reciprocated friendship dyads. While the teacher studied perceived classroom friendship networks (ties and non-ties) that were statistically similar to those reported by her students in four different classrooms over time, she significantly under-perceived student reported friendship ties. This conclusion supports the recent findings of Gest (2006). As a consequence of this disagreement, the results further suggest a significantly large degree of error can be introduced by conflating teacher perceptions and student reports, depending on the network measure involved.

Speculation suggests that the moderate amount of agreement found can be attributed to the fact that both the teacher and students agree, more often than not, on who is *not* friends with whom. The importance of this unit of agreement should not necessarily be minimized. As White et al. (1976) have noted, the absence of a relation is just as important as the relation itself. Nevertheless, given the patterns observed, there is reason to challenge the conventional notion that teachers' perceptions of within-class social relations are accurate.

The findings reported here are preliminary in nature and speak to the larger need of the research community to resume work in this area. There are several questions that come to mind. First, what individual, structural and instructional factors help explain the variability when examining the same teacher across classrooms and time periods? Casciaro (1998: 332) has emphasized the importance of understanding "the variability in people's accuracy in perceiving the informal social networks that develop in their social groups." Second, what information do teachers rely on when perceiving student friendship patterns? Third, does a teacher's network of relations in a school contribute significantly to her ability to perceive student friendship relations? A final question on this short list of many possibilities is whether agreement between teachers' and students' reports correlates with other forms of cognitive–behavioral agreement.

The distinctive contributions of behavioral and cognitive social structures are perhaps best illustrated by testing specific hypotheses that, while based on cognitive processes, are susceptible to being analyzed with behavioral data, thus introducing error. Kilduff and Krackhardt (1994) study of performance reputations is an example of such an approach. In education, so too is Gest's (2006) study of teacher–student sociometric agreement. Indeed, both provide strong empirical support for a longstanding value of social-psychological research: the critical importance of aligning theory and data in empirical investigations.

References

Bonney, M. E. (1943). The constancy of sociometric scores and their relationship to teacher judgments of social success, and to personality self-ratings. *Sociometry, 6*, 409–424.

Bonney, M. E. (1947). Sociometric study of agreement between teacher judgments and student choices. *Sociometry, 10*, 133–146.

Borgatti, S. P., Everett , M. G., & Freeman, L. (2002). *Ucinet 6.0 for Windows*. Natick: Analytic Technologies.

Cairns, R. B., Leung, M. C., Buchanan, L., & Cairns, B. D. (1995). Friendships and social networks in childhood and adolescence: Fluidity, reliability, and interrelations. *Child Development, 66*, 1330–1345.

Casciaro, T. (1998). Seeing things clearly: Social structure, personality and accuracy in social network perception. *Social Networks, 20*, 331–351.

Durland, M., & Fredericks, K. A. (Eds.). (2005). *The use of network analysis in program evaluation: Trends, techniques, and applications*. New Directions in Program Evaluation, *107*. San Francisco: Jossey Bass.

Freeman, L. C. (1977). A set of measures of centrality based on betweeness. *Sociometry, 40*, 35–41.

Gage, N. L., Leavitt, G. S., & Stone, G. C. (1955). *Teachers' understanding of their pupils and pupils' ratings of their teachers*. Psychological Monographs No. 106. Washington, DC: American Psychological Association.

Gest, S. D. (2006). Teacher reports of children's friendships and social groups: Agreement with peer reports and implications for studying peer similarity. *Social Development, 15*, 248–259.

Gronlund, N. E. (1951). The accuracy of teachers' judgments concerning the sociometric status of sixth-grade pupils, *Sociometry Monographs*, No. 25. New York: Beacon House.

Gronlund, N. E. (1955). Social perception and teaching effectiveness. *The School Review, 63*, 277–283.

Gronlund, N. E. (1957). The general ability to judge sociometric status: Elementary student teachers' sociometric perceptions of classmates and pupils. *The Journal of Educational Psychology, 47*, 147–157.

Gronlund, N. E. (1959). *Sociometry in the classroom*. New York: Harper & Brothers.

Gronlund, N. E., & Whitney, A. P. (1958). The Relation between teachers' judgments of pupils' sociometric status and intelligence. *The Elementary School Journal, 58*, 264–268.

Hallinan, M. T., & Smith, S. S. (1989). Classroom characteristics and student friendship cliques. *Social Forces, 67*, 898–919.

Kilduff, M., & Krackhardt, D. (1994). Bringing the individual back in: A structural analysis of the internal market for reputation in organizations. *The Academy of Management Journal, 37*, 87–108.

Krackhardt, D. (1987). Cognitive social structures. *Social Networks, 9*, 109–134.

McFarland, D. A. (2001). Student resistance: How the formal and informal organization of classrooms facilitate everyday forms of student defiance. *American Journal of Sociology, 107*, 612–678.

Moreno, J. L. (1934). *Who shall survive?* Washington: Nervous and Mental Disease Publishing Co.

Pearl, R., Leung, M.-C., Van Acker, R., Farmer, T. W., & Rodkin, P. C. (2007). Fourth- and fifth-grade teachers' awareness of their classrooms' social networks. *Elementary School Journal*, *108*, 25–39.

Snijders, T. A. B., & Borgatti, S. P. (1999). Non-parametric standard errors and tests for network statistics. *Connections, 22*, 61–70.

Takei, Y., Johnson, M. P., & Clark, M. E. (1998). Academic achievement and impression management as factors in the grading of white junior high pupils. *Sociological Perspectives, 41*, 27–48.

White, H. C., Boorman, S. A., & Breiger, R. L. (1976). Social structure from multiple networks. I: Blockmodels of roles and positions. *American Journal of Sociology, 81*, 730–778.

Author Biographies

Matthew Pittinsky, is a Ph.D. candidate in Sociology of Education at Teachers College, Columbia University and a Mellon Graduate Fellow at the Institute for Social and Economic Research and Policy (ISERP) at Columbia University. His research interests include social network analysis, balance theory, and new models of Internet-based data collection.

Brian V. Carolan, Ph.D., is an Assistant Professor of Education at the College of Staten Island, The City University of New York. His research interests include social networks and the sociology of education. He is currently working on a project that examines the effects of adolescents' social networks in small secondary schools.