

Talking and learning physics: Predicting future grades from network measures and Force Concept Inventory pretest scores

Jesper Bruun

Department of Science Education, University of Copenhagen, Copenhagen, Denmark

Eric Brewster

Department of Physics, Florida International University, Miami, Florida 33199, USA
(Received 19 July 2012; revised manuscript received 8 April 2013; published 31 July 2013)

The role of student interactions in learning situations is a foundation of sociocultural learning theory, and social network analysis can be used to quantify student relations. We discuss how self-reported student interactions can be viewed as processes of meaning making and use this to understand how quantitative measures that describe the position in a network, called centrality measures, can be understood in terms of interactions that happen in the context of a university physics course. We apply this discussion to an empirical data set of self-reported student interactions. In a weekly administered survey, first year university students enrolled in an introductory physics course at a Danish university indicated with whom they remembered having communicated within different interaction categories. For three categories pertaining to (1) communication about how to solve physics problems in the course (called the PS category), (2) communications about the nature of *physics concepts* (called the CD category), and (3) social interactions that are not strictly related to the content of the physics classes (called the ICS category) in the introductory mechanics course, we use the survey data to create networks of student interaction. For each of these networks, we calculate centrality measures for each student and correlate these measures with grades from the introductory course, grades from two subsequent courses, and the pretest Force Concept Inventory (FCI) scores. We find highly significant correlations ($p < 0.001$) between network centrality measures and grades in all networks. We find the *highest* correlations between network centrality measures and *future* grades. In the network composed of interactions regarding problem solving (the PS network), the centrality measures *hide* and PageRank show the highest correlations ($r = -0.32$ and $r = 0.33$, respectively) with future grades. In the CD network, the network measure *target entropy* shows the highest correlation ($r = 0.45$) with future grades. In the network composed solely of non-content related social interactions, these patterns of correlation are maintained in the sense that these network measures show the highest correlations and maintain their internal ranking. Using hierarchical linear regression, we find that a linear model that adds the network measures *hide* and *target entropy*, calculated on the ICS network, significantly improves a base model that uses only the FCI pretest scores from the beginning of the semester. Though one should not infer causality from these results, they do point to how social interactions in class are intertwined with academic interactions. We interpret this as an integral part of learning, and suggest that physics is a robust example.

DOI: [10.1103/PhysRevSTPER.9.020109](https://doi.org/10.1103/PhysRevSTPER.9.020109)

PACS numbers: 01.40.Fk, 01.40.Ha

I. INTRODUCTION

The role of social interactions in learning is well established and has been a foundation of learning theory [1,2]. Within physics education research (PER), studies are emerging using social network analysis (SNA) of student interactions, and other types of network analysis in the learning of physics are emerging [3–7]. In the broader field of education, SNA has been used to investigate, for example, how social networks relate to educational

change [8], leadership at schools [9–11], and the relation between student online interactions and academic performance [12–14].

We continue this line of work, and extend it by using self-reported student interactions in an introductory physics course as one variable in a model to predict students' grades at the University of Copenhagen in two subsequent courses, one in physics and one in mathematics. In order to model how students' grades depend on student interactions we surveyed students at the University of Copenhagen weekly by asking them to name with whom they remembered having communicated about solving problems in the physics course, about the meaning of concepts, or with whom they remembered having communicated about social matters *not* related to physics as a subject during

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

scheduled teaching or learning activities. (See Sec. IVA 1 for a discussion about these three categories and why they were chosen.) Using remembered communication as a proxy for interaction,¹ we used the survey data to make three different networks of student interaction. In order to analyze these networked data, we first conceptualize the underlying assumptions about networks and how the analysis can be interpreted. Our analysis then predicts students' grades by using measures calculated from the network data that describe how the individual student is related to the rest of the network via the inferred interactions. We argue that the network perspective presented in this article has unique potential for quantitative analysis of learning as a socially embedded phenomenon. This study of a Danish introductory physics course at the University of Copenhagen was undertaken to investigate the roles social interactions play in learning physics. This study employs social network analysis, which has been used in PER to show the growth and variation of student-student interactions in instruction [15,16], to identify the roles that gender and ethnicity play in an informal network of students in a physics learning center [3], to model student retention [6,7], and to model interactions between students and teachers in nonphysics learning environments [9,17,18]. Bodin recently used network analysis to document changes in student epistemic networks [19] after engaging in numerical modeling of a physics problem [4]. This study differs from the others in that it uses network analysis to predict students' grades in a subsequent course.

Network analysis is a collection of analytic techniques which can be used to visualize, quantify, and test hypotheses based on relationships between entities within a network. In network terminology, the entities are called nodes and the relationships are called links. Visually, we draw nodes as circles and links as arrows. The correspondence between entities nodes and relations links is not necessarily one to one. This means that a node may only represent aspects of the entity, and a link only an aspect of a relation. In our study, the nodes represent students, and links between students signify that one student remembers having communicated with another student in a physics learning context. We see nodes as placeholders representing different aspects of students. In this study, we relate aspects of student performance, grades, and Force Concept Inventory (FCI) pretest scores to measures describing how central they are from a network point of view. The collection of grades, test scores, and the measures of centrality are what we call *node attributes*. From here, we will use the terms *nodes* or *students* for the entities and *links*, *communications*, or *interactions* for the

relations, but we will return to what other entities might be represented in networks in Sec. VI.

The students in this study were given a weekly 11-question survey that asked them to select the names of students with whom they remembered having had interacted in a variety of contexts, including the areas of problem solving, concept communication, and in-class social interactions. The students selected the names from a roster of all students enrolled in the course. The responses to these questions indicate an interaction between students, which, in our analytical framing, are the linkages between the nodes. We collected attribute data including FCI pretest scores and grades in the introductory physics course. Both FCI scores and grades in the introductory physics course have been shown to be predictive of grades in subsequent physics classes [20,21]. We also collected student grades from two subsequent courses, a physics course on rotational Newtonian mechanics and a mathematics course in linear algebra. The collection of nodes, links, and the attributes of the nodes are the constituent parts of any network. Thus, in our study, the students, along with the interactions they reported and the attribute data we collected, constitute the networks. These networks and node-level attribute data were analyzed to address the research question in this article: How do node centrality measures improve the prediction of grades in subsequent physics classes for three different types of networks (problem solving, conceptual discussion, and in-class social communication) over other attributes such as FCI scores or grades in a previous course?

Network perspectives on data are unique in that the data are primarily relational. Thus, we conceptualize students as situated within the learning context, which not only includes physics classes, but also other interactional and social settings. As network analysis is relatively new in physics education research, it is worthwhile to consider how data are collected and how one would interpret the meaning of these data. Further, the interpretation of the meaning of a link is an ongoing debate among social network analysis researchers [22]. In the following sections, we describe two theoretical perspectives that are employed for the interpretation of the data we have collected.

