False positives using social cognitive mapping to identify children's peer groups

Zachary P. Neal, Jennifer Watling Neal & Rachel Domagalski
Michigan State University

Children and adolescents interact in peer groups, which are known to influence a range of psychological and behavioral outcomes. In developmental psychology and related disciplines, social cognitive mapping (SCM), as implemented with the SCM 4.0 program, is the most commonly used method for identifying peer groups from peer report data. However, in a series of four studies, we demonstrate that SCM has an unacceptably high risk of false positives. Specifically, we show that SCM will identify peer groups even when applied to random data. We introduce backbone extraction and community detection as one promising alternative to SCM, and offer several recommendations for researchers seeking to identify peer groups from peer report data.

Introduction

Decades of research demonstrate the importance of peers for child and adolescent development and psychological well-being (Bukowski, Laursen, & Rubin, 2018; Gifford-Smith & Brownell, 2003). Children and adolescents interact in peer groups with structural and behavioral features that are associated with a wide range of psychological, social, and academic outcomes (Birkett & Espelage, 2015; Espelage, Holt, & Henkel, 2003; Ryan, 2001). However, identifying peer groups can be challenging and represents a critical measurement task for developmental and clinical researchers (Kindermann & Gest, 2018). To overcome these challenges, Cairns and colleagues proposed social cognitive mapping (SCM), a method of peer group identification that involves identifying peer groups using multiple peer reports of groups of children that interact together in a setting such as a classroom (Cairns & Cairns, 1994; Cairns, Cairns, Neckerman, Gest, & Gariépy, 1988).

SCM has become a dominant method for identifying children's peer groups from peer report data. The data are easy to collect, the ability to triangulate from peers reduces the impact of non-response, the analysis is easy to perform, and there is some evidence for its validity (Gest, Farmer, Cairns, & Xie, 2003). However, nothing is known about the extent to which SCM can yield false positives, where the method identifies peer groups from data that contain no or only weak evidence of their existence. In this paper, we show that SCM has a high rate of false positives, assigning on average two-thirds of children to peer groups even when it is applied to random peer report data. We conclude that researchers should *not use SCM, particularly as it is implemented in the* `SCM 4.0` *program, to identify peer groups*, and should explore alternative methods for identifying peer networks and peer groups from peer report data (Z. Neal, 2014).

We begin by reviewing SCM, providing an overview of its origins, where and how it has been used, how it works, and evidence for its accuracy. Then, in a series of four related studies, we confirm that SCM can detect true positives, examine SCM's risk of false positives under different conditions, and explore backbone extraction and community detection as an alternative to SCM. In the discussion section, we synthesize the key findings from these studies, offering recommendations for researchers seeking to identify children's peer groups.

Background

How are peer groups measured?

Experiences in peer groups play a significant role in childhood and adolescent development (Howe, 2010; Kindermann & Gest, 2018; Rubin, Bukowski, & Bowker, 2015). Specifically, aspects of peer group structure (e.g., size, hierarchy) or behavior (e.g., norms) have been linked to psychological (e.g., depression), social (e.g., aggression, homophobic name calling, prosocial behavior, resource control) and academic (e.g., motivation, achievement) outcomes (Birkett & Espelage, 2015; Espelage et al., 2003; Ryan, 2001; Zarbatany, Ellis, Chen, Kinal, & Boyko, 2019; Zhao, Chen, Ellis, & Zarbatany, 2016). Therefore, it is important for developmental researchers to have methods for measuring peer

Correspondence concerning this article should be addressed to Zachary Neal, Department of Psychology, Michigan State University, East Lansing, MI 48824; zpneal@msu.edu.
Zachary P. Neal https://orcid.org/0000-0003-3076-4995
Jennifer Watling Neal https://orcid.org/0000-0002-7749-8121
Rachel Domagalski https://orcid.org/0000-0003-2457-9195

groups.[1]

There are multiple ways to operationalize peer groups and empirically measure them. Perhaps the most direct approach is field observation: a researcher directly observes children's interactions with one another in a naturalistic setting such as a classroom or playground (e.g., Cairns, Perrin, & Cairns, 1985; Gest et al., 2003). From these observations, peer groups might be operationalized as sets of children seen to interact at all, to interact some minimum number of times, or to interact for some minimum duration. Although observation can yield the most direct evidence of peer groups, it can also be the most time-consuming approach to measurement. Therefore, more often peer groups are measured using more indirect approaches that vary in terms of the *source* of the report and what is *reported* (see Table 1).

Table 1

Data for measuring peer groups.

	Source	
Reported	Self	Peer
Group	Name List	SCM
Network	SNA	CSS

First, self-reported groups involve asking a child to report a list of the names of each other child in his or her own peer group(s). This approach is simple, and offers the advantage that it can be used to measure a single child's peer group(s) without requiring the collection of data from any other children. However, by relying on self-reported data, it is subject to self-enhancement bias (i.e. a child seeking to appear popular may report being a member of a popular group or may leave less popular members out of their group) (Gifford-Smith & Brownell, 2003). Additionally, when the goal is to identify all peer groups in a setting, self-report data can introduce bias due to missing data when some children in the setting do not participate and do not provide reports.

Second, self-reported networks involve asking each participating child in a setting to report his or her social contacts in the setting. Each child's reports are then combined to yield the setting's peer network. This represents the most common approach to collecting data for social network analysis (SNA), and has been used to such frequently studied peer network data as AddHealth (e.g. Haynie, 2001). Again, by relying on self-report data, it is also subject to biases due to self-enhancement and missingness (Cairns & Cairns, 1994; Gifford-Smith & Brownell, 2003; J. Neal, 2008). Additionally, because networks and not groups are reported, the identification of peer groups requires the application of a community detection algorithm.

Third, peer-reported groups involve asking each participating child in a setting to report the members of each peer group in that setting. By relying on reports of the setting's groups from multiple members of the setting, this approach reduces the risk of bias from self-enhancement and missingness, however it does require a method for triangulating or combining these multiple reports. The most widely used approach for processing peer-reported group data is social cognitive mapping (SCM; Cairns & Cairns, 1994), which is the focus of this paper.

Finally, peer-reported networks involve asking each participating child in a setting to report the social contacts of every other child in the setting, thereby yielding each participating child's view of the setting's network. Again, by relying on multiple reports, this approach reduces the risk of bias from self-enhancement and missingness, but requires a method for triangulating or combining these multiple reports. The most widely used approach for combining multiple network reports in a setting is cognitive social structures (CSS; Krackhardt, 1987; J. Neal, 2008).

What is social cognitive mapping?

Among the various approaches to measuring peer groups, social cognitive mapping (SCM) is one of the most widely used in developmental psychology and related fields. Cairns and colleagues (Cairns & Cairns, 1994; Cairns et al., 1988) proposed SCM as a method for reducing the biases and costs involved in other methods (Cairns & Cairns, 1994; Kindermann & Gest, 2018; J. Neal & Neal, 2013). SCM relies on peer informants to provide reports of groups of children in a particular setting, such as a classroom, that hang out together. Through a series of aggregating and filtering transformations, SCM uses these peer-reported data to identify peer groups (J. Neal & Neal, 2013). Specifically, SCM is intended to answer two questions: *first*, do the children in this setting interact with one another in peer groups, and *second*, if so, which children are members of which groups?

