# "I remember how to do it": exploring upper elementary students' collaborative regulation while pair programming using epistemic network analysis

**Article** *in* Computer Science Education · March 2022

DOI: 10.1080/08993408.2022.2044672



**4 authors**, including:

Jessica Vandenberg
North Carolina State University


Eric Wiebe
North Carolina State University

Computer Science Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ncse20

# “I remember how to do it”: exploring upper elementary students’ collaborative regulation while pair programming using epistemic network analysis

Jessica Vandenberg, Collin Lynch, Kristy Elizabeth Boyer & Eric Wiebe

**To cite this article:** Jessica Vandenberg, Collin Lynch, Kristy Elizabeth Boyer & Eric Wiebe (2022): “I remember how to do it”: exploring upper elementary students’ collaborative regulation while pair programming using epistemic network analysis, Computer Science Education, DOI: 10.1080/08993408.2022.2044672

**To link to this article:** https://doi.org/10.1080/08993408.2022.2044672

Published online: 09 Mar 2022.

COMPUTER SCIENCE EDUCATION
https://doi.org/10.1080/08993408.2022.2044672

# "I remember how to do it": exploring upper elementary students' collaborative regulation while pair programming using epistemic network analysis

Jessica Vandenberg[a], Collin Lynch[a], Kristy Elizabeth Boyer[b] and Eric Wiebe[c]

[a]Department of Computer Science, North Carolina State University, Raleigh, NC, USA; [b]Department of Computer & Information Science & Engineering, University of Florida, Gainesville, FL, USA; [c]Department of Stem Education, North Carolina State University, Raleigh, NC, USA

**ABSTRACT**

**Background and Context:** Students' self-efficacy toward computing affect their participation in related tasks and courses. Self-efficacy is likely influenced by students' initial experiences and exposure to computer science (CS) activities. Moreover, student interest in a subject likely informs their ability to effectively regulate their learning in that domain. One way to enhance interest in CS is through using collaborative pair programming.
**Objective:** We wanted to explore upper elementary students' self-efficacy for and conceptual understanding of CS as manifest in collaborative and regulated discourse during pair programming.
**Method:** We implemented a five-week CS intervention with 4th and 5th grade students and collected self-report data on students' CS attitudes and conceptual understanding, as well as transcripts of dyads talking while problem solving on a pair programming task.
**Findings:** The students' self-report data, organized by dyad, fell into three categories based on the dyad's CS self-efficacy and conceptual understanding scores. Findings from within- and cross-case analyses revealed a range of ways the dyads' self-efficacy and CS conceptual understanding affected their collaborative and regulated discourse.
**Implications:** Recommendations for practitioners and researchers are provided. We suggest that upper elementary students learn about productive disagreement and how to peer model. Additionally, our findings may help practitioners with varied ways to group their students.

**ARTICLE HISTORY**
Compiled 23 February 2022

**KEYWORDS**
Elementary school; pair programming; attitudes; self-efficacy; academic regulation

## 1. Introduction

Students' initial experiences and exposure to computing as well as socio-cultural perceptions around gender and computing influence their attitudes toward computer science (Jepson & Perl, 2002; Mejias et al., 2019). Research suggests that consistent and high-quality experiences with computer use both in- and out-of-school likely inform students' interest in and self-efficacy for computing. Prior studies have shown that initial, low quality experiences may prove boring (see, Goode, 2010; Moyer et al., 2018) or deleterious

**CONTACT** Jessica Vandenberg jvanden2@ncsu.edu Department of Computer Science, North Carolina State University, Raleigh, NC, USA

to one's self-efficacy (Lishinski et al., 2016) and that these experiences affect students' participation in related tasks and courses, and the likelihood that they will pursue a degree program in CS or a career related to computing (e.g. Cassel et al., 2007; Mitchell et al., 2009; Yardi & Bruckman, 2007). Moreover, a lack of access to early and frequent computing activities may influence students away from considering studying computing in the future (see, Tran, 2018).

One strategy employed to provide high-quality, engaging programming experiences, especially among girls and historically underrepresented minorities, is through the use of pair programming (McDowell et al., 2006; Porter et al., 2013). This collaborative approach to learning computer programming has been shown to result in higher self-efficacy, in part, because the stereotype of solitary programmers is attenuated and the two programmers engage in supportive discussion of their thinking (L. L. L. L. Werner et al., 2004). In fact, Zhong et al. (Zhong et al., 2016) found that pair programming improved young students' sense of partnership and collaboration. Pair programming has primarily been studied with high school and university-aged students (e.g. Missiroli et al., 2016; Williams et al., 2000), but there is a growing interest in applying it with younger students since collaborative learning strategies are regularly used in elementary classrooms (e.g. Gillies & Boyle, 2010).

The verbal interchange between students is both seen as a central facet of efficacy of this strategy and a means of eliciting data for a better understand of how and why it may be effective. Pair programmers' collaborative discussions are a focal point here as we are particularly interested in how students' discourse over the course of a programming activity may illuminate individual and pair learning strategies. These strategies are driven, in part, by cognitive regulation on the part of the programming pair. An individual's ability to regulate their learning significantly predicts their group's regulation (Panadero et al., 2015). Moreover, when groups utilize more regulatory strategies, they often perform better on their academic tasks (Janssen et al., 2012).

There is longstanding interest in the relationship of academic self-efficacy, prior knowledge, and learning outcomes (Hinckle et al., 2020). Far less work has been done in this area with younger students in computer science or how these affective and cognitive factors may influence the collaborative regulatory dynamics inherent in pair programming. Although studies with older students indicates that individual regulatory ability predicts group regulatory ability (Panadero et al., 2015), less is known about whether this holds for younger students or the impact of similar or differing levels of prior knowledge and self-efficacy within the dyad.

## 2. Literature review

### *2.1. Theoretical framing: collaborative regulation of learning*

Learning is often a social activity, and through talking with others, students' individual cognitive capacities are recruited, constructed and refined (Mercer et al., 1999). As such, it is important to consider how students regulate themselves and others in group learning tasks, to what extent individual regulation influences group regulation (Panadero et al., 2015), and how group processes influence an individual's acquisition of self-regulated learning skills (Hadwin & Oshige, 2011).

Research into how students collectively engage in academic regulation emphasizes student interactions and fluctuations in group member influence, in particular in task performance and their related social processes. For example, dyadic social processes were monitored by group members more so than task processes on an inquiry-based computer-supported collaborative learning (CSCL) activity (Saab et al., 2012). This finding implies that regulating the (social) collaborative process is paramount to, and perhaps facilitates, effective task performance. In partial contrast, Janssen et al. (2012) found that groups devoted approximately 35% of their time to planning and monitoring their task and 30% of their time to social processes – both creating a shared understanding and social support – however, only the regulation of social processes contributed positively to group performance.

Students' discourse gives insight into what they think, what they want, and how they make sense of tasks (Johnstone, 2017; Potter, 1998). Students' own words while engaged in a programming activity can be especially informative when we also consider students' self-reported self-efficacy toward CS and their performance on a CS assessment. This follows because of findings around prior experience likely fosters both conceptual understanding and self-efficacy (Schunk, 1995), which can lead to better academic regulation (Bradley et al., 2017; Pajares, 2002). This better academic regulation tends to promote better learning outcomes (Schunk & Zimmerman, 2012), which then leads to higher conceptual understanding and self-efficacy, thus completing the virtuous cycle.

To examine students' discursive collaborative regulation, we previously designed, piloted, and refined a framework (Vandenberg et al., 2021; Vandenberg, Tsan, Zakaria et al., 2020); (see, Table 1) that included components of regulation frameworks by Hadwin et al. (2005) and Janssen et al. (2012), and a social talk framework by Kumpulainen and Mutanen (1999). The design and refinement process occurred largely because the existing frameworks included elements not relevant to the population or contexts under study.