II. THEORETICAL FRAMEWORK

We are undertaking an analysis of learning that is dependent on the interactions of individuals within a network of learners. Underlying this methodology is a situated or participationist framework of learning [23,24]. This theoretical framework avers that learning is an ongoing process of transforming participation [25–27], and it has roots in the Vygotskian perspective [1]. Vygotskian theory holds that learning and development, especially development of language, are so interrelated that they should not be considered independently. Thus, studying learning should be a study of the interactions in which a student is involved.

¹In our conception, interaction is to be understood in a very broad manner. It could be a lengthy conversation about the nature of some idea, or it could be simply to look at each others' results after solving an end of chapter problem.

While Vygotsky primarily used clinical interviews or clinical student-teacher interactions, others studying situated cognition extended the analysis to interactions in naturalistic settings [28,29] to help account for the contextual or situated nature of interactions. Rogoff, Matusov, and White ([27], p. 390) describe the central premise of the participationist perspective on learning as “the idea that learning and development occur as people participate in the socio-cultural activities of their community, transforming their understanding, roles and responsibilities as they participate.” The participationist perspective on learning is particularly salient to this study, as we are investigating the effect of the interactions that students engage in within and around their physics class *vis-à-vis* predictors of their success in a subsequent pair of classes.

III. THEORY: NETWORK MEANING AND STRUCTURE

The analysis of social networks developed from quantitative sociology (that began in the 1920s—see Ref. [30]) as a way of investigating the structures within social groups. Taking a social network to represent a social group, researchers can use a variety of approaches to investigate these structures. One example is to use mathematical models on the social network to identify community structure [31,32] or to characterize the whole network in terms of large-scale structures [33]. We have adopted a second approach, which is to look at the structure of interactions within a network, focusing not on the larger structures of the network but on the individual node’s links to the rest of the network [34], which we term the *position* of the node. This latter approach, focusing on nodes and their positions, has been used extensively in social science (for numerous examples, see Ref. [30]) and general education research [8], but hardly at all in physics education research. Some rare and very recent examples can be found in the investigation of Brewe *et al.* [3] of participation in a network of physics learners, in the exploration by Forsman [6] of retention in physics and engineering programs based on complexity thinking, and in the study by Bruun [15] of upper secondary and first year university students’ social and cognitive networks. Particularly, Brewe *et al.* investigated the positional advantage or constraint that a student experiences due to that student’s centrality as a node within the network. In broad terms, a centrality measure gauges the relative importance of individual nodes in a network.² This is done by using the quantitative aspect of links between nodes (for example, how many times an interaction has occurred) to calculate how a node relates to its neighbors. Common centrality measures include degree and closeness, which we utilize in our analysis and describe in detail in the Supplemental Material [35].

²In social network analysis, centrality is often related to the notion of power.

Other centrality measures are probabilistic, but still describe the relative importance of a node within a network. We utilize three such centrality measures, PageRank, target entropy, and hide, which we also describe in the Supplemental Material [35].

One of the basic assumptions upon which network analysis is built is that interconnected nodes influence each other. The interpretation of any network measure or property rests on how one chooses to conceptualize the meaning of links and nodes. In other words, a basic understanding of a network shapes how we interpret centrality measures, node attributes, and correlations between them. McCormick *et al.* describe the mechanism by which nodes influence each other in a network by describing a link as representing a process of meaning making [36].³ From our point of view, this interpretation best applies to our study: we asked students to recall interactions, so links represent that the student has identified a relationship with another student. There is evidence that self-reported interactions are biased towards salient and recent events [37,38]. This means that links based on self-reports may inform researchers about the recent interactions students value and remember. In this process of meaning making, advice and ideas are interchanged and changed as students interact with each other. In this study we consider links to be interactions, during which students exchange and modify their ideas.

A. Measures of student centrality

Network theory offers several different centrality measures, all of which have different interpretations. In this section, we briefly describe the centrality measures that we use in this article. All of the centrality measures used in this article are probabilistic. When we sum up the number of interactions students reported during different weeks (see Sec. IVA 2), it becomes possible that student i mentioned student j more than once. As a result, the link between student i and student j becomes weighted. The weighting of links between students leads to probabilistic interpretations of centrality. We have placed formal mathematical definitions of these measures in the Supplemental Material [35] along with code, written in the statistical programming language R, that was used to generate these measures.

1. Degree

Degree is the simplest centrality measure, and thus the most widely used. The degree of a node is defined as the number of other nodes that are directly connected to it. Often degree is split into what is known as indegree and outdegree. Indegree is the number of links coming into a node and outdegree is the number of links going out of a node.

³McCormick *et al.* actually write a “process of transformation” ([36], p. 115), but later argue that meaning is the substance of that transformation process.

2. Strength

In networks such as ours, where one student can select another in multiple weeks, the links between nodes are said to be weighted (see Sec. IV A 2). The strength of a node can be considered the weighted analog to degree. Strength is the sum of weights attached to links belonging to a node. Often two types of strength are considered, instrength and outstrength. Instrength is the sum of the weighted links coming into a node and outstrength is the sum of the weighted links going out from a node.

3. PageRank

PageRank is an algorithm conveying the general idea that a node gains importance via links *to* the node and distributes importance via the links *from* the node. Thus, PageRank incorporates both local connections and global position. Brin and Page [39] developed PageRank to mathematically model the connections between different Web sites on the internet. As an example, CNN.com is a commonly visited site and it thus has a high centrality. When CNN.com provides a link to a less well-trafficked site, that linked site gains centrality by being linked to CNN.com. A student can earn a high PageRank by having many students naming him or her or by having a few students with a high PageRank name him or her. Thus, a student does not need to be popular to gain PageRank. PageRank is designed to reflect this notion of centrality.

4. Target entropy

Target entropy is a centrality measure based on the flow of information through a network. Target entropy gauges the predictability of the traffic around a node in a network [40], based on the assumption that nodes produce and disperse messages. Target entropy is a centrality measure that presumes that when messages pass through a node that node is more “important.” Interacting with many different people increases target entropy. The target entropy will be further increased if they in turn interact with many people who remember them. Thus, if a student is part of a network where people are involved in many interactions and remember their interactions, that student will end up with a high target entropy.

5. Hide

Hide is a measure of how difficult it is, on average, for any node in the network to “find” a particular node [40]. In our study, hide quantifies some of the constraints a student experiences from a network perspective, due to his or her lack of interaction with other students. A student who is, on average, difficult to locate within the network may not have the same access to ideas related to physics as a student who is easier to locate. In the interaction picture, the interpretation of hide is straightforward. With a high hide, a student has not participated in many interactions that other students can remember. In contrast, students with a low hide

have participated in many meaningful interactions so that they are easy to locate.

IV. METHODS

This study took place at the University of Copenhagen with 171 students taking their first university physics courses. The semester at this university is split into two blocks. In the first block students take two courses, an introductory mechanics course and an introductory mathematics course. In the second block, students take another two courses, an advanced mechanics course dealing with rotations and oscillations and a linear algebra course. The physics course in each block includes two 90 minute lectures each week with all students, two scheduled problem solving sessions each week, and one weekly laboratory session (3 hours). Students were divided into seven exercise classes of 20–30, who took the problem solving sessions and laboratory sessions together. Six of the seven exercise classes were physics majors, and the seventh consisted of other majors. The majority of students attended both blocks of the introductory physics course. FCI and network data were collected during the first block.