SCM developed in roughly three phases. First, during the *development* phase in the late 1980s, a research team at the University of North Carolina at Chapel Hill led by Robert Cairns and Beverley Cairns experimented with ways to triangulate multiple children's reports of peer groups into a single picture of a setting's social structure and its peer groups (Cairns, Cairns, & Neckerman, 1989; Cairns et al., 1988, 1985). This work concentrated on examining a matrix of children's co-occurrence in reported peer groups, evolving from "a decision rule procedure [in which] arbitrary standards were adopted" to a more objective set of steps "with minimal reliance on intuitive judgements" (Cairns et al., 1988, p. 817). Second, during the *formalization* stage

[1] In the peer relations literature, peer groups have been conceptualized based on social interactions or identity-based "crowds" (e.g., jocks, goths, nerds, etc.) (Brown, Mounts, Lamborn, & Steinberg, 1993; Rubin et al., 2015). In this paper, we only focus on the measurement of peer groups that are defined based on social interactions.

in the early 1990s, these steps were refined into a consistent procedure that appeared in an essentially identical form across multiple papers that included fully-worked examples (Cairns & Cairns, 1994; Farmer & Cairns, 1991; Farmer, Stuart, Lorch, & Fields, 1993). This phase also included the development of software to facilitate the use of SCM (Leung & Alston, 1998).[2] Finally, the *application* phase from the mid-1990s onward has involved the use of SCM throughout developmental psychology and related fields focused on studying children's peer relations and groups, as well as the development of variations on the steps developed during the formalization stage. In particular, two of Cairns' colleagues developed their own variations: one relying on conditional probabilities and a binomial z test (Kindermann, 1993), and another relying on principal components analysis (Gest, Moody, & Rulison, 2007).

How often is SCM used?

To determine the specific variant of SCM that is most commonly used, we started with a dataset of 201 papers from a recent review of social network data collection methods in developmental psychology (J. Neal, 2020). These papers were initially identified using Google Scholar and reflect papers published or online (a) prior to February 2019 (b) in the 30 top-ranked journals classified by Web of Science as "Psychology, developmental" in 2016 (c) that contained the phrase "social network" and one or more network-relevant keywords (e.g. density, centrality, clique, etc.) J. Neal (in press). We reviewed each paper and identified 73 that attempt to identify network-based peer groups or cliques. A majority of these papers (N = 46, 63%) used SCM to identify peer groups. Among those using SCM, most (N = 38, 83%) used the specific variant described by Cairns and Cairns (1994).[3] Finally, among the papers using Cairns and Cairns' version of SCM, nearly half (N = 16, 42%) explicitly noted that they used SCM 4.0, a DOS-based program that implements this version of SCM. Because it is the most widely-used version of this method, *in this paper we use 'SCM' to refer specifically to the method described by Cairns and Cairns (1994) and implemented in SCM 4.0 (Leung & Alston, 1998).*

To determine how widely SCM is used and the extent to which it is used outside developmental psychology, where it wad first developed, we examined each paper citing Leung and Alston (1998). Using Google Scholar we located an additional 26 papers appearing in such youth-focused fields as school psychology (Farmer, Hall, Petrin, Hamm, & Dadisman, 2010), social psychology (Wolfer, Bull, & Scheithauer, 2012), special education (Avramidis, 2010), STEM education (Radovic, Black, Salas, & Williams, 2017), and substance use (Sheppard, Golonka, & Costanzo, 2012). We also observed that it is used outside North America, by researchers in Latvia (Levina & Ivanova, 2012), Korea (Ahn & Shin, 2011), Norway (Fandrem, Ertesvag, Strohmeier, & Roland, 2010), and Spain (Bacete & Perrin, 2013).

Combining the results from J. Neal (in press) and our own search, we identified a total of 42 papers using SCM. They were published in such flagship journals as *Developmental Psychology* and *Child Development* between 1995 and 2019 (M = 2009.6, SD = 5). We therefore conclude that SCM is among the most widely- and currently-used methods for identifying peer groups.

What is SCM used to study?

Researchers use the peer groups identified by SCM in multiple ways. First, some researchers use SCM-derived peer groups to generate group-level behavioral norms, then estimate mixed models to examine associations between these group norms and individual psychological and social outcomes (Chung-Hall & Chen, 2010; Zhao et al., 2016). For example, Zhao et al. (2016) found that children who participated in SCM-derived peer groups with higher levels of average social withdrawal exhibited less social competence, less positive school attitudes, and higher levels of depression. Second, some researchers study the association between compositional (e.g., ethnic composition) or organizational features (e.g., hierarchization) of SCM-derived peer groups and psychological or social outcomes (Shi & Xie, 2014; Zarbatany et al., 2019). For example, Shi and Xie (2014) found that the socialization of aggression differed depending on the ethnic composition of SCM-derived peer groups. Finally, some researchers have used SCM-derived peer groups to examine the extent to which teachers are accurate observers of classroom peer relationships (i.e. teacher attunement) (Gest, 2006; Hoffman, Hamm, & Farmer, 2015). For example, Hoffman et al. (2015) found that elementary school teachers' reports of classroom peer groups exhibit only modest attunement to SCM-derived peer groups.

How does SCM work?

SCM begins by collecting peer reports by asking participating children a question like *Are there people in school who hang around together a lot? Who are they?* Each participating child is permitted to report any number of "hanging around" groups, and each group they report can contain any number of children including themselves. For example, Child A might report the existence of a "hanging around" group composed of children A, B, and C (report 1), and another group composed of children W, X, Y, and Z (report 2). Additionally, Child B might report the existence of a group

[2] Usually SCM 4.0 is attributed to Leung only, however here we cite both Leung who wrote the manual and original program, and Alston who is identified in the program itself as the programmer.

[3] Of the remaining papers, 3 use the variant described by Kindermann (1993), 2 use the variant described by Gest et al. (2007), and 3 provided insufficient detail to determine the variant.

composed of children A, B, C, and D (report 3). Thus, different reporters may report different numbers of groups (e.g. Child A provided two reports, while Child B provided one report), and these reports may partially overlap (e.g. Child A's first report and Child B's first report overlap). Then a series of aggregations and transformations are applied to these raw data to define a peer network and identify peer groups.

First, these peer report data are organized as a setting-wide "recall matrix" **R** that contains a row i for each child in the setting and a column for each report j, so that each cell in the matrix R_{ij} contains a 1 if child i appeared in report j, and otherwise is 0.[4] Returning to the example above, cell $R_{B1} = 1$ and cell $R_{B3} = 1$ because child B appeared in both reports 1 and 3, but $R_{X1} = 0$ because child X did not appear in report 1. The SCM 4.0 program imposes some restrictions on the recall matrix that are not necessarily required by SCM in general: it can only contain data on up to 2000 peer-reported groups and up to 400 distinct children, and each peer-reported group can contain up to 20 members.

Second, the reports in the recall matrix **R** are transformed into a symmetric "co-occurrence" matrix **C** using

$$\mathbf{C} = \mathbf{RR}' \tag{1}$$

where C_{ij} and C_{ji} contain the number of times child i and child j were reported to be in the same group, and C_{ii} contains the number of times child i was reported to be in any group. In this step, SCM mirrors a classic example from the social network literature in which children's potential interactions are inferred from their co-participation in school clubs using bipartite projection (Breiger, 1974).

Third, **C** is transformed into a "similarity" matrix **S** using

$$\mathbf{S} = cor(\mathbf{C}) \tag{2}$$

where S_{ij} is the Pearson correlation coefficient of child i's and child j's column (or row) in **C**. In this step, SCM mirrors the CONvergence of iterated CORrelations (CONCOR) algorithm for group detection (Breiger, Boorman, & Arabie, 1975). However, unlike CONCOR, which repeatedly computes the correlation of the matrix (e.g. cor(cor(cor(**C**)))), SCM performs this operation only once. Additionally, unlike CONCOR and most other network analytic techniques, SCM includes the diagonal of **C** when performing this step (Z. Neal & Neal, 2013).