**Table 1.** Collaborative regulation of learning framework.

| Dimension | Code (*k*) | Definition | Examples |
|---|---|---|---|
| Task | Planning the task (.84) | Discussion of the task, how to complete it, deciding which strategies to employ, responsibilities students will take on | Let's start by picking a background. |
| | Monitoring task progress (.90) | Discussion of performance and progress, specific mention of strategies being used to approach the task, mentions of time | The glide block worked better last time, so let's try that. |
| | Evaluations of task progress (.78) | Review of performance and progress, includes appraisals of task difficulty | That was harder than I thought it would be. |
| Social | Collaborative (.82) | Actively engaging with partner, attempts to maintain symmetrical contributions | Let's change it so she says "hello" for longer, don't you think? |
| | Agreement (.78) | Acknowledgements and affirmations, most often in response to a partner's contribution | Oh yeah! |
| | Tutoring (.82) | Asking for or offering help/assistance | Hey, how do I add another sprite? |
| | Disagreement (.84) | Social or academic conflict | I'll delete it if you write that in there. |
| | Confusion (.73) | Failure to understand the partner or the task, often accompanied by a question | What are you talking about? |
| | Individualistic (.90) | Working independently with no clear attempt to involve the partner | [these examples often looked like self-talk in proximity to another] |

### 2.2. Assessing young students' CS self-efficacy and conceptual understanding

There is an established link between students' self-efficacy for a content area and their performance in that area (Brosnan, 1998; Lishinski et al., 2016). In a study with six year old students, even brief experiences with programming have the potential to enhance not only their individual technology motivation and interest, but such experiences also reduce students' perception that programming is only for boys (Master et al., 2017). These findings are important as they support the links between interest, self-efficacy, and intentionally designed experiences, all of which, Bandura et al. (1999), increases students' effort and contributes to higher achievement. Less is known about how a student's outcome expectancy and self-efficacy interact in a collaborative context and how these may contribute to the student's individual and group regulatory performance.

There is limited work on assessing upper elementary students' self-efficacy towards CS, coding or programming in particular. In fact, there are, to date, only three published, validated instruments for 4th and 5th grade students whose underlying constructs are interests, attitudes, or self-efficacy. Kong et al. (2018) developed a 15-item 5-point Likert-scale survey that queries students on programming meaningfulness and impact, and creative and programming self-efficacy. They also assessed if programming interests were related to these four factors; they found that students with more interest in programming found it to be more meaningful and impactful and had higher self-efficacy. Moreover, boys had higher interest than girls. Mason and Rich (2020) developed a 24-item 6-point Likert-scale survey that queries students on what they call *attitudes toward coding* that comprises students' coding confidence and interest, the perceived utility of coding, and students' perceptions of coders and social influence. They found that when students self-report high interest in coding, they have greater coding self-efficacy, supporting research in other subject areas (e.g. Grigg et al., 2018; Sheldrake, 2016). Of note, gender differences in coding attitudes were small, although statistically significant, and the authors suggest that younger students may have less exposure to gender bias in coding. This interest in young students' coding attitudes is also taken up by Vandenberg et al. (2021a). This 11-item 5-point Likert-scale survey queries 4th and 5th grade students on the two psychological factors of self-efficacy and outcome expectancy toward CS, coding in particular. These factors are empirically connected; individuals are motivated by their beliefs in their capabilities to complete a specific task (self-efficacy) and that completing that task will produce a desired outcome (outcome expectancy; Bandura et al., 1999). We found that boys had statistically significantly higher CS self-efficacy and outcome expectancy than girls and that there were no statistically significant differences by race/ethnicity.

As noted earlier, prior experience and its associated knowledge development also influences regulatory processes (Schunk & Zimmerman, 2012). Regarding CS conceptual understanding, as with measuring the affective domain, there has been minimal work done with developing and validating instruments appropriate for upper elementary aged students. Román-González et al. (2017) created a largely block-based assessment for 5th through 10th grade students, which emphasizes computational thinking rather than explicit computer science concepts. Another instrument, developed for 10 to 14 year old students, used block-based coding in Alice as a way to score students' computational thinking in a performance assessment, rather than a more traditional multiple-choice instrument (Werner et al., 2012). Most recently, Vandenberg et al. (2021b) validated a measure appropriate for upper elementary students that utilized mostly block-based programming language and which targets core CS concepts, such as loops, conditionals, variables, and algorithms. It is this instrument that we used in our work.

### 2.3. Pair programming

In CS education, collaborative work often takes the form of pair programming. Traditional pair programming entails two students working on a single computer, each student has a designated role – the *driver* who has control of the input devices and the *navigator* who strategically guides the work (Williams et al., 2000). Both programmers are expected to talk continuously about their work, engage in collaborative problem solving, and to switch roles after a set amount of time or portion of the task has been completed. This pedagogical configuration has been used in industry (Canfora et al., 2007), in undergraduate classes (Williams et al., 2000), and in high school (Missiroli et al., 2016). As interest in CS education has moved to earlier grades, there is a growing interest in using pair programming with younger students (e.g. Denner et al., 2014; Shah et al., 2014; Tsan et al., 2020). Research suggests that the pair programming approach may be particularly helpful for females (Werner et al., 2004) and may increase pair programmers' confidence in, and enjoyment of programming (McDowell et al., 2006).

### 2.4. Epistemic network analysis

Close study of the collaborative work of pair programmers inevitably means structured, systematic analysis of their discourse. Although qualitative methods of coding and analyzing dialogue have long been used in this type of research, typically relied on manual processes of linking individuals' talk temporally and across members of the group. Newer methods have emerged that provide unique and powerful insights into the discursive dynamics of dyads. Epistemic Network Analysis (ENA) is a mixed methods and data visualization technique for modeling and analyzing the temporal connections in qualitatively tagged discursive data within and across individuals. This modeling is based upon quantifying the co-occurrence of codes/tags within a conversation. This in turn generates a weighted network of co-occurrences and related visualizations for each unit of analysis. ENA permits researchers to compare these networks visually, statistically, and qualitatively by analyzing networks concurrently. ENA was originally developed to explore the interdependence of cognition, culture, and discourse (Shaffer et al., 2016); use of the technique assumes that connections in the tagged discourse – spoken or typed – are meaningful. That is, temporally adjacent statements are likely cognitively linked. Others have used ENA to explore university students' design thinking (Arastoopour et al., 2016), surgical residents' speech and inclusion of error checklists (Ruis et al., 2018), and socioemotional group interactions in an online STEM education course (Wang et al., 2020).

ENA, coupled with self-efficacy and prior knowledge data, provides us a particularly powerful approach to look at the role of these affective and cognitive dimensions in regulatory processes of collaborative programming. The self-efficacy and prior conceptual understanding data on individuals allows us to set up groupings of homo- and heterogenous dyads to be studied in depth using ENA. As noted, ENA provides us tools to systematically compare and contrast these different dyadic groupings. This work will be guided by the following research questions.

## 3. Research questions

Guided by the literature above, we set out to answer the following questions:

**1**. How do dyads' individual self-efficacy and conceptual understanding scores influence collaborative regulated discourse?

1*a*) How do individuals within dyads with similarly high self-efficacy and conceptual understanding scores collaboratively regulate their learning?

1*b*) How do individuals within dyads with similarly low self-efficacy and conceptual understanding scores collaboratively regulate their learning?

1*c*) How do individuals within dyads with mixed self-efficacy and conceptual understanding scores collaboratively regulate their learning?

## 4. Method

The purpose of this study is to explore how students' collaborative regulatory discourse differed based on CS self-efficacy and CS conceptual understanding within dyads of pair programmers. Students' scores on two self-report measures, taken at the outset of the study, sorted them into High, Low, and Mixed dyadic categories. We then tagged students' collaborative discourse using an academic and social regulation framework to discern how these different categories manifest in students' collaborative talk. Using Epistemic Network Analysis (ENA) to visually and qualitatively model dyads' discourse, we study whether High and Low dyads spoke and academically regulated in ways anticipated by their individual psychological dispositions. Prior literature provides little guidance as to how we would expect Mixed dyads to regulate. Thus a more exploratory approach will be used to provide additional insight into how young students collaborate in such pairings.

### *4.1. Participants and context*

Consented participants included 60 4th and 5th grade students (out of a total of 76 students in both grades) from a school in the southeastern United States. The school's socio-demographic data included 75% White/Caucasian, 10% Black/African American, 5.5% Multiracial, 6.5% Latinx/Hispanic, 3% Asian, with 4% of the student population deemed low-income and 51% of the students identified as female. The educational theme of the school centered around global education and awareness in addition to project-based learning. In alignment with others' ENA work (Bressler et al., 2019) with similarly aged student and in acknowledgement of the qualitatively-intensive analysis, we analyzed 24 students, organized into 12 dyads; these dyads were selected based on recorded audio quality and completion of all of the instruments.