A. Data collection

The data collected for this study include student surveys as well as attribute data for individual students. Data were collected across the two blocks as shown in Fig. 1.

1. Survey data

Each week during the first block, students were asked to answer two online network surveys. Both surveys were administered during the laboratory section of the course. In each survey, students were allowed to select names from a drop-down menu which included the names of all students enrolled in the physics course. There was no limit on the number of names that could be selected. The links generated between students in this way are dependent on student recollection of interactions [37,38], understanding of the interaction categories [9], and the order in which the categories are presented [10]. Following Eagle *et al.*, students will have a tendency to remember salient interactions [38], and following Pustejovsky and Spillane [10], the

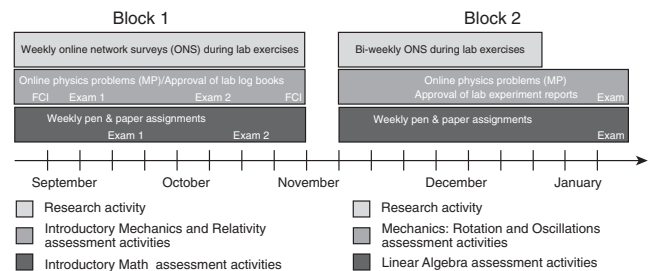


FIG. 1 (color online). Timeline of assessment activities during the first two blocks along with the research activity relevant to this study. The FCI assessments and the approvals of lab log books or reports are not part of the grade.

TABLE I. Survey categories that students used to name other students. Each week students chose whom they remembered having had interactions with in each category. These three categories were part of a larger survey consisting of in total 12 interaction categories.

Survey	Category	Question (translated from Danish)
1.1	Problem solving	We communicated about how to solve a task in physics. (How to perform calculations, what formulas are needed, how to read graphs and the like.)
1.2	Concept discussion	We communicated about understanding one or more physics concepts. (What current is, what the normal force is, how radioactivity works, and the like.)
2.1	In-class social	We communicated socially in connection with a lecture, problem solving session, or laboratory exercise.

order in which interaction categories are presented may affect the number of student names selected by a given student.

All surveys were administered in Danish, so questions have been translated for this article. Three questions were selected for this article and were part of a larger survey. In Table I, questions 1.1 and 1.2 probe for problem solving and conceptual discussion interactions, respectively, while question 2.1 probes for social interactions. We focused on student interactions with regards to problem solving, discussion of physics concepts, and social interactions during classes because the network centrality measures in these networks showed the greatest correlations with the variable we wanted to predict (see Sec. IV D). Using Forsman's framework [6], the first two interaction networks, problem solving and concept discussion, can be taken to reflect academic student interactions, which are related to success in physics courses. The third, in-class social, can be taken to reflect engagement in social interactions that are not strictly related to the content of the physics classes. The student interactions in each network form the basis for calculating different centrality measures. It should be noted that, because these are self-reported interactions, it is possible that students were biased toward listing their friends [41–43]. Three efforts to reduce this bias were used: first, the surveys were anonymous for everybody but the researchers, second, the questions do not imply that interactions should be ranked, and third, the formulation of the questions does not imply that one student is seeking advice from another. See Table I.

2. Transforming survey data to networks

The survey data were used to construct three networks each week, one based on the responses to each of the following questions: 1.1, 1.2, and 2.1. For each week, we construct a problem solving network, a concept discussion network, and an in-class social discussion network (see Fig. 2). When student i indicates that they remember having had an interaction with student j , there is a directed link, l_{ij} , from student i to student j . The link is unweighted, meaning that all existing links have weight $l_{ij} = 1$. For a single week, the problem solving, concept discussion, or in-class social discussion network is directed and

unweighted. The weekly networks were summed to create three networks that include the responses from all weeks. This means that the links between students became weighted links, with the maximum weight between node

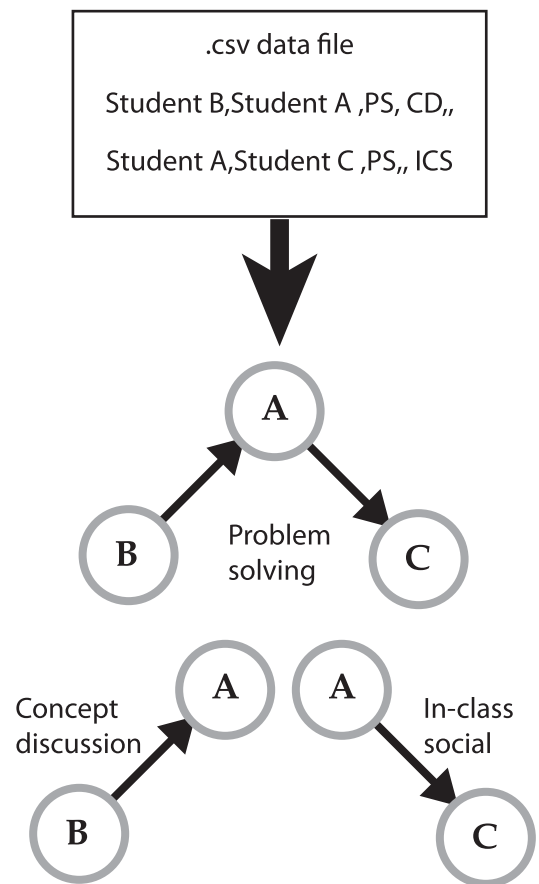


FIG. 2 (color online). Illustration of how the survey answers were transformed to networks. The survey software outputted the data as a comma separated variables (.csv) file. The first entry in a line is the selecting student, the second entry is the selected students. The following entries list the categories in which the students were selected. In this example, student B selected student A in the problem solving (PS) and concept discussion (CD) categories. In network terms B is the source of a connection and A is the target. A is also the source of a PS link to the target C, and an in-class social communication link also to C.

TABLE II. Summary of centrality measures, attributes, and grade variables used in this study.

Centrality measures	Symbol	Attributes	Symbol	Grade variable	Symbol
Instrength	s_{in}	FCI pre	FCI_{pre}	Sum of grades in block 2	SOG
Indegree	k_{in}	Mechanics grade in block 1	GR_{mech1}		
PageRank	PR				
Hide	\mathcal{H}				
Target entropy	T				

i and node j in any network weight $l_{ij} = 9$. Because some students did not attend every laboratory meeting, they did not complete the survey every week and thus not all students are included in each weekly analysis. Summing the weekly networks can be seen as analogous to a long exposure time in a CCD chip in photography. The summed-up networks hold more information about each “cell” (here student), because the instrument has been active for a longer time. However, the summed-up networks do not reveal variations that happen during the time of exposure. Furthermore, the summed-up networks are a way of quantifying connection strength between students; if student i selects student j in most or all weeks, then we infer that the connection is stronger and the interaction more frequent, from i 's point of view than if i selects j fewer times. By summing up, we include a larger sample of students and thus achieve a more representative network than any one of the weekly networks. At the same time, the summed-up networks give a more reliable picture of all the interactions during the nine weeks upon which to base network centrality calculations.