Fourth, a binary peer network **N** is constructed by defining child i and child j as connected if

$$S_{ij} \geq 0.4 \tag{3}$$

The threshold value of 0.4 was first recommended, without a justification, by Cairns and Cairns (1994). This value is used by the SCM 4.0 program (Leung & Alston, 1998), which does not allow an alternate value to be specified. Although both SCM and its implementation in SCM 4.0 use a threshold value of 0.4, see Appendix A for a sensitivity analysis examining other values.

Finally, groups of peers are identified from the peer network described by **N**. These groups are permitted to overlap so that a single child may be a member of none, one, or more than one peer group. Cairns and Cairns (1994) do not describe a specific method, and seem to suggest that the identification of groups in **N** is a trivial task that can be performed by visual inspection. At least two different descriptions of the method used to identify peer groups appear in the literature. Farmer et al. (1993) and others report that it identifies peer groups so that each member of the group is connected to "at least 50% of the members in the cluster" (p. 234; see Avramidis, 2010; Fandrem et al., 2010; Rodkin, Farmer, Pearl, & van Acker, 2006). Separately, Bacete and Perrin (2013) report that it identifies peer groups by adding members to a group until no one "is found who has a correlation profile equal to or greater than r = .40 with any of the members who have previously been incorporated into the group" ([translated] p. 64-65). However, by examining the output generated by SCM 4.0 in the analyses described below, we have verified that it does not use either of these methods for identifying peer groups from a peer network. Thus, SCM 4.0 remains a black box (Z. Neal & Neal, 2013); we (and, seemingly, others) do not know exactly how SCM identifies peer groups from a peer network.

All five of the steps involved in SCM are automated by the SCM 4.0 program (Leung & Alston, 1998), which has been used in at least 42 published studies to identify peer groups. Although we do not know exactly how SCM 4.0 performs the final 'group identification' step, in our analyses below we simply use the output generated by the program.[5]

Is SCM accurate?

Because all methods measure social phenomena with error, the goal of measure development is to understand and minimize those errors. Therefore, it is often important to ask whether a given measurement approach is "accurate." When SCM is used to study peer groups, there are at least three distinct varieties of accuracy that might be investigated.

First, *validity* describes the extent to which peer groups identified via SCM match peer groups that can be directly ob-

[4]The recall matrix does not contain information about *who* provided any given report; this information is not used by SCM when identifying peer groups.

[5]The user's guide dated 20 August 1998 does not contain this information. We attempted to contact both the developer listed in the user's guide (Man-Chi Leung) and the developer identified in the program (Anthony Alston), but were unable to reach either of them.

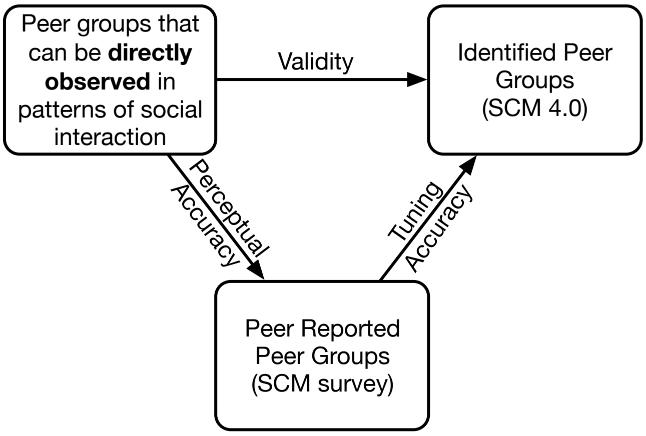

Figure 1

Types of accuracy relevant when social cognitive mapping is used to identify peer groups from peer report data.

served in patterns of children's social interactions. Because such observational data are difficult to collect (Gest et al., 2003), there have been limited attempts to evaluate SCM's validity. One early study of a 7th grade classroom of 26 children, observed that the children were "more likely to interact with members of their own [SCM-identified] subgroup than with members of other [SCM-identified] clusters" (Cairns & Cairns, 1994, p. 107). Similarly, a larger study of 72 children in 4th through 7th grade found that "Children were observed to interact with members of their SCM-identified social cluster at a rate four times higher than with other same-sex classmates" (Gest et al., 2003, p. 513).

Second, *perceptual accuracy* describes the extent to which the peer groups identified by a particular child on an SCM survey and that appear in the recall matrix match the peer groups that can be directly observed. Because "respondents have only a limited knowledge of the classroom" it is assumed that "there are frequent errors" in the reports that each child provides, and thus that perceptual inaccuracy is common (Cairns & Cairns, 1994, p. 104). For example, J. W. Neal, Neal, and Cappella (2016) found that children have inaccurate perceptions about which of their classmates hang out together ($\kappa = 0.371$), but that girls, older children, and children in smaller classrooms were more accurate. Such inaccurate perceptions can still be informative, for example as an indicator of a child's social awareness (Cappella, Neal, & Sahu, 2012), and because an individual's behaviors are shaped by their perceptions of reality as much as by reality itself (Krackhardt, 1987). However, SCM aims to overcome the expected perceptual inaccuracies of individual children by pooling and triangulating multiple children's reports.

Finally, a third type of accuracy which we call *tuning accuracy* describes the extent to which the peer groups identified by SCM accurately summarize or triangulate the information contained in the peer reports collected via an SCM survey. To use the analogy of a radio tuner, the peer report data is a combination of signal (i.e. information about directly observable peer groups) and noise (i.e. random error due to the children's perceptual inaccuracies) (Shannon & Weaver, 1963). The purpose of SCM, like a tuner, is to filter out the noise to yield a clear signal. There are two ways that SCM might exhibit tuning accuracy: first it can identify peer groups for which the peer reports contain evidence (true positives), and second it can fail to identify peer groups for which the peer reports do not contain evidence (true negatives). There are also two ways that SCM might exhibit tuning inaccuracy: first it can identify peer groups for which the peer reports do not contain evidence (false positives; type I error), and second it can fail to identify peer groups for which the peer reports do contain evidence (false negatives; type II error).

Each of these forms of accuracy is important. However, perceptual accuracy and validity are both challenging to establish because collecting observational data from many diverse settings would be cost- and time-prohibitive. In contrast, tuning accuracy can be evaluated without observational data by using simulated data as we describe below. Moreover, tuning accuracy is a critical prerequisite for validity. Without the ability to tune in the signal, and tune out the noise, no amount of perceptual accuracy will allow SCM-identified peer groups to be valid. Therefore, in this paper, we are interested in investigating SCM's tuning accuracy, and specifically its risk of false positives. Although false positives may not be of direct substantive interest for developmental psychologists seeking to identify peer groups, their presence is methodologically important because it can lead researchers to draw erroneous substantive conclusions. In the case of SCM, false positives may lead a researcher to conclude that peer groups exist when the data to not justify such a conclusion.

Three prior works offer some insight into SCM's risk of false positives. First, Watts (2008) explains that when a recall matrix is transformed into a co-occurrence matrix using Equation 1 "even a random [recall matrix] – one that has no particular structure built into it at all – will be highly clustered" (p. 128). Second, Pijl, Koster, Hannink, and Stratingh (2011) found that compared to identifying peer groups from a network of reciprocated self-reported friendships, SCM assigned all children to a peer group, which was "quite surprising, as it is known from the literature that 4-10% of children do not have friends in primary classrooms" (p. 484). Finally, J. Neal and Neal (2013) demonstrated using illustrative data that "distinct peer groups always appear to be present, no matter what responses children give" during data collection (p. 605). Guided by this past work, *we hypothesize that SCM has a high rate of false positives*, identifying peer groups even from peer report data that lack evidence of peer groups. To investigate this hypothesis, we report on four separate studies: The first study confirms that SCM can detect true positives, the second and third studies examine SCM's risk of false positives under different conditions, and the fourth study explores backbone extraction and community detection as an alternative to SCM for identifying peer groups. Replication code for all studies is available at https://osf.io/txgph.