The students attended weekly technology classes and learned a series of block-based coding lessons, as guided by their technology teacher. The intervention was designed to incorporate five total lessons for each group of students. The lessons instructed students on foundational CS concepts such as how to use conditionals, loops, and variables. The

**Table 2.** Participant pseudonyms and status indicators.

| Dyad Number | Pseudonym | E-CSA Self-Efficacy Status | E-CSCA Status | Dyad Status |
|---|---|---|---|---|
| 1 | Melanie | high | high | High |
| 1 | Poppy | high | high | High |
| 2 | Mila | high | low | Mixed |
| 2 | Nathan | low | low | Mixed |
| 3 | Max | low | low | Low |
| 3 | Joshua | low | low | Low |
| 4 | Samantha | low | high | Mixed |
| 4 | Andi | high | low | Mixed |
| 5 | Rylee | low | high | Mixed |
| 5 | Amber | high | low | Mixed |
| 6 | Phoebe | low | low | Mixed |
| 6 | Kylie | high | high | Mixed |
| 7 | Emma | high | low | Mixed |
| 7 | Malachi | high | low | Mixed |
| 8 | Louis | low | high | Mixed |
| 8 | Ashley | low | high | Mixed |
| 9 | Sahil | low | high | Mixed |
| 9 | Ezra | high | low | Mixed |
| 10 | David | high | low | Mixed |
| 10 | Leo | low | low | Mixed |
| 11 | Allegra | high | low | Mixed |
| 11 | Chloe | high | low | Mixed |
| 12 | Alaina | low | low | Mixed |
| 12 | Arden | low | high | Mixed |

teacher paired the students based on her assessment of who might work well together, and the students retained these partnerships over the duration of the study. As part of their participation in the study, all students completed several self-report surveys. The E-CSA (Vandenberg et al., 2021a) and ECSCA (Vandenberg et al., 2021b) were administered pre-intervention to assess their knowledge, self-efficacy, and outcome expectancy coming into the study. Table 2 presents the students' pseudonyms and relevant data for the following analysis.

## 4.2. Instruments

### 4.2.1. Elementary-Computer Science Attitudes (E-CSA) (self-efficacy items)

This 11-item 5-point Likert scale survey intended for upper elementary use queries students on their self-efficacy and their outcome expectations for learning CS, coding more specifically, and was adopted from the middle grades version developed and validated by Rachmatullah, Wiebe, et al. (2020). Based on a previously validated questionnaire, this version underwent revision and validation through cognitive interviewing (Vandenberg, Tsan, Boulden et al., 2020) to ensure appropriate wording for young students and has undergone confirmatory factor analysis and item response theory-Rasch analysis for establishing validity and reliability (Vandenberg et al., 2021a). The CS self-efficacy subscale consisted of 4 items ($\alpha$ = .812) and the CS outcome expectancy subscale consisted of 7 items ($\alpha$ = .838). For our analysis, we only used the students' answers to the subscale containing the 4 self-efficacy items. These include statements such as, "I am good at fixing code". Students were then organized into high-low categories via a median split.

#### 4.2.2. Elementary-Computer Science Concepts Assessment (E-CSCA)

This 18-item multiple choice assessment intended for upper elementary use queries students on their knowledge of foundational CS concepts such as loops, conditionals, and variables and does so using mostly block-based language. Based on a validated middle grades version that assessed the same concepts (Rachmatullah, Akram et al., 2020), the elementary version's results (Vandenberg et al., 2021b) indicate psychometrically sound items with no statistically significant item bias by gender or grade. Item response theory (IRT) analysis indicated an appropriate range of item difficulty for the target ages, thus capable of differentiating between a range of abilities. We used the students' scores on the assessment to organize them into high-low categories via a median split (Table 2).

### 4.3. Procedure and analysis

We used a mixed methods design (Creswell & Clark, 2017) to our work. Our data included quantitative student self-report scores and qualitatively coded student transcripts. Moreover, we used ENA to create visual network models, with both quantitative and qualitative information.

Dyads were video and audio recorded each time they collaboratively programmed. We used Open Broadcaster Software (https://obsproject.com/) to align the dyads' webcam video, their audio (gathered through headsets attached to the laptop), and their screen capture. The videos, approximately 40 to 50 minutes in length, were transcribed verbatim and qualitatively tagged using the collaborative regulation of learning framework (see, Table 1). The videos were selected from the second or third day of the intervention; thereby providing the students time to acquaint themselves with their partner and to learn certain CS concepts. The task most of the students analyzed here focused on was user input, or coding a sprite to query the user and then use that information to respond. The number of discursive moves made by the dyads ranged from 431 to 1,020, with an average across the 12 dyads of 712 moves.

Each task-related utterance a student made was tagged with one code from the Task Regulation dimension, as the students demonstrated which phase of the regulation cycle they were in, and at least one code from the Social dimension, as the students used their language to communicate for specific purposes, such as to disagree or express confusion. The first author trained a second researcher on the framework and, after resolving misunderstandings of the codes, they dual coded 25% of the dataset (three transcripts/videos). Kappas were computed on each category (see Code column in Table 1 for individual category kappas), ranging from a low of .73 in the Confusion category to a high of .90 in both the Monitoring and Individualistic categories. An overall kappa of .82 and agreement of .96 was reached (McHugh, 2012). Having established acceptable kappa and in adhering to the literature on dual coding at minimum 20% of the data (Syed & Nelson, 2015), the first author then proceeded to solo code the remainder of the transcripts/videos.

The qualitatively tagged transcripts were then imported into ENA along with necessary metadata, including dyad number, pseudonyms, the individual students' high/low status based on their scores on the E-CSA (self-efficacy) and the E-CSCA, and overall dyad status. The individual student scores were assigned a high or low status designation determined by a median split of all the students' data (see, Table 2). Dyad Status was determined by comparing the individual members of the dyads' CS self-efficacy status (low or high) and conceptual understanding status (low or high). Dyads where both members achieved high on both instruments were deemed High Dyad Status (n =1). Low Dyad Status was ascribed to those dyads where both members achieved low status on both instruments (n =1); Mixed Dyad Status was applied to the remaining dyads whose members had a mix of high and low scores on the two instruments (n =10). We then created a series of ENA models hierarchically organized by the dyads' collective performance on the two instruments.

To answer our research questions, we created our models using the following information. The units of analysis were all lines of data associated with a single value of Dyad Status subsetted by Dyad and by individual Student. By hierarchically organizing our units in this way, we could visually compare the three types of Dyad Status: Low, Mixed, and High. That is, we could analyze how the dyads' regulated discourse differed according to their dyadic self-efficacy and performance status.

Next, we made determinations about how ENA would make connections within the students' dyadic discourse. Our data were segmented into turns of talk, wherein everything a single student said was bracketed on either side by what was said by their partner. These turns of talk could be brief utterances such as "huh?" to longer, multi-sentence statements that included task relevant and irrelevant responses. Temporal proximity in students' discourse likely indicates meaningful connection (Siebert-Evenstone et al., 2017). Connections within the networks are determined by creating a scanning window for the co-occurrence of tags in the current statement and within a set number of previous lines, in our case the window was 8 lines (or 8 turns of talk). We settled on this window length because qualitative analysis of the pairs' discourse for rates of connections indicated they often conversed at length, suggesting a larger window size would capture the richest connections. Our ENA model included the following tagged collaborative regulation categories (from Table 1): Planning, Monitoring, Evaluation, Collaborative, Agreement, Tutoring, Disagreement, Confusion, and Individualistic. In each network model, each node represents a collaborative regulation category. A dimensional reduction via a single value decomposition (SVD) algorithm was used to rotate the model, similar to what occurs in principal components analysis, so that the x-axis explains the greatest variance among the units and the y-axis explains the second greatest variance (Arastoopour et al., 2016). Our model had co-registration correlations of 0.99 (Pearson) and 0.99 (Spearman) for the first dimension and co-registration correlations of 0.96 (Pearson) and 0.96 (Spearman) for the second.