B. Attribute data

Though network perspectives on data assume that the data are relational and interdependent, each node within a network has characteristics, or attributes, which are independent of the other nodes. These can be used in conjunction with networked measures to constitute predictive outcomes. In this study, we used two attributes, Force Concept Inventory pretest scores (FCI_{pre}) and physics grades in the first course block (GR_{mech1}). FCI_{pre} scores have been shown to correlate with academic success in introductory courses [21]. Also, grades have been used to predict future grades [20], so we use these two attributes as benchmarks in our models. The data were provided by the physics department.

1. Force Concept Inventory pretest score

The Force Concept Inventory [44] was administered as a pre- and posttest in the first block of the course. We used the FCI_{pre} score attained during the first week of the first block of the course as a predictive measure in the sum of grades in the second block of the course. Using FCI_{pre} as a predictor variable in the regression analysis is in line with the findings of Henderson [21].

2. Grades

In the Danish system, grades are numeric values taken from the set $G = \{-3, 0, 2, 4, 7, 10, 12\}$. A grade of 2 or above means that the course is passed, while 12 is the maximum grade. Grades 4 and 7 are meant as average grades. A student achieving a -3 or 0 has failed the course and will need to retake it or pass a reexamination. We use the numeric grades in two ways. First, we use the mechanics grade in the first block as a baseline predictor attribute in the linear modeling component of the study. The possible values are the elements of G . Second, we use the sum of the physics and math grades (SOG) in the second block as the dependent variable in the linear modeling part of the study. Because this is a sum of two grades, the possible values for this grade are the elements of $G + G = \{g_i + g_j; g_i, g_j \in G\}$.⁴ The sum of grades variable is only applicable for students enrolled in both the rotations and oscillations course and the linear algebra course. Students who were only enrolled in one of these two courses were left out of the predictive calculations, but not from the centrality calculations on the networks from the introductory course.

C. Summary of variables

This study made use of a large number of variables, which include the centrality measures instrength, indegree, PageRank (PR), hide, and target entropy, attributes, and the dependent variable, the sum of grades in block 2. Table II summarizes these.

D. Correlations

The initial analysis of each network was to create a Pearson's r correlation matrix including each of the centrality measures, the attributes, and sum of grades in the second block. Initially, these correlations were carried out between the sum of grades in block 2 and centrality measures calculated on networks based on all the other interaction categories in the survey. However, we reduced the number of correlates after finding that centralities from the three networks for problem solving (PS), concept

⁴The reason for this cumbersome notation is that not all integers between -6 and 24 are possible when summing the two grades.

discussion (CD), and in-class social (ICS) communication showed the highest correlations with the SOG. By keeping only the highest correlates, we reduced the complexity of our data, even if many of the removed correlations were significant in terms of Pearson's r . Furthermore, the three networks investigated in this article—PS, CD, and ICS—reflect a large part of the activity related to learning physics with which students engage at the university level. We would expect them to engage with problem solving, to discuss conceptual meaning, and to interact socially in the context of a university education in physics. Because not all students participated in each network, each of the three networks include differing numbers of students. Further, students who were named in a category at some point during the two blocks of data collection but who did not name any other student during the first block were removed from the analysis. The number of removed students were $N_{PS} = 25$, $N_{CD} = 43$, and $N_{ICS} = 40$. All analyses were carried out in the R environment [45], primarily with the iGraph package [46] for analyzing network data. The scripts for running these analyses are included in the Supplemental Material [35]. These correlation matrices were used to identify centrality measures that held the greatest promise for predicting SOG, so all correlations with $p \geq 0.001$ were eliminated. This left five centrality measures as candidates to construct linear models for predicting SOG in the second block. To further understand the connections between variables, we employed bootstrapped hierarchical multiple linear regression to create models that predict the SOG. Such models add to our analysis by allowing us to make statistical inferences. Using bootstrapping, they take into account that in network statistics the independence of measures assumption is violated [47]. Thus the linear models in the next sections were calculated using bootstrap methods.

E. Hierarchical multiple regression on sum of grades

We used hierarchical multiple regression [3] to create models with the SOG in the second block as the dependent variable. The centrality measures and the attributes serve as independent variables within each of the three networks considered. Hierarchical multiple regression is a regression technique used to generate and compare the predictive models for a continuous dependent variable using different sets of independent variables. Independent variables are entered in a specific order according to logical or theoretical considerations. In this study, we entered variables in order of decreasing correlation with the SOG variable. A new predictive model is created with each new independent variable (or pair of variables) entered. These models were then compared using an F-test for the differences between correlations as described in Tabachnick and Fidell [48]. Again, the analyses were carried out in the R statistical programming language, this time using the boot package [49]. In addition, all models were compared with

the regression model using the mechanics grade in block 1. This comparison provided us with a baseline that was interpreted as “Does the regression model using centrality measures do as good of a job predicting the SOG as the grades in a previous block of the course?”

V. RESULTS: CENTRALITIES AND FUTURE GRADES

The results can be grouped into two sections. The first section shows correlations between attributes, centrality measures, and the grades from the first block. In the second section, we evaluate the predictive power of different linear combinations of attributes and centrality measures.

A. Centrality and attribute correlations

Having two different attribute variables, five centralities, and one outcome variable for each network, we had 84 possible correlations. In order to show the correlations graphically, identify significant correlations, and report differing degrees of freedom with each variable, we use a correlation network [50]. Correlation networks are diagrams that show the correlations between various measures. In a correlation network, two measures are connected if they are correlated and the strength of the link is taken to be proportional to the strength of the correlation. Figure 3 shows the correlation network for the measures in our study, picking out only significant correlations. The measures for each network are color coded, and the value in each node represents the number of degrees of freedom in the correlation. The values of the correlation coefficients are represented on the links. We have adjusted the layout manually, so for each variable the distance from the measures to the SOG is ranked. This means that since GR_{mech1} shows the highest correlation in each of the three distinct networks, it is closest to the SOG in the drawing.

The GR_{mech1} score correlates with the SOG with $r \approx 0.6$, reflecting that performance in the first block is a strong predictor of performance in the classes in the subsequent block. The FCI pretest score (FCI_{pre}) correlates with the SOG with a coefficient of $r \approx 0.4$. In each network, FCI_{pre} also correlates with the GR_{mech1} score.

1. Problem solving network correlations

In the PS network, only two network measures, PageRank and hide, correlated significantly with SOG. The hide (\mathcal{H}) correlates negatively with the SOG with a coefficient $r = -0.32$ indicating that the more difficult a student is to locate within the problem solving network, the lower the grades in the subsequent block of courses. The PageRank correlates with a high level of significance with both the SOG and with the GR_{mech1} , indicating that students who are more central in the problem solving network also tend to have higher grades in the second block of courses.