Study 1: True positives in a benchmark classroom

Methods

To examine whether SCM is able to detect true positives, correctly identifying when children are members of peer groups, we use data described by Cairns and Cairns (1994). These data consist of a set of 61 peer reports collected from 17 children (11 girls, 6 boys) in a 26 child (15 girls, 11 boys) 7th grade classroom.[6] They offer an ideal benchmark dataset for two reasons. First, these are the data originally used by Cairns and Cairns (1994) to demonstrate and validate SCM, and therefore ought to offer SCM the best opportunity for correctly identifying peer groups. Second, independently of collecting these data, Cairns and Cairns (1994) also conducted "direct observation of social interactions among children" (p. 107), finding that all children in the classroom belonged to a peer group, and that the classroom contained five distinct peer groups. This observational data provides a criterion against which to judge SCM's validity.

In this and subsequent studies, we focus on one statistic of interest, P, the proportion of children that SCM identifies as a member of a peer group composed of at least 3 children.[7] In this benchmark classroom, based on the independent observational data that all children are members of a peer group, if SCM can detect true positives then P should equal 1.

Results

Figure 2 shows the peer network obtained by applying SCM to data from a benchmark 7th grade classroom, while the shaded regions outline the peer groups identified by SCM. The identified groups match what Cairns and Cairns (1994) observed: cohesive groups of 4 girls, 4 boys, and 7 boys, as well as a larger cluster of 7 girls and 3 girls that are bridged by Heather, who belongs to both groups but here is shown as a member of the larger group. Because all children are identified by SCM as a member of a peer group, $P = 1$, which matches our expectation based on observational data and confirms that SCM detects the true positives in this classroom. From this, we conclude that at least in this benchmark classroom, SCM is able to detect true positives with respect to whether or not children are members of peer groups.

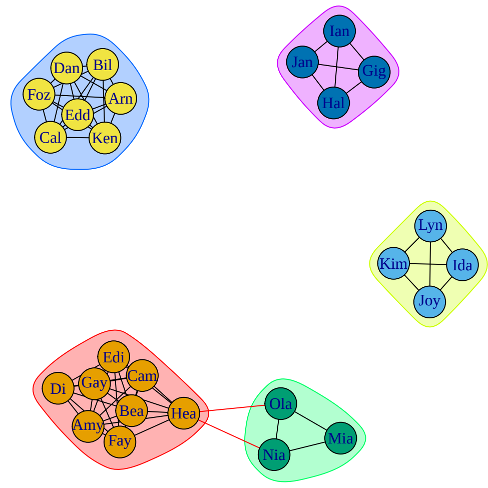

Figure 2

Applying SCM to peer report data from a benchmark 7th grade classroom.

Study 2: False positives in a benchmark classroom

Methods

Evaluating SCM's risk of false positives is more challenging than examining its ability to detect true positives. In this context, a false positive occurs when SCM identifies children as a members of peer groups when the peer report data lacks evidence of such peer group membership. The most extreme example of peer report data that lacks evidence of peer groups is random peer report data, which a researcher might obtain if the responding children simply guess about peer groups, do not take the data collection seriously and report nonsense, or lack any perceptual accuracy. Although obtaining random peer report data may be unlikely in practice, it provides the most conservative test of SCM's risk of false positives because it is the case where its risk of false positives should be smallest.

To test SCM's false positive rate, we generated 1000 simulated recall matrices. In each simulated recall matrix, we randomize which children have been reported by their peers to

[6] The classroom contained 27 children, but one child (Pam) never appeared in any of the peer reports, and therefore is excluded from these analyses.

[7] The requirement that peer groups contain at least 3 members is common in the developmental psychology literature (Shi & Xie, 2014; Zarbatany et al., 2019; Zhao et al., 2016)

be members of which groups. However, for the sake of comparability and to ensure these simulated data sets are plausible, each simulated recall matrix preserves some features of the original 7th grade classroom described above (Strona, Nappo, Boccacci, Fattorini, & San Miguel-Ayanz, 2014). First, the simulated data contain the same number of children (i.e. 26) and the same number of peer reports (i.e. 61). Second, they preserve the salience of each child. For example, in both the original Cairns and Cairns (1994) data and in every simulated dataset Arn is a high-salience child who was named in 15 peer reports, while Ken is a low-salience child who was named in only 3 peer reports. Finally, they preserve the sizes of the peer reported peer groups. For example, in both the original Cairns and Cairns (1994) data and in every simulated dataset 28 of the 61 peer reported peer groups contained 4 children, while 1 contained 12 children. This approach yields simulated recall matrices that have many of the same features as the original 7th grade classroom, except that by randomizing which children are reported as members of which groups, should contain no evidence of actual peer groups.

We then use SCM 4.0 to identify peer groups from each of these 1000 simulated recall matrices. In each case, we compute our statistic of interest, P, the proportion of children that SCM identifies as a member of a peer group composed of at least 3 children.

Because these simulated recall matrices are random, they should generally not contain any evidence of peer groups. Therefore, when SCM is applied to one of these random recall matrices, it should fail to identify children as members of peer groups. We can use this logic to determine how often SCM yields true negatives or false positives. If P is skewed toward 0 across the 1000 simulated recall matrices, this means that SCM is correctly failing to identify children as members of peer groups when applied to random data, and that SCM yields true negatives. Alternatively, if P is not skewed toward 0 and $P \gg 0$, this means that SCM is incorrectly identifying children as members of peer groups when applied to random data, and provides evidence that SCM yields false positives.

Results

Figure 3 summarizes the value of P obtained by using SCM to identify peer groups in 1000 random recall matrices with characteristics similar to the 7th grade classroom originally studied by Cairns and Cairns (1994). Because these recall matrices are known to be random, SCM should not find evidence in them that children are members of peer groups, and therefore P should be skewed toward 0. However, as Figure 3 illustrates, P is *not* skewed toward 0, indicating that SCM frequently yields false positives. How severe are the false positives when applied in this a setting like this? We find that on average SCM assigns two-thirds (M = 0.67, SD = 0.12, min = 0.27, max = 1) of children to a peer group, when in fact the recall matrix contains no evidence of peer groups because it is random.

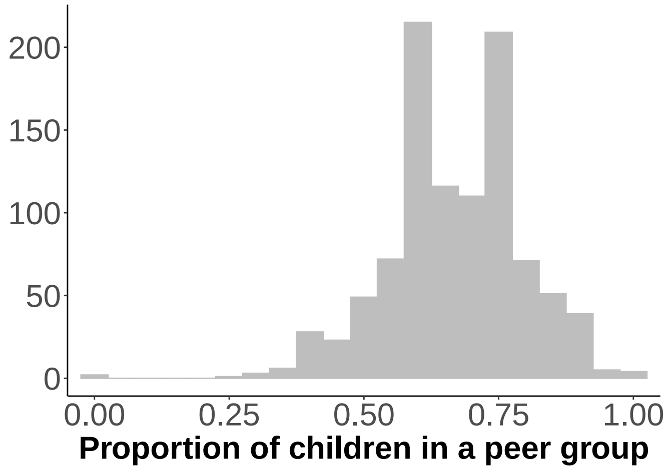

Figure 3

False positives applying SCM in a hypothetical 7th grade classroom.