We analyzed the resulting normalized models according to the node size, edge (or line) thickness, and node placement. Node size indicates the frequency of the corresponding collaborative regulation category that occurred relative to all category co-occurrences. Edge thickness, which appears both as width and color intensity, indicates the relative frequency of the connected collaborative regulation categories, or nodes. ENA places nodes using an optimization routine that minimizes the distance between the representative point of the network and the centroid of that network model (Shaffer et al., 2016). Thus, a centroid "summarizes the network as a single point in the projection space that accounts for the structure of connections in the specific arrangement of the network model" (Shaffer et al., 2016, p. 16). In our models, the dyad centroids are represented as unconnected squares, with red denoting High, blue denoting Mixed, and purple denoting Low, and each represent the mean of all dyads within that group.

Because we were interested in looking at the role inter- and intra-dyad differences in self-efficacy and content knowledge might play in regulatory discourse while pair programming, median splits with E-CSA (self-efficacy) and the E-CSCA (content knowledge) scores were used to form analysis groupings. This approach was driven by empirical connections between student self-efficacy and CS knowledge or experience (Beyer & Haller, 2006; He & Freeman, 2010; Hinckle et al., 2020), and these two constructs and regulation (Moos & Azevedo, 2008; Pintrich, 2000; Zimmerman, 2000). Based on the prevailing literature, we anticipated that students would largely score similarly on both measures – high or low on both. However, of the 24 students under analysis, 15 had mixed scores. Therefore, we started by creating the mixed dyad group, comprising 10 dyads with both inter- and intra-differences in self-efficacy and knowledge scores. As the next step, we start our analysis by closely examining the mixed cases to see what patterns and thematic groupings emerge within this grouping. Prior work shows that these mixed dyads may provide more insight into how self-efficacy and prior knowledge affects discursive patterns while collaborating (cf., Denner et al., 2014; Mohamed, 2019).

## 5. Findings

The nodes within ENA permit us to compare several networks and to ascribe meaning to where dyads' centroids reside in the model, relative to the nodes. Figure 1 displays singular, extreme cases of High and Low dyads. That is to say, one dyad has each member score as high on both instruments, noted as the red High square in the figure, and one dyad had each member score as low on both instruments, noted as the purple Low square in the figure. We begin our findings by providing a within-case report for each of these extreme cases. Table 1 provides our codes, definitions, and examples.

### *5.1. Extreme cases: high to low*

#### *5.1.1. High: dyad 1*

Dyad 1 students, Poppy and Melanie, collectively uttered more co-occurring statements tagged as Monitoring and Collaborative, Monitoring and Agreement, Collaborative and Agreement, and Monitoring and Tutoring. Collaborative regulation researchers maintain that students' use of task-regulating behaviors (e.g. monitoring) contributes to effective collaboration (Janssen et al., 2012; De Jong et al., 2005). Excerpts from the students' work demonstrate how the two smoothly move from collaboratively agreeing on their problem-solving approach to questioning the other's strategy to talking through their individual desires to resolving minor disagreements, such as in the excerpt in Table 3.

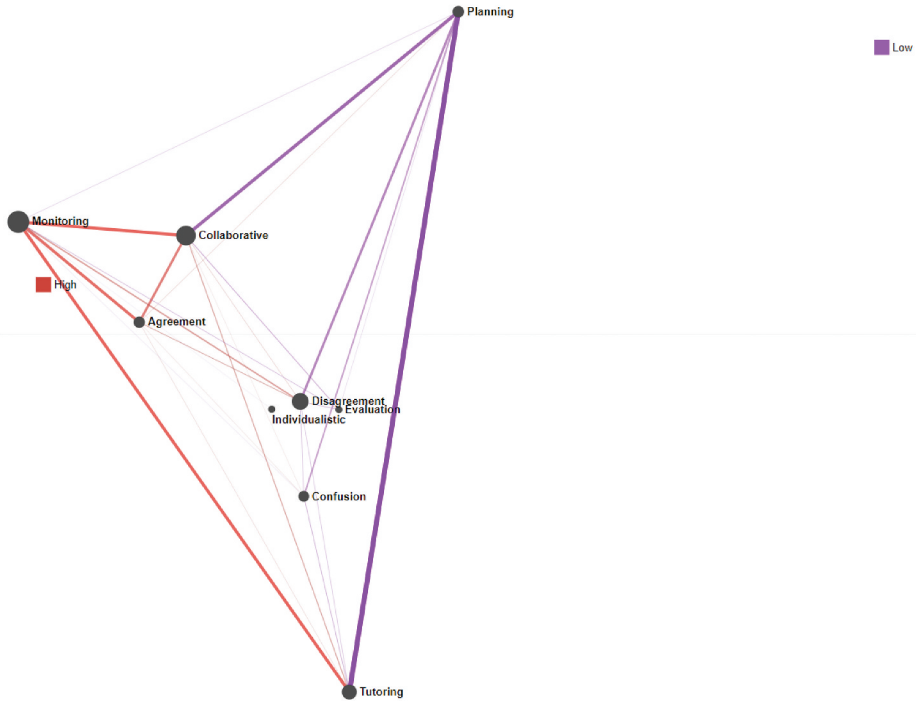


**Figure 1.** ENA network diagram of high and low dyads displaying high dyads in red and low dyads in purple.

**Table 3.** High dyad: excerpt 1.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Melanie (N) | Oh, we need to put a space in, in front of . . . I wonder if we should do a say [block] . . . Wait, why'd you put that? |
| Poppy (D) | Oh. |
| Melanie (N) | Back space and then space. I wonder if we should do that on another line so that, like another say, so that it's not like, so awkward 'cause it's like all in one, you know? |
| Poppy (D) | [reading the code] What's your name? PM, enter. Hello, PM. Nice to meet you. |
| Melanie (N) | I think we should do like- |
| Poppy (D) | No, I think, I think, I think it's good like this 'cause it, 'cause- |
| Melanie (N) | But it's all jumbled, I feel. I feel like it could just be a little neater. |
| Poppy (D) | Hm. |
| Melanie (N) | What do you think? |
| Poppy (D) | Should we do it? I feel like we should- |
| Melanie (N) | Can we just try it and then- |
| Poppy (D) | Hello, PM, then it would go off. Then it would say nice to meet- [runs the code] |
| Melanie (N) | Yeah. |
| Poppy (D) | . . . you. Yeah, I guess that would work. |

In this excerpt, Melanie, who is navigating, encourages Poppy, the driver, to clean up the code so it appears neater. Poppy initially resists, but after Melanie asks to "just try it", Poppy realizes that the "jumbled" code blocks are not simply messy looking but that they are not permitting the program to run how the girls desired.

**Table 4.** High dyad: excerpt 2.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Melanie (N) | So we need to use a loop. |
| Poppy (D) | But which? |
| Melanie (N) | Oh this one- *points at the screen* If space key pressed- |
| Poppy (D) | Okay. |
| Melanie (N) | This is- |
| Poppy (D) | So, if space key pressed, this happens. And then you need to put that in a loop. |
| Melanie (N) | Space key pressed. |
| Poppy (D) | So pick three letters. How about A, M, and T? |
| Melanie (N) | Okay. |
| | . . . . . . *the girls debate which fruits to represent the letters during which time Poppy references class time ending soon* |
| Melanie (N) | But we need to work, but, okay, so we need to put this in a loop but we need to make it so this changes every time. |
| Poppy (D) | Eh, that's hard. |
| Melanie (N) | But for now, let's try it. |
| Poppy (D) | Say "mango" for 2 seconds |
| Melanie (N) | Wait. It needs to like ask [the user] something- |

Grounding, or seeking shared understanding of the task goal or next steps, is essential for collaboration and to prevent unproductive conflict (Erkens et al., 2006; Jeong & Hmelo-Silver, 2016). In Poppy and Melanie's discourse, grounding occurs through their use of Tutoring and Agreement statements, in particular. The Tutoring statements, as noted in Table 4, were often simple requests for clarification and task-related assistance so the two could remain grounded.

In this excerpt, the girls were tasked with building code such that a user would be given options of letters to select and, upon selecting one, a written response would appear in addition to an image, in this case fruit. Poppy is the driver again and receives tutoring assistance from Melanie in the form reminders of needing to use a loop, which one to use, and why that loop block is necessary.