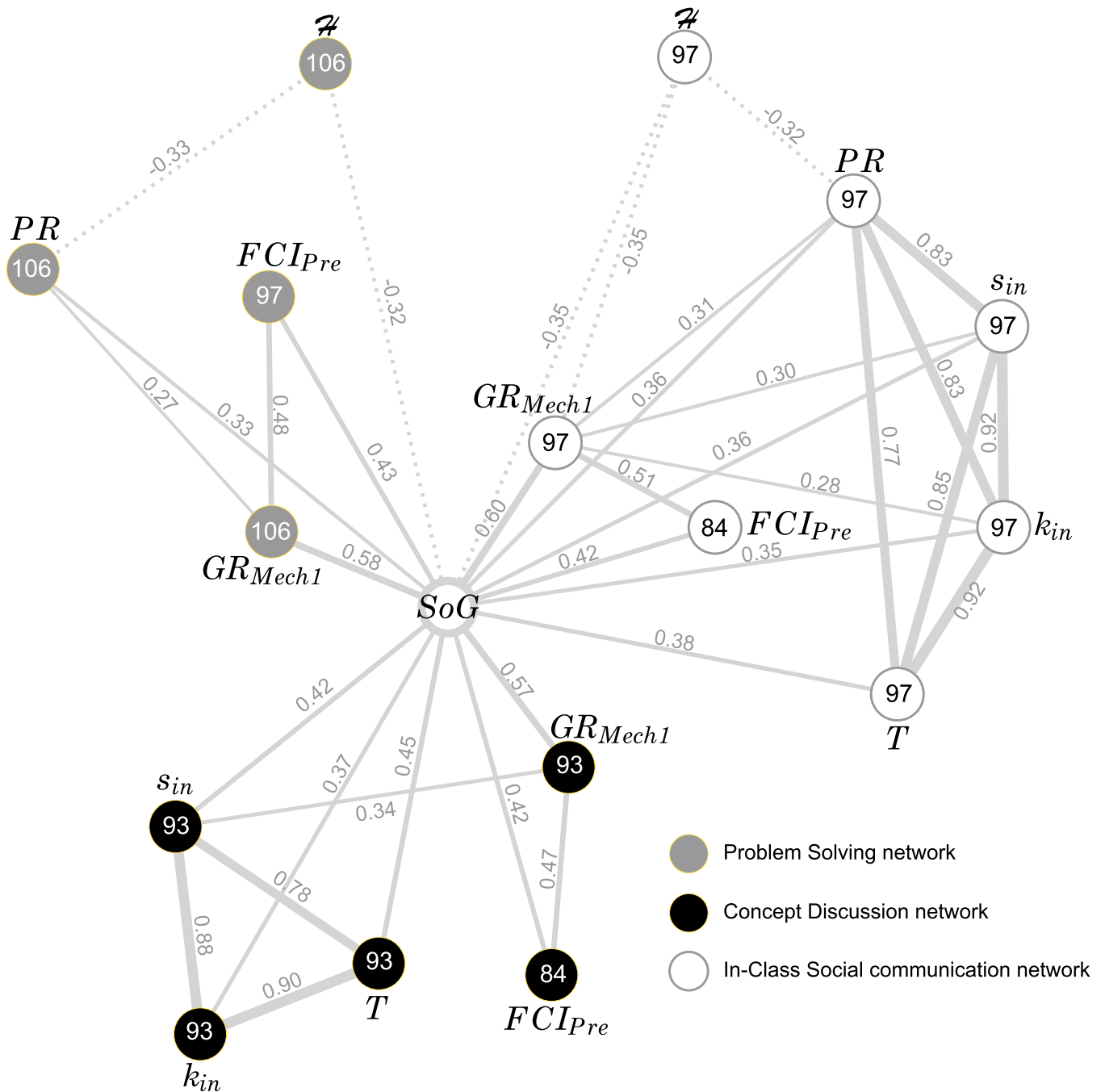


FIG. 3 (color online). Correlation network of measures that correlate significantly with SOG ($p < 0.001$). Correlation coefficients are included on the links, the number of degrees of freedom are shown in the nodes, and the layout is determined by the strength of correlation with SOG.

2. Concept discussion network correlations

In the CD network, three network measures, indegree, instrength, and target Entropy, correlate with SOG. One feature we find interesting is that in the concept discussion network two centrality measures, instrength and indegree, are both about how others view the student. This indicates that a student who is memorable to other students tends to have a higher grade in subsequent courses. Finally, we

note that the target entropy has a slightly higher correlation with the SOG ($r = 0.45$) than does FCI_{pre} ($r = 0.42$).

3. In-class social network correlations

In the ICS network, PageRank, instrength, indegree, and target entropy correlate with the SOG with relatively similar strength ($r = 0.35-0.38$). The ICS network has centrality features similar to those of the other networks: \mathcal{H} has a

negative correlation with SOG, and again the instrength and indegree are positively correlated with SOG. One proposed interpretation of this is that students who are identified by other students as participating in social interactions in the classroom setting tend to earn higher grades and students who are not easily found engaging in social interactions within the in-class social network tend to earn lower grades.

B. Hierarchical multiple regression models

In each of the three networks, the GR_{mech1} score shows the highest correlation with SOG. Thus, we created models using GR_{mech1} as a benchmark, and we searched for models using networked variables with equivalent predictive power. Since most of the measures correlate significantly with the GR_{mech1} attribute, we did not expect that adding centrality measures or FCI pretest scores to a GR_{mech1} model would increase its predictive power significantly. Instead, we used hierarchical multiple regression [3]. The results, including R^2 , F-tests, and Δ values, are listed in Tables III, IV, and V.

1. Problem solving network modeling

The attribute FCI_{pre} correlated most highly with the SOG in the problem solving network, so FCI_{pre} was used to create the base model. As seen in Table III, this base model was significantly better than a constant only model. When PageRank was added and a new model was created, an F-test indicated that this model was significantly different from the FCI_{pre} only model. However, the Δ values in Table III indicate that the model with FCI_{pre} and PageRank

only improves prediction of SOG by 1.52 standard deviations. Thus, we consider the model with FCI_{pre} only as the best predictive model. When compared with a model that includes the grades in the first block of the course, the model with FCI_{pre} is 1.95 standard deviations worse than the grade-based model. This suggests that a model with FCI_{pre} only is close to being significantly worse than a model with the physics grade in the first block. Finally, adding all three attributes that significantly correlated with SOG, FCI_{pre} , PageRank, and \mathcal{H} , did not further improve the prediction of SOG.

2. Concept discussion network modeling

In the concept discussion network, the target entropy (T) was the attribute most highly correlated with the SOG. Thus, T was used to create the base model, as seen in Table IV. This base model improved the prediction of SOG in the second block. In this network, FCI_{pre} was added and a second model was created. An F-test indicated that this model was not significantly different from the model with T only. The difference in prediction of SOG indicates that the model with T and FCI_{pre} improves prediction of SOG by 1.49 standard deviations. Thus, in this case, the best predictive model is the model with T only. When compared with a model that includes the grades in the first block of the course, the model with T only is 1.33 standard deviations worse than the grade-based model, which suggests that a model with T only is not significantly different from a model with the physics grade in the first block.

TABLE III. Comparison of linear models in the problem solving network. In all tables, the model in bold is the simplest model that cannot be improved significantly by adding further variables.