Study 3: False positives in other classrooms

Methods

In study 2 we investigated SCM's risk of false positives in simulated classrooms that were similar to the benchmark classroom originally studied by Cairns and Cairns (1994). However, classrooms can vary widely, and SCM's risk of false positives may be more or less severe in certain types of classrooms. To investigate this possibility, we generated an additional 1000 random recall matrices, varying five features of the simulated classroom: (1) the number of children in the classroom, (2) the number of peer reports provided, (3) the probability that a child is named in a peer report, (4) the amount of skew in the number of times children were named in a peer report, and (5) the amount of skew in the number of children named in a peer report. In this context, skew in the number of times children were named is a measure of child salience; when skew is positive, this corresponds to a classroom where a few children are highly salient and receive many nominations, but most receive few nominations. Skew in the number of children named in a peer report is a measure of group size; when skew is positive, this corresponds to a classroom where some reported groups are large, but most are small. Then, following the same process as study 2, we use SCM 4.0 to identify peer groups from each recall matrix and compute our statistic of interest, P. Finally, we estimate a regression to examine how characteristics of classroom settings are associated with P.

Results

Figure 4 summarizes the value of P obtained by using SCM to identify peer groups in 1000 random recall matrices with the characteristics of classrooms that differ in size, density, child salience, and reported group size. As in study 2, because these recall matrices are known to contain random data, SCM should not find evidence in them that any children are members of peer groups, and therefore in each case P should be near 0. These results can be used to answer two questions about SCM. First, how often does SCM yield false positives when applied in different types of classroom settings (i.e. how often is $P > 0$)? We find that across all settings $P > 0$ in 81.2% of the 1000 simulated datasets, and therefore that SCM *frequently* yields false positives in a range of classroom settings. Second, how severe are the false positives when applied in this a setting like this? Among datasets where SCM identfied peer groups, it assigned nearly two-thirds (M = 0.63, SD = 0.30, min = 0.075, max = 1) of children to a peer group, when in fact the recall matrix contains no evidence of peer groups because it is random.

Table 2 reports the results of a regression predicting P as a function of the classroom characteristics that we varied in these datasets, as well as the range of these classroom characteristics across the 1000 simulated datasets. We find that in random data SCM assigns more children to peer groups (i.e. yields more false positives) when the classroom is larger, when children are more likely to be reported as group members, when some children are highly salient, and when some reported groups are large. We also find that SCM assigns fewer children to peer groups when participating children provide more group reports. These estimates are concerning because they highlight that whether SCM assigns a child to a peer group is not based only on patterns in the recall matrix, but is also driven by unrelated characteristics of the data. Moreover, some of the characteristics associated with false positives are precisely the challenges for which SCM was developed to overcome: larger settings where direct network data collection is impractical, and limited numbers of group reports due to low rates of parental and child consent to participate.

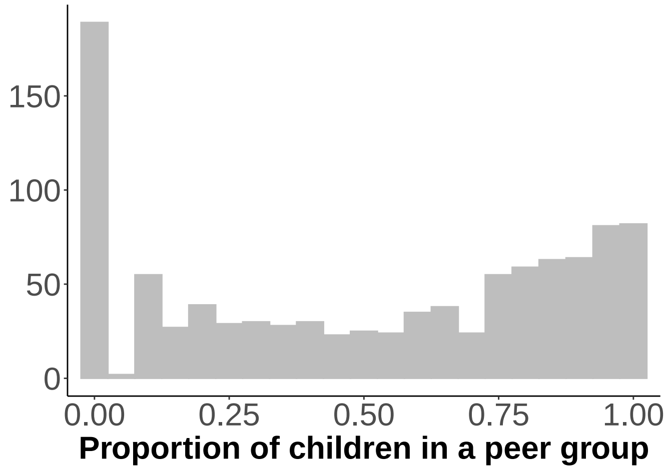

Figure 4

False positives applying SCM in classrooms with varying characteristics.

Study 4: Backbone extraction and community detection as an alternative

Methods

Studies 2 and 3 suggest that SCM has a high risk of false positives, identifying children as members of peer groups even in random data, under a wide range of circumstances. As we discuss in the background section, there are many different ways to measure peer groups that do not rely on peer-reported group data, and therefore do not require SCM. However, in this final study we explore one potential alternative to SCM for identifying peer groups from this type of data: backbone extraction and community detection.

Backbone extraction methods offer an alternative way to transform a recall matrix **R** into a binary peer network **N** (Z. Neal, 2014). Like SCM, these methods begin by transforming data such as a recall matrix (known as a *bipartite* matrix) into a co-occurrence matrix (known as a bipartite *projection*). However, unlike SCM, they apply a statistical test to these transformed data. The co-occurrence matrix contains, for each pair of children, the number of times they were reported by their peers to be members of the same peer group. Backbone extraction statistical tests are designed to determine whether this value (i.e. their number of group co-occurrences) exceeds the value that would be expected at random. If two children are reported by their peers to be members of the same peer group *more often than would be expected at random*, this is interpreted as evidence that these two children likely socially interact. Repeating this statistical test for each pair of children allows a backbone network of inferred social interactions to be derived. Backbone extraction methods differ in how they determine the number of group co-occurrences that would be expected at random (i.e. the null model), but here we focus on the Stochastic Degree Sequence Model (SDSM) because it is fast, well-documented, and easy to compute using the `R` `backbone` package (Domagalski, Neal, & Sagan, 2019, 2020).

Community detection methods offer an alternative way to identify peer groups from a binary peer network. These

Classroom characteristic	b	se	β	min	max
Intercept	0.252	0.043	—	—	—
Number of children	0.021	0.001	0.413	15	40
Number of group reports	-0.008	0.000	-0.880	15	200
Probability of nomination	0.013	0.001	0.189	10	45
Skew in nominations	0.100	0.015	0.115	-1.77	1.99
Skew in reported group size	0.095	0.016	0.114	-0.52	2.37

R^2 = 0.7147; P-values are not presented because the datasets are simulated.

Table 2

Impact of classroom characteristics on the number of false positives from SCM.

methods aim to identify cohesive groups in a network such that the majority of relationships are located within group and few relationships are located between groups. There are several methods for identifying these groups, including methods that allow children to have multiple group memberships, and methods that allow groups to have fuzzy boundaries (Fortunato, 2010). In this study we use the `igraph` package's `cluster_optimal()` function to identifies the *optimal* way to assign children to groups that maximizes ties within groups and minimizes ties between groups (Csardi, Nepusz, et al., 2006). Like SCM, this method identifies groups with distinct rather than fuzzy boundaries, but unlike SCM it requires group memberships to be mutually exclusive. While this is an important difference, in practice it may play a limited role because most studies using SCM already focus only on each child's one primary peer group (Berger & Rodkin, 2012; Chung-Hall & Chen, 2010; Zarbatany et al., 2019; Zhao et al., 2016).

In this study, we repeat studies 1 - 3 using a combination of backbone extraction and community detection (BE-CD) to identify peer groups rather than SCM. Before turning to the results, we briefly illustrate how BE-CD can be used in R. The first time BE-CD is used, the `backbone` and `igraph` packages must be installed in R by typing:

```
install.packages("backbone")
install.packages("igraph")
```

These two packages must be loaded so that R can use them by typing:

```
library(backbone)
library(igraph)
```

If the recall matrix is stored as a CSV file called `recall.csv`, then the BE-CD approach to identifying peer groups involves typing:

```
R <- read.csv(recall.csv)
N <- sdsm(R)
N <- backbone.extract(N, signed=F)
N <- graph_from_adjacency_matrix(N,
     diag=F, mode="undirected")
N.groups <- cluster_optimal(N)
```

This series of five commands imports the recall matrix data, conducts the SDSM statistical test, extracts the peer network, converts the network into a form that `igraph` can understand, and identifies peer groups. After these commands, the results can be examined using:

```
membership(N.groups)
plot(N.groups, N)
```

The first command will show the peer group membership of each child, while the second command will plot the peer network and show the boundaries of the peer groups.