### *5.1.2. Low: dyad 3*

Dyad 3 students, Max and Joshua, offered more co-occurring statements that included Planning. That is, Planning and Collaborative, Planning and Disagreement, Planning and Confusion, and Planning and Tutoring. Collaborative regulation research often reports that students rarely plan (Järvelä & Hadwin, 2013), so these consistent co-occurring statements that include Planning are intriguing. However, upon closer examination of the dyad's transcripts, in conjunction with watching their video, the boys seldom progressed beyond the planning stage of the task as they intermittently sang songs, engaged in off-task conversations about popular YouTube videos, and, after discovering the decibel scale for the research audio collection, made loud noises to get the scale to "go red". They only focused on their coding task when they saw an adult nearby or when their teacher sat and took them through the lesson step-by-step.

The boys' use of Tutoring statements largely revolved around exclamations for help (ie., "Max! Help me!") and questions such as "Max look, I need help, where do we go now?" and "Can you help me?" A brief excerpt that includes Planning, Tutoring, Confusion, and Disagreement statements appears in Table 5.

**Table 5.** Low dyad: excerpt 1.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Joshua (D) | Max! Help me! |
| Max (N) | Dude, you're the driver. |
| Joshua (D) | I don't know how to do this. |
| Max (N) | Wait. What are we supposed to do? |
| Joshua (D) | *Reading the directions on the screen* Create a program that takes- that takes . . . |
| Max (N) | Okay. |
| Joshua (D) | *Reading the directions on the screen* . . . in your user's name and greets them. |
| Max (N) | So the first thing we do- You go to the costumes |
| Joshua (D) | Wait. I don't know. Is it-? |
| Max (N) | Today . . . *Begins talking about a YouTube challenge video* |

Following this excerpt, the boys engage in off-task banter interrupted twice by Joshua asking Max for help. Joshua drags one block to the scripting area during this five-minute time period, after which Joshua states, "Okay, we're going to start". The teacher appears in the video and guides the two toward selecting the next block which is intended to ask the user their name. The boys then begin to chant "what's your name?" and make nonsense sounds repeatedly.

The two only engaged in Monitoring behaviors once during the almost 42-minute coding session. That excerpt appears in Table 6, with co-occurring Confusion and Disagreement utterances. It is important to note that Max, as driver, is the only one interacting with the programming environment at the beginning of this excerpt.

For the remaining 20 minutes of the coding session, the boys worked without adult assistance only twice; once they worked to change what the sprite said and the other time they changed the sprite's size. However, three times, adults intervened for a total of 9 minutes and 37 seconds. Of the five tasks the students were expected to complete, the boys correctly completed two and both were done with adult assistance.

### *5.2. Mixed dyad status*

The remaining 10 dyads comprise the Mixed Status. Noted previously, there is minimal literature to guide how young students' individual CS self-efficacies and performance may influence their collaborative and regulated discourse, so we explored these dyads for patterns to see what findings emerged. The following ENA model (see, Figure 2) displays

**Table 6.** Low dyad: excerpt 2.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Max (D) | This is- I don't get this. So I put all of these in the if [block]? |
| Joshua (N) | *Inaudible (not wearing headset)* |
| Max (D) | So all of them in? Oh- just the say [block]. |
| Joshua (N) | *Inaudible (not wearing headset)* |
| Max (D) | If- |
| Joshua (D) | *Shifts laptop toward himself, becomes audible, and assumes the Driver role* I know, I know. |
| Max (N) | Answer is right. Answer is right. Answer. |
| Joshua (D) | *Pretends to remove the codeblocks Max just added* No, no dang. |
| Max (N) | Dude. Stop. Dude. |
| Joshua (D) | What, dang? Just kidding. |
| Max (N) | I'm so confused. |
| Joshua (D) | *Opens a new window and Google searches "I'm so confused" and both boys laugh* |

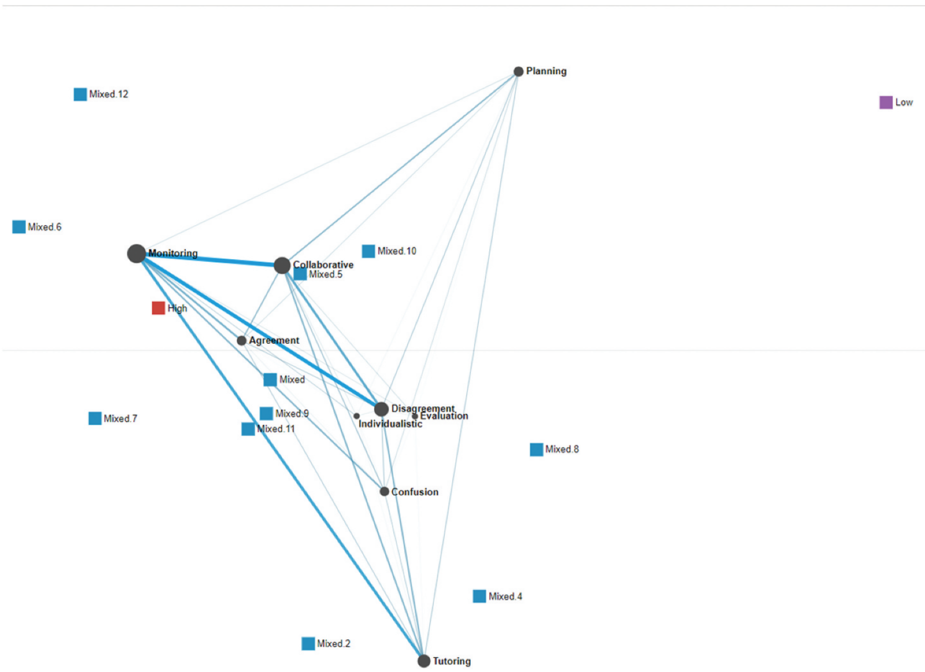


**Figure 2.** ENA mixed dyad network diagram displaying mixed dyads in blue, high dyads in red, and low dyads in purple.

these dyads. Here, our analysis made use of patterns we saw when we explored the scores the individuals within the dyads earned (Table 2). For example, students within dyads 2 and 10 each had one member who scored high on self-efficacy and low on conceptual understanding and the other who scored low on both measures. After exploring the data to identify patterns, we found three distinct groups: one in which one member has high self-efficacy (ie. dyads 2 and 10), one in which the students have opposite scores (ie., dyads 4, 5, and 9), and one in which the students are both high in self-efficacy and low in conceptual understanding (ie., dyads 7 and 11). The remaining dyads (6, 8, and 12) did not form a clear group, although at least one member of the dyad had high conceptual understanding. Table 2 provides individual and dyad status information.

We now explore the three main groups, in particular their similarities in terms of regulated collaborative discourse. Moreover, we present additional network models that visually highlight the nature of these mixed groups' discourse.

#### *5.2.1. Mixed: dyads 2 and 10 (one student with high self-efficacy)*

Dyad 2 (Mila and Nathan) and 10 (David and Leo) each had one member who scored high on the self-efficacy measure, Mila and David, respectively. The two dyads share some commonalities in their use of collaborative discourse (see, Figure 3) – they both have frequent co-occurrences that include Disagreement, Confusion, and Tutoring. We have conceptualized disagreement in our work as both social and academic conflict. Dyad 2 engaged in more social disagreements, whereas dyad 10 had more academic, or task-based, conflicts.


**Figure 3.** ENA mixed dyad network diagram displaying mixed dyad group with 1 high self efficacy member (in green) compared to high dyads (in red).

Dyad 2's (Mila and Nathan) social disagreements largely centered around Nathan telling Mila to shut up (five times) and him resisting either the switch to being the navigator or to taking input from Mila when he was driver, as seen in Table 7.

An excerpt from David and Leo, demonstrating their use of the disagreement utterances, appears in Table 8. The goal of the activity was to have Alonzo (the sprite) put on sunglasses when the sun appeared and to remove them when the clouds appeared.

In this case, the boys disagreed about whether they had completed the activity and, upon realizing they had not, how to code Alonzo correctly. Immediately after this exchange, Leo changed the subject and began pretending to be an airline pilot, leaving David to complete the work alone.