Model	Variables	R^2 [95% C.L.]	F-tests	Δ
PS_F	FCI_{pre}	0.18 [0.04–0.31]	$F(1, 97) = 21.6^b$	
PS _{FP}	FCI _{pre} , PR	0.29 [0.12–0.43]	$F(2, 96) = 18.4^b$	1.52
PS _{FPH}	FCI _{pre} , PR, \mathcal{H}	0.28 [0.12–0.41]	$F_{F/FP}(132, 132) = 0.66^a$ $F(3, 95) = 12, 3^b$	1.59
PS _G	GR _{mech1}	0.34 [0.18–0.48]	$F_{FG/FPH}(132, 132) = 0.91$ $F(1, 106) = 53.7^b$ $F_{F/G}(132, 151) = 0.35^b$	–1.95

^a $p < 0.05$.

^b $p < 0.001$.

TABLE IV. Comparison of linear models in the concept discussion network.

Model	Variables	R^2 [95% C.L.]	F-tests	Δ
CD_T	T	0.20 [0.07–0.33]	$F(1, 93) = 23.7^a$	
CD _{TF}	T, FCI _{pre}	0.30 [0.13–0.44]	$F(2, 83) = 17.7^a$	1.49
CD _{TFP}	T, FCI _{pre} , PR	0.30 [0.12–0.43]	$F_{T/TF}(133, 114) = 0.75$ $F(3, 82) = 11.7^a$	1.49
CD _G	GR _{mech1}	0.32 [0.16–0.47]	$F_{T/TFP}(133, 114) = 0.76$ $F(1, 93) = 43.8^a$ $F_{T/G}(133, 132) = 0.52^a$	–1.33

^a $p < 0.001$.

TABLE V. Comparison of linear models in the in-class social communication network.

Model	Variables	R^2 [95% C.L.]	F-tests	Δ
ICS _F	FCI _{pre}	0.18 [0.02–0.31]	$F(1, 89) = 19.1^b$	
ICS _{FT}	FCI _{pre} , T	0.29 [0.13–0.42]	$F(2, 88) = 17.9^b$	1.69
ICS _{FTH}	FCI _{pre} , T , \mathcal{H}	0.35 [0.18–0.49]	$F_{F/FT}(117, 117) = 0.54^b$ $F(3, 87) = 15.3^b$ $F_{F/FTH}(117, 117) = 0.39^b$	2.33^a
ICS _G	GR _{mech1}	0.35 [0.19–0.51]	$F(1, 97) = 53.2^b$ $F_{FTH/G}(117, 134) = 0.79$	0.10

^a $p < 0.05$.^b $p < 0.001$.

3. In-class social network modeling

In the in-class Social network, FCI_{pre} was the attribute most highly correlated with the SOG, so it was used to create the base model, as seen in Table V. This base model improved the prediction of SOG in the second block. In this network, T was added and a second model was created. An F-test indicated that this model was significantly different from the model with FCI_{pre} only. The difference in prediction of SOG indicates that the model with T and FCI_{pre} improves prediction of SOG by 1.69 standard deviations. However, when both T and \mathcal{H} are added, the third model improves prediction of SOG by 2.33 standard deviations over the base model. In this network, the best model includes not only FCI_{pre}, but also the networked attributes T and \mathcal{H} . The results of the linear modeling in this network are interesting, in that when compared with a model that includes the grades in the first block of the course, the best model with ICS_{FTH} is 0.10 standard deviations better than the grade-based model. This suggests that a model with FCI_{pre}, T , and \mathcal{H} predicts the SOG as well as a model with the physics grade in the first block.

VI. DISCUSSION: UNDERLYING STRUCTURES

A. Understanding relations among networks through correlations

We argue that the three networks are distinct: the problem solving and concept discussion networks measure different aspects of student interactions, while the correlation patterns between attributes and centrality measures in the ICS network suggest that this network measures a combination of PS and CD. The structure of the correlation network in Fig. 3 indicates that in the PS network, PR, and hide (\mathcal{H}) correlate with equal strength with the SOG and significantly with each other. However, in the CD network, instrength (s_{in}), indegree (k_{in}), and target entropy (T) correlate with roughly equal strength with SOG, and to a high degree with each other. In the ICS network, all measures are present and show the same structure as in each of the other two networks: PR and \mathcal{H} are related and s_{in} , k_{in} , and T are related. Furthermore, PR is related to the s_{in} , k_{in} , T variables. Thus, the three networks give information about different aspects of student interaction.

1. Problem solving network correlations

Being easy to locate (low \mathcal{H}) in the PS network tends to be associated with academic success. We posit that students who are easy to find are the same students who are proficient with physics and share their knowledge with others in a memorable way. High \mathcal{H} students may not be proficient with physics or their contributions to problem solving interactions are not memorable to other students. The way this study is constructed, a student has to be recognized by another student as being memorable to get a link. We take this to reflect what Wenger terms as mutual engagement [2]. The named student engaged with the other students in problem solving practices, which, to some extent, were part of a process of meaning making [36]. The student might have been mentioned for several reasons, although if he or she never added to the meaning-making process of others, it is difficult to understand why other students would continue naming him or her.

Adding to this view, we can also imagine high \mathcal{H} students are the ones who do not engage in the problem solving practices of their local physics student community. In that case, we can imagine that they would not have access to other students' problem solving strategies or to worked out solutions to problems. On the other hand, the positive correlation between PR and SOG indicates to us that if others, in general, recognize collaborators (i.e., someone who participates in the practices of the class), then one is more likely to gain experience with others' problem solving strategies. Using the idea of a community of practice [2], students might develop a shared repertoire, to which the students who occupy peripheral positions in the network have limited access, because they do not engage with others. Note that there is no correspondence between being peripheral to the network (however we measure this) and the notion of legitimate peripheral participation of Lave and Wenger [25]. If physics is a trade, then the students in this study are all at the periphery of that trade.

2. Concept discussion network correlations

It is not surprising that T and k_{in} are highly correlated since T relies on the number and distribution of shortest paths to a node. A high k_{in} will in itself contribute to a

large T . The fact that T correlates better with SOG than k_{in} shows that taking the whole network into account does seem to make a difference. Being in a part of the network where ideas and concepts are discussed (and where these discussions are remembered) by many students has a positive impact on later grades.

While strong connections, as represented by the instrength, also seem to have an impact, using the whole network perspective with T has slightly more predictive power when it comes to sum of grades.⁵ In our interpretation, we can associate being a part of the network where ideas are frequently discussed by many students with academic achievement. In that part of the network students have the opportunity to put forth ideas, use physics vocabulary actively, and critically examine others' arguments. These are qualities one could, and one should, associate with good grades. Again, we see the links as the context needed for students engaging mutually with each other in meaning-making processes.