Results

Figure 5A replicates study 1 by using BE-CD to identify peer groups in the benchmark data described by Cairns and Cairns (1994) (c.f. Figure 2). The identified peer groups almost perfectly match those identified by SCM, and those described by Cairns and Cairns (1994) from their direct observations. However, there are two exceptions: Heather and Ken are not assigned to peer groups. These exceptions offer an opportunity to compare the tuning accuracy of SCM and BE-CD by considering whether the recall matrix contains sufficient evidence to believe Heather and Ken are members of peer groups (consistent with SCM), or insufficient evidence than Heather and Ken are members of peer groups (consistent with BE-CD). Peer reports about Heather's hanging out behaviors were mixed: Four children reported that Heather was a member of the larger group of girls only, another four reported she was a member of the smaller group only, and no one reported she was a member of both groups. SCM views this as sufficient evidence to conclude that Heather is a member of both groups, while BE-CD views this evidence as too mixed to conclude she is a member of either group. Ken offers a similarly ambiguous case: only three children reported about Ken's hanging around behavior: two reported he hangs around with the larger group of boys, while one reported he hangs around with the smaller group of boys. SCM views this as sufficient evidence to definitively conclude Ken is linked to all the boys in the larger group, and none in the smaller group, while BE-CD again finds the evidence too mixed to draw a conclusion. We are unable to determine whether SCM or BE-CD is 'right,' but based on Cairns and Cairns (1994) description of their observations and on the groups reported by the children, both seem plausible.

Figure 5B replicates study 2 by using BE-CD to identify peer groups in 1000 randomized versions of Cairns and Cairns (1994) data (c.f. Figure 3). These results can be used to ask: how often does BE-CD yield false positives when applied in a classroom setting like that originally observed by Cairns and Cairns (1994) (i.e. how often is $P > 0$)? We find that BE-CD yields false positives in only 5 of the 1000 datasets. Among those few cases where it did yield false positives, they were not severe, assigning only 11.5% of children to a peer group when the correct value is 0%.

Figure 5C replicates study 3 by using BE-CD to identify peer groups in 1000 classrooms with varying characteristics (c.f. Figure 4). Again, these results can be used to ask: how often does BE-CD yield false positives when applied in a classroom setting like that originally observed by Cairns and Cairns (1994) (i.e. how often is $P > 0$)? We find that BE-CD yields false positives in only 10 of the 1000 datasets. Among those cases where it did yield false positives, they were not severe, assigning between 7.5% and 13.6% of children to a peer group when the correct value is 0%.

Discussion

Social cognitive mapping (SCM) is a method for identifying peer groups from peer report data. As formalized by Cairns and Cairns (1994) and implemented in SCM 4.0 (Leung & Alston, 1998), it is the most common method for identifying peer groups in developmental psychology, and is widely used in other fields including school psychology, social psychology, special education, and substance use. Our findings suggest one reason that SCM has enjoyed such widespread use: researchers wish to identify peer groups, and SCM *always* finds evidence of them. However, our findings also demonstrate why SCM is problematic: it will find evidence for peer groups *even when such evidence does not exist*.

In study 1, we found that SCM can identify peer groups that are known from direct observation to actually exist (true positives). However, in studies 2 and 3, we found that SCM also frequently identifies peer groups that are known *not* to exist (false positives). The results from study 3 also demonstrates that the severity of false positives is greatest in those settings where SCM was designed to be used: larger classrooms with lower participation rates. Finally, in study 4, we introduced backbone extraction and community detection as an alternative to SCM, and found that it has a similar ability as SCM to detect true positives, but a much lower risk than SCM of detecting false positives.

Based on our detailed review of SCM, the associated SCM 4.0 program, and the results of these four studies, we offer four recommendations to developmental psychologists and others wishing to identity peer groups from peer report data. First, because SCM as implemented in the SCM 4.0 program has a very high risk of false positives and because key parts of the SCM 4.0 are undocumented, *researchers should not use the SCM 4.0 program*. Second, and for the same reasons, findings about peer groups reported in *papers using SCM 4.0 should be viewed with caution*. Third, because multiple variants of SCM exist (e.g., Cairns & Cairns, 1994; Gest et al., 2007; Kindermann, 1993), *researchers using SCM should be explicit and detailed about the exact procedures they employ*, including reporting the cut-off threshold and the method for identifying peer groups from a binary network. Finally, although in 1994 "the scientific assessment of peer groups [was limited] by a gap in methods available for social network analysis" (Cairns & Cairns, 1994, p. 100), this is no longer the case. Therefore, when seeking to identify peer groups from peer report data, *researchers should consider using well-documented and statistically-informed network analytic methods*. In this paper we have illustrated how backbone extraction and community detection (BE-CD) might be used, however this is only one example among many alternatives.

Studies of peer groups identified via SCM, and in particular SCM 4.0, remain common in the developmental literature. However, SCM has an unacceptably high rate of false positives, casting doubt on whether the peer groups it identifies actually exist. Because understanding peer groups remains essential for understanding a wide range of developmental processes, developmental researchers must adopt alternative methods for identifying peer groups.

Contributions: ZN Developed the research design and performed the analysis, ZN and RD Developed the backbone package, ZN and JWN drafted the initial manuscript, All authors

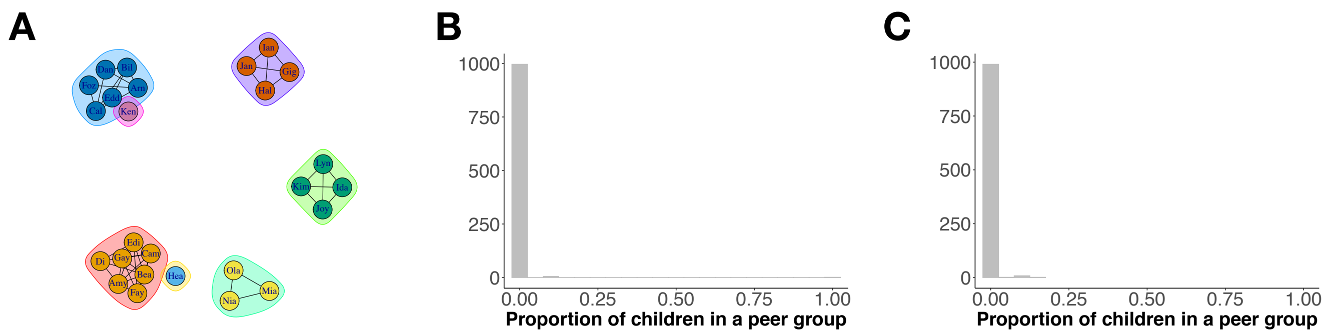

Figure 5

*Identifying peer groups using backbone extraction and community detection (BE-CD). (**A**) Applied to Cairns and Cairns (1994) 7^{th} grade classroom; (**B**) Applied to 1000 simulated 7^{th} grade classrooms; (**C**) Applied to 1000 classrooms with varying characteristics.*

revised the manuscript.

Funding information: This work was supported in part by National Science Foundation awards #1851625 & #2016320.

Competing interests: The authors declare they have no competing interests.

Data accessibility statement: All data and code necessary to replicate these analyses is available at https://osf.io/txgph.

Earlier versions: Earlier versions of this manuscript were posted at https://arxiv.org/abs/1911.05703 and https://psyarxiv.com/yfmzd/

Appendix A: Sensitivity analysis of similarity thresholds

The social cognitive mapping method, as described by Cairns and Cairns (1994) and as implemented in SCM 4.0 by Leung and Alston (1998), defines two children as connected in a binary peer network if the Pearson correlation coefficient between their group co-occurrence profiles is greater than or equal of 0.4 (see equation 3). This threshold value cannot be changed when using the SCM 4.0 program, and therefore is likely always the value used by researchers using SCM to identify peer groups.