### *5.2.2. Mixed: dyads 4, 5 and 9 (students with opposite scores)*

This group includes dyads whose individual members have opposite scores on a measure – –where one was low, the other was high. A qualitative exploration of this groups' transcripts revealed intriguing instances of Tutoring, although this network (see, Figure 4) is visually similar to the prior network. Because tutoring was conceptualized as both requesting and offering of task-related assistance, we found it beneficial to examine which dyads and who within the dyad made which type of tutoring statement.

Dyad 4's Samantha and Andi made tutoring-tagged statements that fell into both categories of requesting and offering assistance. Samantha had low self-efficacy and high conceptual understanding, whereas Andi had high self-efficacy and low conceptual

**Table 7.** Mixed dyad: Mila and Nathan.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Nathan (N) | No, okay, okay, okay, three, okay, so press, press that. Press that. |
| Mila (D) | Press what? |
| Nathan (N) | Press the, that. |
| Mila (D) | Oh, this? You can't see it, Nathan. |
| Nathan (N/D) | Damn it. *Takes control of the laptop* |
| Mila (N) | Uh, it's my turn. |
| Nathan (D) | Just freaking kill this guy [the sprite]. |
| Mila (N) | Hey, it's my turn. |
| Nathan (D) | Um, can we just freaking kill this guy somehow? How'd you throw him to the trash can? |
| Mila (N) | Oh, I think I know how to do it. |
| Nathan (D/N) | Yeah, do it, then. *Pushes laptop back to Mila* |
| Mila (D) | I actually don't, but I just wanted the computer. Wait, where's the mouse? What the- oh, whoa. |
| Nathan (N) | *Tries to take the laptop back, but Mila resists*. Do you want me to try- no, don't move it, over there. Oh, yes, kill that guy. |
| Mila (D) | Stop it. |
| Nathan (N) | No, make it, make it, like, right there. |
| Mila (D) | Okay. |

**Table 8.** Mixed dyad: Leo and David.

| Student (Driver or Navigator) | Utterance |
|---|---|
| David (D) | Wait, what? How did he switch to . . . How did he get sunglasses? |
| Leo (N) | Oh, we did it. |
| David (D) | No, we didn't. |
| Leo (N) | Yeah, we did. It's . . . You only need to modify [Alonzo], so put your- |
| David (D) | Oh, I know what I did. |
| Leo (N) | If block inside. If the sun is out, [Alonzo] puts on glasses. |
| David (D) | Yeah, Leo, look, you can't put that in there and if you do it like that- |
| Leo (N) | Well then, [Alonzo] doesn't . . . |
| David (D) | Yeah, [Alonzo] is brainless. |
| Leo (N) | . . . Dude, just grab the sunglasses and put them on. |

understanding. Twice Samantha sought help, seemingly from an adult in the room, without directly requesting it (ie., "Um, we need help" and "I need help again" said loudly enough for an adult to hear, but without saying their name). In the exchange in Table 9, Andi has just become driver and Samantha is offering very explicit help. The statements about colors refer to the coding categories, such as Motion, Sound, or Variables.

Rylee and Amber are students in Dyad 5. Rylee scored low on self-efficacy and high on conceptual understanding, whereas Amber had high self-efficacy and low conceptual understanding. Rylee used slightly more tutoring and monitoring statements than Amber, and like Samantha in Dyad 4, she generally offered very direct help to her partner, such as in the excerpt in Table 10.

Sahil and Ezra were dyad 9. Sahil scored low on self-efficacy, but high on conceptual understanding, whereas Ezra had high self-efficacy and low conceptual understanding. Ezra drove the dyad's tutoring statement usage. Twice, he requested help (ie., "how do we make him smaller?" and "What are we supposed to do on task three?"); however, the bulk of his tutoring-tagged utterances fell into the category of offering assistance to Sahil. One such exchange appears in Table 11.

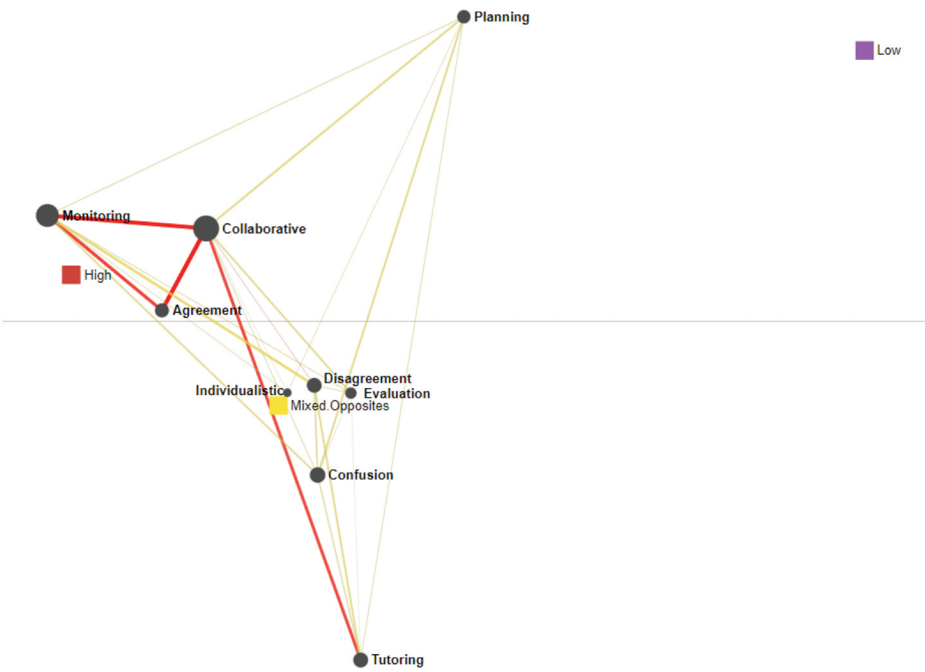


**Figure 4.** ENA mixed dyad network diagram displaying mixed dyad group with individual students with opposite scores (in yellow) compared to high dyads (in red).

**Table 9.** Mixed dyad: Samantha and Andi.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Samantha (N) | Yeah. Okay, so now, um, go to the clear, clear . . . |
| Andi (D) | What, like, what, um- |
| Samantha (N) | It's, like, teal. |
| Andi (D) | Oh, so, pen? |
| Samantha (N) | Um, yeah, clear. |
| Andi (D) | Oh, stop it. There we go. |
| Samantha (N) | Okay, now it's point in direction. |
| Andi (D) | Yeah, but- |
| Samantha (N) | And that's blue. |
| Andi (D) | And that's blue, like, look . . . Motion? |
| Samantha (N) | Yeah. |
| Andi (D) | And what is it? |
| Samantha (N) | Um, point and direction. |
| Andi (D) | Point . . . Oh. |
| Samantha (N) | And it's 90. |
| Andi (D) | Which is already 90, so that's perfect. |
| Samantha (N) | Okay, so go to . . . It's the same thing. Go to X100. |
| Andi (D) | So I have to do the numbers? |
| Samantha (N) | Yeah. |

In this instance, Ezra attempts to help Sahil by telling him where to go next (Sounds), despite Sahil not requesting such assistance. Sahil clarifies his coding logic, to which Ezra responds that he wanted to complete that code. Immediately after this exchange, Ezra shifted his attention to a classmate.

**Table 10.** Mixed dyad: Rylee and Amber.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Rylee (N) | I remember how to do it. Who should we use, though? |
| Amber (D) | And just do, let's just do Alonzo since it's easy, you know. |
| Rylee (N) | Okay. |
| Amber (D) | I don't like it but- |
| Rylee (N) | Make him 75, though. |
| Amber (D) | (laughs) No, I wanna make, can we make him giant? |
| Rylee (N) | Move him over, though. |
| Amber (D) | Yeah. |
| Rylee (N) | So we can fit the words. |
| Amber (D) | The words will fit- |
| Rylee (N) | No, hit move. Hit move. |
| Amber (D) | Hold on. Where? |
| Rylee (N) | Tap arrow. Like . . . |
| Amber (D) | Oh. There. |
| Rylee (N) | Yeah, that's good. |

**Table 11.** Mixed dyad: Sahil and Ezra.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Ezra (N) | Sounds. |
| Sahil (D) | No, move and then makes sounds. You have to move first. |
| Ezra (N) | Move what? |
| Sahil (D) | Um, the guy. |
| Ezra (N) | But I wanna move him. |
| Sahil (D) | You are not the driver (laughs) |

*5.2.3. Mixed: Dyads 7 and 11 (both students have high self-efficacy and low conceptual understanding*

Dyads in this group had members who scored high on the self-efficacy measure, but low on the conceptual understanding instrument. The most commonly tagged utterances across the two dyads included Confusion, Individualistic, and Disagreement (see, Figure 5).