3. In-class social network correlations

The patterns of correlation, as shown in Fig. 3, between centrality measures from the three networks and SOG tell us that engaging in noncontent related social interactions at a lecture, a problem solving session, or laboratory exercise is connected to students' physics understanding or ability to solve problems. It is not clear to us why a student who is socially memorable (central) in class also achieves good grades. This is seen in the in-class social network, which we suggest is a critical component of successful physics learning. Examining the results from the in-class social network from a participationist perspective, we are able to conclude that the practice of physics is not only about discussing concepts and solving problems, but it is also about engaging with others in social realms. It thus seems reasonable to propose that "off-topic" social interactions during class time are an important component in overall engagement in physics because they are a component of building a community around the practice of physics. These results suggest two important follow-up questions:

- (1) What are the mechanisms by which social engagement relates with achievement and learning?
- (2) How can social engagement lead to on-topic engagement or, conversely, how can on-topic engagement lead to social engagement?

B. Linear modeling

It is interesting that the network with the most predictive power is the in-class social network. The linear model that combines FCI_{pre} with T and \mathcal{H} predicts SOG significantly better than FCI_{pre} or any of the proposed network measures alone. This was not the case with the PS and CD networks,

where the best models were the models with one measure, FCI_{pre} and T , respectively. In the ICS network, one interpretation of the linear model is that FCI_{pre} functions as a measure of the individuals' understanding of physics, while T and \mathcal{H} seem to measure different aspects of student participation in learning activities. This model seems like a hybrid of the models that best predicted SOG in the PS and CD networks. This hybrid interpretation is similarly reflected in the correlation network, shown in Fig. 3. Seen in this way, the ICS network could serve as a proxy for the PS and the CD networks, which preserves some of the features of both problem solving and concept discussion.

The linear model for the ICS network allows us to consider that students engage in social interactions, both when solving problems and when discussing concepts (something which may or may not happen at the same time). Thus, when we ask students to name with whom they remember communicating socially in the classroom context, it is not surprising that students include the names of people who they engaged with during both PS and CD activities. That is, we suspect that if you remember working together to solve a problem or discussing a physics concept, you also remember communicating socially. This may indicate that engaging in social interactions and in disciplinary interactions are not easily seen as separate. They may interact or share a common, more general, underlying predictive variable. This line of thinking supports current research that combines classical cognitive thinking with sociocultural theory.

Linear models involving network measures can account for some of the variance in our data. However, the rationale for using linear models for these kinds of problems derives from a view of learning that is linear. Linear modeling allows researchers to see trends in a limited regime of the variables included in the model, but linear models do not include feedback mechanisms or interactions between variables. In contrast to the linear view, recent theories of learning (e.g., Refs. [2,25,36,51,52]) conceive learning as an activity involving many kinds of processes within and between learners and artifacts using multiple modalities. As a consequence, it seems very difficult to conceive of a linear model of learning that could incorporate the development of a shared repertoire of problem solving practices and how this repertoire manifests itself in student performances. In our view, seeing learning as linearly dependent on any macroscopic variable like test scores, grades, or network centralities severely limits our scientific progress in understanding learning. More sophisticated ways of modeling physics learning environments as more complex systems might arise from recent developments in thinking about educational systems, for example, complexity thinking [52].

C. Network analysis methods in PER

With this study we have undertaken an analysis of networked measures and student grades. We find that

⁵We have not yet developed the target entropy measure to incorporate both shortest paths and instrengths. This may impact the correlation between T and SOG.

correlations ($0.3 < r \leq 0.45$, $p < 0.001$) exist between network centrality measures for students at one point in time and their grades at a later point in time. The measures yielding the highest correlations are all probabilistic measures taking the whole network into account.

The in-class social network correlation network can be seen as a hybrid of the problem solving and concept discussion correlation networks, as it seems that the ICS captures both problem solving and concept discussion interactions in addition to social interactions. This is also reflected in the fact that we can combine hide and target entropy calculated on the ICS network to yield significantly better predictions than with the FCI pretest scores alone. This suggests that future studies can use student recall of in-class social interaction as a proxy for both problem solving and concept discussion interactions. This reduction of the number of categories would ease the cognitive load of the survey and would also make it simpler to use networks in intervention studies or to relate networks of social interaction to cognitive networks describing different aspects of students' conceptualizations of physics.

Insofar as the FCI can tell researchers whether a student has a Newtonian view of the world or not, the FCI_{pre} data can be seen as a crude way of integrating the knowledge state with a social network. However, we see many further possibilities for integrating student knowledge with their network position to give a description of both how students act and what they know. Following the ideas of Ref. [5] or Ref. [4], who explore cognitive networks, we propose to use network analyses of cognitive networks to characterize student knowledge states in addition to the position in social networks explored in this article. Comparing student cognitive networks may be the basis for identifying cognitive connections between them, which can then be compared with the social networks.

With the ICS network data, we can create a linear model with FCI_{pre} , target entropy, and hide as variables from the first block of students' courses to predict the sum of grades of the courses in the next block of classes. This model is significantly better at predicting grades than a model using only FCI_{pre} . The success of this model indicates that network measures can be used to generate predictions. While this may (or may not) seem obvious, as the use of network analysis in physics education research evolves, it is important to identify places where network analysis has been successful and useful. Further, the finding that academic success can be predicted based on centrality within a primarily social network indicates the intrinsically social nature of learning. This has implications for instruction and curriculum design. It provides a mechanism, creation of social networks, that supports our understanding of the roles that interactive engagement methods play in learning.

Our results can be taken to show that social interactions correlate positively with physics learning. The counter-argument, claiming that there are other underlying variables responsible for both network positions and grades, needs to incorporate how these variables affect both social interactions and grades. Future work will investigate mechanisms for becoming central in academic (PS and CD) and social (ICS) networks and to further investigate attributes that are correlated with students' academic success as measured by grades.

ACKNOWLEDGMENTS

We would like to thank Renee Michelle Goertzen for her careful reading of the manuscript. We would like to thank Ian Bearden for his efforts with data collection. This research was supported in part by the Center for High Energy Physics Research and Education Outreach (NSF No. 0802184).

-
- [1] L. Vygotsky, *Mind in Society* (Harvard University Press, Cambridge, MA, 1978).
 - [2] E. Wenger, *Communities of Practice Learning, Meaning, and Identity* (Cambridge University Press, Cambridge, England, 1998), 17th ed.
 - [3] E. Brewé, L. Kramer, and V. Sawtelle, Investigating student communities with network analysis of interactions in a physics learning center, *Phys. Rev. ST Phys. Educ. Res.* **8**, 010101 (2012).
 - [4] M. Bodin, Mapping university students' epistemic framing of computational physics using network analysis, *Phys. Rev. ST Phys. Educ. Res.* **8**, 010115 (2012).
 - [5] I. Koponen and M. Pehkonen, Coherent knowledge structures of physics represented as concept networks in teacher education, *Sci. Educ.* **19**, 259 (2010).
 - [6] J. Forsman, Licentiate thesis, Uppsala University, 2011.
 - [7] J. Forsman, C. Linder, R. Moll, D. Fraser, and S. Andersson, A new approach to modelling student retention through an application of complexity thinking, *Stud. Higher Educ.*, doi:10.1080/03075079.2011.643298 (2012).
 - [8] A. Daly, *Social Network Theory and Educational Change* (Harvard Education Press, Cambridge, MA, 2010).
 - [9] V. Pitts and J. Spillane, Using social network methods to study school leadership, *Int. J. Res. Meth. Educ.* **32**, 185 (2009).
 - [10] J. Pustejovsky and J. Spillane, Question-order effects in social network name generators, *Soc. Netw.* **31**, 221 (2009).
 - [11] J.P. Spillane and C.M. Kim, An exploratory analysis of formal school leaders' positioning in instructional advice and information networks in elementary schools, *Am. J. Educ.* **119**, 73 (2012).