In practice, researchers using SCM 4.0 have not been able to specify a different similarity threshold. Nonetheless, in this appendix we explore whether SCM's false positive rate is sensitive to the similarity threshold value and whether using a different value could reduce SCM's false positive rate. Because the similarity threshold in the SCM 4.0 program is fixed at 0.4 and cannot be changed, we cannot use the SCM 4.0 program to conduct this sensitivity analysis, and instead must develop our own code to emulate it. However, as we explain above, "we do not know exactly how SCM 4.0 identifies peer groups from a peer network" in the fifth and final step of the SCM process. This step is performed by SCM 4.0, but is not accurately described in any documentation. We believe the following R code emulates SCM 4.0, and therefore use it for the purposes of this sensitivity analysis:

```
R <- read.csv(recall.csv)
C <- R %*% t(R)
S <- cor(C)
N <- (S >= .4) + 0
N <- graph_from_adjacency_matrix(N,
     diag=F, mode="undirected")
N.groups <- fastgreedy.community(N)
```

The first four lines exactly duplicate the method implemented in SCM 4.0. The fifth line converts the resulting peer network to an iGraph object, while the sixth line uses the `igraph` package's `fastgreedy.community()` function to assign children to groups (Csardi et al., 2006). The 'fastgreedy' function yields an fast approximation of the `cluster_optimal()` function we recommend in study 4, and is used here to reduce the computational time for this sensitivity test. Both of these approaches to identifying peer groups in a network differ from the unknown method implemented in SCM 4.0, but both yield results that are nearly identical to those produced by SCM 4.0 in studies 1-3.

Using this code, we first repeated study 1, applying SCM to benchmark data with known peer groups (Cairns & Cairns, 1994), but varying the value of the similarity threshold from

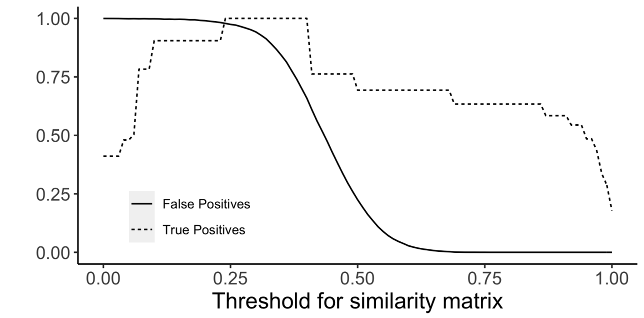

Figure 6

Sensitivity of true and false positives to similarity thresholds.

0 to 1 in increments of 0.01. For each value of the threshold, we computed the proportion of children that were assigned to the 'correct' peer group to which they are known by direct observation to belong. The dashed line in Figure 6 plots this proportion, which measures SCM's true positive rate, and which should ideally be near 1. We observe that SCM's ability to detect true positives in these data is greatest when the threshold is between about 0.25 and 0.45. When a lower threshold is used, the binary peer network is denser, leading all children to be assigned to a single peer group, which is incorrect. Conversely, when a higher threshold is used, the binary peer network is sparser, leading all children to be assigned to unique peer groups, which is also incorrect.

We then repeated study 2, applying SCM to the 1000 random recall matrices, again varying the value of the similarity threshold from 0 to 1 in increments of 0.01. For each value of the threshold, we computed the mean of our statistic of interest, P, over the 1000 random recall matrices. The solid line in Figure 6 plots this value, which measures SCM's false positive rate, and which ideally should be near 0. We observe that the false positive rate is sensitive to the value similarity threshold value. Specifically, the false positive rate is lowest when a high threshold is used. This occurs because a high threshold yields a parse peer network and therefore few peer group assignments.

Although a lower false positive rate can be achieved by using a higher similarity threshold, the results of this sensitivity test suggest that this is not an appropriate solution to the challenges identified in studies 2 and 3. Specifically, although using a larger threshold will indeed reduce SCM's false positive rate, it does so at the expense of also reducing its true positive rate.

References

Ahn, L.-H., & Shin, M.-S. (2011). A study on the relevance of social cognitive map analysis to peer subgrouping of elementary school students. *Korean Journal of Elementary Counseling*, *10*, 151–165.

Avramidis, E. (2010). Social relationships of pupils with special educational needs in the mainstream primary class: peer group membership and peer-assessed social behaviour. *European Journal of Special Needs Education*, *25*, 413–429. https://doi.org/10.1080/08856257.2010.513550

Bacete, F., & Perrin, G. (2013). Social cognitive maps. un método para identificar los grupos sociales en contextos naturale. *Psychosocial Interventions*, *22*, 61–70. https://doi.org/10.5093/in2013a8

Berger, C., & Rodkin, P. C. (2012). Group influences on individual aggression and prosociality: Early adolescents who change peer affiliations. *Social Development*, *21*(2), 396–413. https://doi.org/10.1111/j.1467-9507.2011.00628.x

Birkett, M., & Espelage, D. (2015). Homophobic name-calling, peer-groups, and masculinity: The socialization of homophobic behavior in adolescence. *Social Development*, *24*, 184–205. https://doi.org/0.1111/sode.12085

Breiger, R. (1974). The duality of persons and groups. *Social Forces*, *53*, 181-190. https://doi.org/10.2307/2576011

Breiger, R., Boorman, S., & Arabie, P. (1975). An algorithm for clustering relational data with application to social network analysis and comparison with multidimensional scaling. *Journal of Mathematical Psychology*, *12*, 328–383. https://doi.org/10.1016/0022-2496(75)90028-0

Brown, B. B., Mounts, N., Lamborn, S. D., & Steinberg, L. (1993). Parenting practices and peer group affiliation in adolescence. *Child development*, *64*(2), 467–482.

Bukowski, W., Laursen, B., & Rubin, K. (2018). Peer relations: Past, present, and promise. *Handbook of Peer Interactions, Relationships, and Groups*, 3–20.

Cairns, R., & Cairns, B. (1994). *Lifelines and risks: Pathways of youth in our time*. Cambridge University Press.

Cairns, R., Cairns, B., & Neckerman, H. (1989). Early school dropout: Configurations and determinants. *Child Development*, *60*, 1437–1452. https://doi.org/10.1111/j.1467-8624.1989.tb04015.x

Cairns, R., Cairns, B., Neckerman, H., Gest, S., & Gariépy, J. (1988). Social networks and aggressive behavior: Peer support or peer rejection? *Developmental Psychology*, *24*, 815–823. https://doi.org/10.1037/0012-1649.24.6.815

Cairns, R., Perrin, J., & Cairns, B. (1985). Social structure and social cognition in early adolescence: Affiliative patterns. *Journal of Early Adolescence*, *5*, 339–355. https://doi.org/10.1177/0272431685053007

Cappella, E., Neal, J., & Sahu, N. (2012). Children's agreement on classroom social networks: Cross-level predictors in urban elementary schools. *Merrill-Palmer Quarterly*, *58*, 285–313. https://doi.org/10.1353/mpq.2012.0017

Chung-Hall, J., & Chen, X. (2010). Aggressive and prosocial peer group functioning: Effects on children's social, school, and psychological adjustment. *Social Development*, *19*, 659–680. https://doi.org/10.1111/j.1467-9507.2009.00556.x

Csardi, G., Nepusz, T., et al. (2006). The igraph software package for complex network research. *InterJournal, complex systems*, *1695*(5), 1–9.

Domagalski, R., Neal, Z., & Sagan, B. (2019). `backbone`: Extracts the backbone from weighted graphs [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=backbone

Domagalski, R., Neal, Z., & Sagan, B. (2020). backbone: An r package for extracting the backbone of bipartite projections. *arXiv:1912.12779[cs.SI]*.