Dyad 7's Emma and Malachi's disagreements seemed to largely stem from Emma's confusion and Malachi's working individualistically, without the input of Emma. Fifteen times across the coding activity, Emma asked "what are you doing?" or "wait, what?" or "then what?" An excerpt (Table 12) demonstrates how the two are trying to recall the steps their teacher took when modeling the concepts.

Allegra and Chloe, in Dyad 11, show similar patterns as Dyad 7. In their case, Allegra utters more confusion-tagged statements, whereas Chloe utters more disagreement-tagged statements. In their excerpt (Table 13), the girls are trying to complete the first user input activity.

## 6. Discussion

This study was driven by the knowledge that student academic regulation is important for collaborative functioning and that both prior conceptual understanding and self-efficacy are likely to influence group dynamics in general and collaborative regulation in

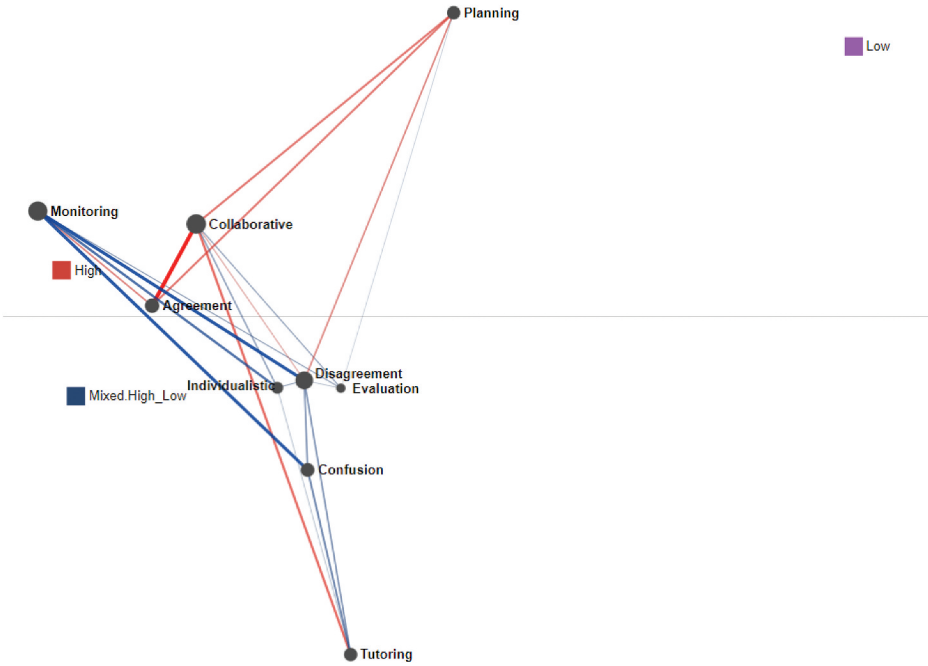


**Figure 5.** ENA mixed dyad network diagram displaying mixed dyad group with high self efficacy and low conceptual understanding (in dark blue) compared to high dyads (in red).

**Table 12.** Mixed dyad: Emma and Malachi.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Emma (N) | No, it did so- she did something- |
| Malachi (D) | Oh, no, it's just an if. |
| Emma (N) | No, it was an if else. |
| Malachi (D) | No, it was just an if. |
| Emma (N) | Oh yeah, it was just an if, I think. I don't know. |
| Malachi (D) | It is. Trust me. I'm smart. |
| Emma (N) | Okay. Okay. |
| Malachi (D) | I talk- |
| Emma (N) | It's fine, Malachi. It's fine. And then . . . |
| Malachi (D) | Di- di- di- di- |
| Emma (N) | And then what? I can't remember. |
| Malachi (D) | Um, I know. We do, uh, operate right? Uh . . . |
| Emma (N) | What are you doing? |
| Malachi (D) | Um, um, this. |
| Emma (N) | No, it's supposed to say your name, not, like . . . Just keep working. Leave it there. |
| Malachi (D) | Just watch. Um, okay, right, no. It's, um . . . |
| Emma (N) | What are you doing? |
| Malachi (D) | Sensing, that's right. |
| Emma (N) | That one, yes. |
| Malachi (D) | No, it's this one. |
| Emma (N) | Wait what are we doing? |

particular. Here the context was a rapidly emerging instructional strategy – pair programming in elementary classrooms. As such, we utilized validated measures of conceptual

**Table 13.** Mixed dyad: Allegra and Chloe.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Chloe (N) | No, no it has to say . . . "hello world", in the "hello" |
| Allegra (D) | In the "hello"? |
| Chloe (N) | Yes. |
| Allegra (D) | I can't get this. Why is it not . . . |
| Chloe (N/D) | I know that I'm not supposed to do this but can I try? *They switch roles* |
| Allegra (N) | Yeah, just . . . yeah. There I just . . . Four, let's just make it . . . |
| Chloe (D) | No she said two. Just click the green flag. And see what's happening, what is your name, Allegra? |
| Allegra (N) | You have to say "hello Allegra" |
| Chloe (D) | Hello. No, no, no. |
| Allegra (N) | Yes we do. We have to say "hello Allegra". |
| Chloe (D/N) | Oh my goodness, that is not what we are supposed to do, trust me. Go to "green". *Chloe relinquishes the laptop to Allegra* |
| Allegra (D) | Aren't you going to try this? |
| Chloe (N) | It will work, but not the way you want it to. |
| Allegra (D) | I am just, I am just going to try. |
| Chloe (N) | It will work, but not the way we want . . . are supposed to. |

understanding and self-efficacy to cluster students into high, mixed, and low dyadic groupings, after which their collaborative discourse was closely examined for discursive patterns. These patterns are discussed further here, by the three sub-research questions.

***6.1. How do individuals within dyads with similarly high self-efficacy and conceptual understanding scores collaboratively regulate their learning?***

The one High dyad (both members had high self-efficacy and conceptual understanding) readily engaged in grounding through question-asking. A brief excerpt appears in Table 14. Here the girls are deciding how their sprite will ask the user their name.

Such incremental grounding occurs when the participants must determine each other's meaning, understanding, or expectations in order to move forward together (Brennan, 1998). In this case, the girls talked through how they envisioned the sprite saying hello. Moreover, they two made consistent use of each other's knowledge, sometimes acknowledging that they did not know the answer or the correct next step but moving ahead with "let's just try it". Otherwise, they sought adult assistance twice during the coding activity. The first time, the girls could not find a certain block, but by the time the teacher appeared, they had found the correct block and had moved on. The other time, a researcher was walking by as Melanie said "we don't understand it really" which prompted the researcher to stop and assist. In many ways, this dyad works and speaks in ways we would anticipate for students with high self-efficacy and understanding of the task; they collaborate well, are

**Table 14.** High dyad: grounding excerpt.

| Student (Driver or Navigator) | Utterance |
|---|---|
| Melanie (D) | And then how do we want to say that, what's your name? |
| Poppy (N) | Just like hello. |
| Melanie (D) | Well no, but like how do we want to say, "Hi"? |
| Poppy (N) | Oh, do you wanna be like French or- |
| Melanie (D) | You can say, "Bonjour". We can say, "Hi, Hello, Hola". What do you wanna do? |
| Poppy (N) | Hm. I'm not sure. Maybe hola. Oh, or we could do all of them. |

balanced in their use of asking for help and simply learning from mistakes, and both offer and receive tutoring considerately. They are also willing to take risks and to learn from those ventures.