- [12] S. Dawson, A study of the relationship between student social networks and sense of community, *Educ. Technol. Soc.*, **11**, 224 (2008).
- [13] L. Macfadyen and S. Dawson, Mining LMS data to develop an “early warning system” for educators: A proof of concept, *Comput. Educ.* **54**, 588 (2010).
- [14] L. Lockyer, S.P. Dawson, L. Macfadyen, and D. Mazzochi-Jones, Using social network metrics to assess the effectiveness of broad based admission practices, *Australas. J. Educ. Tech.* **27**, 16 (2011).
- [15] J. Bruun, Ph.D. thesis, University of Copenhagen, 2012.
- [16] E. Brewe, L. Kramer, and G. O’Brien, Changing participation through formation of student learning communities, *AIP Conf. Proc.* (**1289**, 85 2010).
- [17] S. Dawson, ‘Seeing’ the learning community: An exploration of the development of a resource for monitoring online student networking, *Br. J. Educ. Technol.* **41**, 736 (2010).
- [18] J. Enriquez, Fluid centrality: A social network analysis of social-technical relations in computermediated communication, *Int. J. Res. Meth. Educ.* **33**, 55 (2010).
- [19] D.W. Shaffer, D. Hatfield, G.N. Svarovsky, P. Nash, A. Nulty, E. Bagley, K. Frank, A. Rupp, and R. Mislevy, Epistemic network analysis: A prototype for 21st-century assessment of learning, *Int. J. Learn. Media* **1**, 33 (2009).
- [20] P. Sadler and R. Tai, Success in introductory college physics: The role of high school preparation, *Sci. Educ.* **85**, 111 (2001).
- [21] C. Henderson, Common concerns about the Force Concept Inventory, *Phys. Teach.* **40**, 542 (2002).
- [22] S. Borgatti and V. Lopez-Kidwell, in *The SAGE Handbook of Social Network Analysis*, edited by J. Scott and P. Carrington (SAGE, London, 2011), Chap. 4, pp. 40–55.
- [23] M. Packer and J. Goicoechea, Sociocultural and constructivist theories of learning: Ontology, not just epistemology, *Educ. Psychol.* **35**, 227 (2000).
- [24] A. Sfard, On two metaphors for learning and the dangers of choosing just one, *Educ. Res.* **27**, 4 (1998).
- [25] J. Lave and E. Wenger, *Situated Learning: Legitimate Peripheral Participation* (Cambridge University Press, Cambridge, England, 1991).
- [26] B. Rogoff, *Apprenticeship in Thinking: Cognitive Development in Social Context* (Oxford University Press, New York, 1990).
- [27] B. Rogoff, E. Matusov, and C. White, in *Handbook of Education and Human Development: New Models of Learning, Teaching and Schooling*(Blackwell, Oxford, 1996), pp. 388–414.
- [28] M. Cole, *Cultural Psychology: A Once and Future Discipline* (Belknap Press, Cambridge, MA, 1998).
- [29] S. Scribner, *Everyday Cognition: Its Development in Social Context* (Harvard University Press, Cambridge, MA, 1984), Chap. 1, pp. 9–40.
- [30] P.J. Carrington and J. Scott, in *The SAGE Handbook of Social Network Analysis* (Ref. [22]), Chap. 1, pp. 1–8.
- [31] D. Watts, *Six Degrees: The Science of a Connected Age* (WW Norton & Company, New York, 2003).
- [32] M. Rosvall and C. Bergstrom, Maps of random walks on complex networks reveal community structure, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1118 (2008).
- [33] M. Newman, A. Barabasi, and D. Watts, *The structure and Dynamics of Networks* (Princeton University Press, Princeton, NJ, 2001).
- [34] A. Marin and B. Wellman, in *The SAGE Handbook of Social Network Analysis* (Ref. [22]), Chap. 1, pp. 11–25.
- [35] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevSTPER.9.020109> for a detailed description of combinatoric and probabilistic measures of centrality, for sociograms of the problem solving, concept discussion, and in-class social networks, and for the code developed to use R for the calculations made in this article.
- [36] R. McCormick, A. Fox, P. Carmichael, and R. Procter, *Researching and Understanding Educational Networks* (Routledge, New York, 2011).
- [37] N. Eagle, A. Pentland, and D. Lazer, in *Social Computing, Behavioral Modeling, and Prediction*, edited by H. Liu, J. Salerno, and M. Young (Springer, New York, 2008), pp. 79–88.
- [38] N. Eagle, A. Pentland, and D. Lazer, Inferring friendship network structure by using mobile phone data, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15 274 (2009).
- [39] S. Brin and L. Page, The anatomy of a large-scale hypertextual Web search engine, *Comput. Netw. ISDN Syst.* **30**, 107 (1998).
- [40] K. Sneppen, A. Trusina, and M. Rosvall, Hide-and-seek on complex networks, *Europhys. Lett.* **69**, 853 (2005).
- [41] D. Liben-Nowell, Ph.D. thesis, Massachusetts Institute of Technology, 2005.
- [42] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley, and Y. Åberg, The web of human sexual contacts, *Nature (London)* **411**, 907 (2001).
- [43] P.V. Marsden, in *The SAGE handbook of Social Network Analysis* (Ref. [22]), Chap. 25, pp. 370–388.
- [44] D. Hestenes, M. Wells, and G. Swackhamer, Force Concept Inventory, *Phys. Teach.* **30**, 141 (1992).
- [45] R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria 2012), ISBN 3-900051-07-0.
- [46] G. Csardi and T. Nepusz, The igraph software package for complex network research, *InterJournal* **18**, 1695 (2006).
- [47] S. Wasserman and K. Faust, *Social Network Analysis, Methods and Applications* (Cambridge University Press, Cambridge, England, 1994).
- [48] B.G. Tabachnick and L.S. Fidell, *Using Multivariate Statistics* (Allyn and Bacon, Boston, 2007).
- [49] A. Canty, Sampling methods in R: The boot package *R News* **2**, 2 (2002).
- [50] W. De Nooy, A. Mrvar, and V. Batagelj, *Exploratory Social Network Analysis with Pajek* (Cambridge University Press, Cambridge, England, 2011).
- [51] A. Gupta, D. Hammer, and E.F. Redish, The case for dynamic models of learners’ ontologies in physics, *J. Learn. Sci.* **19**, 285 (2010).
- [52] B. Davis and D.J. Sumara, *Complexity and Education: Inquiries into Learning, Teaching, and Research* (Lawrence Erlbaum, Mahwah, NJ, 2006).