Espelage, D., Holt, M., & Henkel, R. (2003). Examination of peer-group contextual effects on aggression during early adolescence. *Child Development*, *74*, 205–220. https://doi.org/10.1111/1467-8624.00531

Fandrem, H., Ertesvag, S., Strohmeier, D., & Roland, E. (2010). Bullying and affiliation: A study of peer groups in native and immigrant adolescents in norway. *European Journal of Developmental Psychology*, *7*, 401–418. https://doi.org/10.1080/17405620802335075

Farmer, T., & Cairns, R. (1991). Social networks and social status in emotionally disturbed children. *Behavioral Disorders*, *16*, 288-298. https://doi.org/10.1177/019874299101600404

Farmer, T., Hall, C., Petrin, R., Hamm, J., & Dadisman, K. (2010). Evaluating the impact of a multicomponent intervention model on teachers' awareness of social networks at the beginning of middle school in rural communities. *School Psychology Quarterly*, *25*, 94–106. https://doi.org/10.1037/a0020147

Farmer, T., Stuart, C., Lorch, N., & Fields, E. (1993). The social behavior and peer relations of emotionally and behaviorally disturbed students in residential treatment: A pilot study. *Journal of Emotional and Behavioral Disorders*, *1*, 223–234. https://doi.org/10.1177/106342669300100404

Fortunato, S. (2010). Community detection in graphs. *Physics reports*, *486*(3-5), 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Gest, S. (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. https://doi.org/10.1046/j.1467-9507.2006.00339.x

Gest, S., Farmer, T., Cairns, B., & Xie, H. (2003). Identifying children's peer social networks in school classrooms: Links between peer reports and observed interactions. *Social Development*, *12*, 513–529. https://doi.org/10.1111/1467-9507.00246

Gest, S., Moody, J., & Rulison, K. (2007). Density or distinction? the roles of data structure and group detection methods in describing adolescent peer groups. *Journal of Social Structure*, *6*.

Gifford-Smith, M., & Brownell, C. (2003). Childhood peer relationships: social acceptance, friendships, and peer networks. *Journal of School Psychology*, *41*, 235–284. https://doi.org/10.1016/S0022-4405(03)00048-7

Haynie, D. L. (2001). Delinquent peers revisited: Does network structure matter? *American journal of sociology*, *106*(4), 1013–1057.

Hoffman, A., Hamm, J., & Farmer, T. (2015). Teacher attunement: Supporting early elementary students' social integration and status. *Journal of Applied Developmental Psychology*, *39*, 14-23. https://doi.org/10.1016/j.appdev.2015.04.007

Howe, C. (2010). *Peer groups and children's development*. West Sussex,UK: Wiley.

Kindermann, T. (1993). Natural peer groups as contexts for individual development: The case of children's motivation in school. *Developmental Psychology*, *29*, 970–977. https://doi.org/10.1037/0012-1649.29.6.970

Kindermann, T., & Gest, S. (2018). The peer group: Linking conceptualizations, theories, and methods. In *Handbook of peer interactions, relationships, and groups* (pp. 84–105). Guilford Press.

Krackhardt, D. (1987). Cognitive social structures. *Social Networks*, *9*, 109–134. https://doi.org/10.1016/0378-8733(87)90009-8

Leung, M.-C., & Alston, A. (1998). Scm 4.0 [computer software]. *Center for Developmental Science, University of North Carolina*.

Levina, J., & Ivanova, N. (2012). The self-concept of adolescents with different social status in peer cliques. *Baltic Journal of Psychology*, *13*, 98–112.

Neal, J. (2008). "kracking" the missing data problem: Applying krackhardt's cognitive social structures to school-based social networks. *Sociology of Education*, *81*, 140–162. https://doi.org/10.1177/003804070808100202

Neal, J. (2020). *Systematic review of social network methods in developmental psychology dataset.* figshare. [dataset]. https://doi.org/10.6084/m9.figshare.11728863.v1

Neal, J. (in press). A systematic review of network methods in high impact developmental psychology journals. *Social Development*. https://doi.org/10.1111/sode.12442

Neal, J., & Neal, Z. (2013). The multiple meanings of peer groups in social cognitive mapping. *Social Development*, *22*, 580–594. https://doi.org/10.1111/j.1467-9507.2012.00656.x

Neal, J. W., Neal, Z. P., & Cappella, E. (2016). Seeing and being seen: Predictors of accurate perceptions about classmates' relationships. *Social Networks*, *44*, 1 - 8. https://doi.org/10.1016/j.socnet.2015.07.002

Neal, Z. (2014). The backbone of bipartite projections: Inferring relationships from co-authorship, co-sponsorship, co-attendance, and other co-behaviors. *Social Networks*, *39*, 84–97. https://doi.org/10.1016/j.socnet.2014.06.001

Neal, Z., & Neal, J. (2013). Opening the black box of social cognitive mapping. *Social Development*, *22*, 604–608. https://doi.org/10.1111/j.1467-9507.2012.00668.x

Pijl, S., Koster, M., Hannink, A., & Stratingh, A. (2011). Friends in the classroom: a comparison between two methods for the assessment of students' friendship network. *Social Psychology of Education*, *14*, 475–488. https://doi.org/10.1007/s11218-011-9162-2

Radovic, D., Black, L., Salas, C., & Williams, J. (2017). Being a girl mathematician: Diversity of positive mathematical identities in a secondary classroom. *Journal for Research in Mathematics Education*, *48*, 434–464.

Rodkin, P., Farmer, T., Pearl, R., & van Acker, R. (2006). They're cool: Social status and peer group supports for aggressive boys and girls. *Social Development*, *15*, 175–204. https://doi.org/10.1046/j.1467-9507.2006.00336.x

Rubin, K., Bukowski, W., & Bowker, J. (2015). Children in peer groups. In *Handbook of child psychology and developmental science* (Vol. 4, pp. 175–222). Wiley.

Ryan, A. (2001). The peer group as a context for the development of young adolescent motivation and achievement. *Child Development*, *72*, 1135–1150. https://doi.org/10.1111/1467-8624.00338

Shannon, C., & Weaver, W. (1963). *The mathematical theory of communication*. University of Illinois Press.

Sheppard, C., Golonka, M., & Costanzo, P. (2012). Evaluating the impact of a substance use intervention program on the peer status and influence of adolescent peer leaders. *Prevention Science*, *13*, 75–85. https://doi.org/10.1007/s11121-011-0248-z

Shi, B., & Xie, H. (2014). Moderating effects of group status, cohesion, and ethnic composition on socialization of aggression in children's peer groups. *Developmental Psychology*, *50*, 2188–2198. https://doi.org/10.1037/a0037177

Strona, G., Nappo, D., Boccacci, F., Fattorini, S., & San Miguel-Ayanz, J. (2014). A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications*, *5*, 4114. https://doi.org/10.1038/ncomms5114

Watts, D. J. (2008). *Six degrees: The science of a connected age*. W. W. Norton.

Wolfer, R., Bull, H., & Scheithauer, H. (2012). Social integration in youth: Insights from a social network perspective. *Group Dynamics: Theory, Research, and Practice*, *16*, 138–147. https://doi.org/10.1037/a0024765

Zarbatany, L., Ellis, W. E., Chen, X., Kinal, M., & Boyko, L. (2019). The moderating role of clique hierarchical organization on resource control by central clique members. *Journal of youth and adolescence*, *48*(2), 359–371. https://doi.org/10.1007/s10964-018-0972-9

Zhao, S., Chen, X., Ellis, W., & Zarbatany, L. (2016). Affiliation with socially withdrawn groups and children's social and psychological adjustment. *Journal of Abnormal Child Psychology*, *44*, 1279–1290. https://doi.org/0.1007/s10802-015-0120-x