***6.2. How do individuals within dyads with similarly low self-efficacy and conceptual understanding scores collaboratively regulate their learning?***

The one Low dyad, Max and Joshua, (both members had low self-efficacy and conceptual understanding) were regularly off-task and worked on task only when externally regulated by an adult. When an adult was near or sitting with them, their discourse shifted to focus more so on the teacher; that is, the student who was driving interacted with the teacher who talked as though they were navigator, often leaving the student navigator silent. There were instances in which the boys sought their partner's assistance. Joshua asked or demanded Max help him a total of 10 times during the coding activity. No such appeals came from Max to Joshua, and in response to the 10 requests for assistance, Max either ignored them and changed the subject (e.g. "What's up guys. Back to the kitchen. We're always cooking some videos".) or deflected responsibility (e.g. "Dude, you're the driver".). Neither of the boys explicitly asked for teacher assistance, but it was given when it became apparent that the boys were struggling. The Low dyad behavior aligns in ways we would expect for students with low self-efficacy and low conceptual understanding. We are unaware if they completed the self-report measures with fidelity; as such we are left to question if they lacked interest in the tasks but could have worked better on more engaging activities, or if they needed more explicit modeling from the teacher, or if other in-system supports would have helped. We posit that a dyad with matching low self-efficacy for and conceptual understanding of a topic may not make an ideal partnership. At minimum, they likely need additional support from the teacher and/or system.

***6.3. How do individuals within dyads with mixed self-efficacy and conceptual understanding scores collaboratively regulate their learning?***

The Mixed dyads presented three clear sub-groupings. The first (dyads 2 and 10) featured one student in each dyad with high self-efficacy and low conceptual understanding and the other student with low scores on both measures. These dyads had more co-occurrences of Monitoring and Disagreement. However, how disagreement appeared in their discourse differed. Dyad 2 had more social disagreement, including the rude command to "shut up" when an idea was shared. In this case, the student with the high self-efficacy, Mila, rarely started the disagreements and was the one being told to shut up. Dyad 10's disagreements were more academic, or task-based, and the student with the low self-efficacy, Leo, often changed the subject right after a disagreement occurred. Conflict is challenging for most students, perhaps especially so for those engaged in a new learning activity and where they may have little experience from which to pull. Here we see students with low self-efficacy engaging in either aggressive off-task behavior or task avoidance when asked to engage. Students of this age (9 to 11) likely lack conflict resolution skills and this may be why we see them defaulting to demands to shut up or the shifting of topics (Johnson et al., 1994). Because of this, we believe curricular support around the value of disagreeing with a partner (ie., to learn of different perspectives or problem solutions) ought to be implemented and modeled (Okada & Matsuda, 2019).

The next Mixed dyad grouping included students with opposite scores from their partner. Dyads 4, 5, and 9 each had one high self-efficacy and low conceptual understanding student and one low self-efficacy and high conceptual understanding student. The discursive commonality among these dyads was their use of Tutoring and Monitoring statements. Two of the three dyads (4 and 5) used tutoring to both request and offer assistance, whereas dyad 9 saw one student (Ezra, with high self-efficacy) almost exclusively offer assistance to his partner, who did not reciprocate. In all cases, however, the offering of assistance generally appeared as directives. In other words, the students were telling their partner exactly what to do next, step by step. We hypothesize that, despite one student having higher conceptual understanding of the CS concepts needed for the task, students of this age struggle with scaffolding learning for others such that the other student can engage in a supported, inquiry process. In fact, Webb et al. (1995) found that the amount of elaboration one student offers another while peer tutoring predicts the level of constructive activity the tutee carries out in response. That is, when a peer tutor explained how to complete a task in a supportive fashion, the tutee was more likely to work constructively on the problem. Utilizing peer modeling during the teaching of concepts is one way to familiarize students with seeing one another as instructors. Additionally, prompts or sentence starers such as "when did we see something like this before?" may help engage both students in the inquiry process.

The final Mixed dyad group included two dyads (7 and 11) in which both students had high self-efficacy and low conceptual understanding. The discursive features of this group included the use of Confusion, Disagreement, and Individualistic-tagged statements. In both dyads, one student offered more confusion-tagged statements, which contributed in some way to the disagreements and/or their partner opting to work individualistically. Knowing that self-efficacy tends to predict performance, we are intrigued by this group and their collaborative discourse. We are left to question how accurately the students completed the self-efficacy measure, or if their understanding of coding (referenced in the measure) was incongruous with either the conceptual understanding measure or the coding activity, or both. Additionally, we posit that having two students with high self-efficacy, but low conceptual understanding may not make for a productive pairing. While they may have the motivation to work through the task, the lack of collective core knowledge does not provide the foundational tools needed to resolve the confusion and work towards a solution path. Their overall level of confusion may have been mitigated by classroom or in-system supports, such as reminders of task goals or work previously completed (see, Weinstein et al., 2000). Additionally, students need to learn the value of collaboration and, when asked for example, "what are you doing?" they need to realize this may mean their partner no longer has a shared understanding of goals and they are no longer collaborating effectively.

## 7. Conclusions

In this study, we examined upper elementary student pair programmers' scores on two self-report measures (CS self-efficacy and CS conceptual understanding) in addition to their collaborative discourse, tagged using an academic and social regulation framework. The dyads' scores on the measures grouped them into High, Low, and Mixed status categories. We then used ENA to visually model and qualitatively examine the three dyadic categories. Results indicate that the High and Low categories spoke to one another in ways expected, given their CS self-efficacy and conceptual understanding scores. However, the Mixed category provided insight into how 4th and 5th grade students' self-efficacy and understanding of a domain interact as they collaborate.

Self-efficacy is a predictor of performance, with positive estimates of one's competence likely bolstering effort and contributing to higher achievement (Bandura et al., 1999). Maladaptive, or inaccurate, estimates of one's efficacy, however, can be problematic as they may lead to a lack of awareness of when to seek help and when to apply appropriate learning strategies (Bandura et al., 1999). Our study extended findings at the individual level to the dyadic level that groups who were uniformly high or low in both measures performed much as expected. Panadero et al. (2015) report on the empirical relationship between individual and group regulation, such that higher individual regulation predicts higher group regulation. Our findings partially support this work. Using discourse markers as evidence of both social and task regulation, our High and Low dyads regulated themselves well and poorly, respectively. More interesting were the mixed dyads. The subgroup where both were high in self-efficacy but low conceptual understanding (7 and 11) paralleled findings found at the individual level that some base level of foundational knowledge is necessary regardless how motivated one might be (Baek, Xu, Han & Cho, 2015). The hope that complementary pairings of high/low scores (4, 5, and 9) might, as a dyad, make up for deficits of their partner, had mixed findings. This subgroup reinforced prior literature that both task and social regulation play important roles with these young students (Hadwin et al., 2010; Zimmerman, 2000). As also seen in the last sub-group (2 and 10), low self-efficacy seems to be tied to anti-social regulatory behavior, to the detriment of task-related discourse. The findings with these Mixed dyads indicate that although there may be one individual whose regulatory behaviors are optimal, they may not be sufficient to positively drive the dyadic regulation. This may be a feature of the students' age in our study or working in pairs. Additional research is needed to unpack this dynamic with a larger, more diverse population.

Our findings may support the work of others such as Okal et al. (2020) who found that the younger the students are in a coding education intervention, the more likely they are to report statistically significant differences in self-efficacy (Okal et al., 2020). Therefore, the earlier students are exposed to programming through mastery experiences, the more their self-efficacy will be positively affected (Mazman & Altun, 2013; Resnick et al., 2009). The students in our study all participated in the same weekly intervention at school, but may have had different at-home and out-of-school experiences that influenced their interest in and self-efficacy for CS. Efforts to improve low student self-efficacy are varied. Crippen and Earl (2007) found that students in an online Chemistry class had improved self-efficacy and performance when provided a worked example and the requirement to self-explain. This type of intervention would be straightforward and appropriate to integrate into a CS setting, especially one that uses pair programming where students are expected to talk through their thinking.

## 8. Limitations and future directions

Regarding students' assessment of their capabilities to successfully complete certain CS-specific actions, we were struck by the finding that the only High group was made up of girls and the only Low group was made up of boys. Although these pairings work against the prevailing literature that boys are generally overconfident in their assessment of their capabilities (Beyer et al., 2003; Cheryan et al., 2009) and that they tend to perform better than girls in CS (Kallia & Sentance, 2018), the small sample of only two dyads prevents us from drawing broad conclusions. We do believe that this is worth exploring further with this young age group, however. Both self-efficacy and CS conceptual understanding played a role in shaping dyadic regulation and discourse. They did not have to be equally distributed within the dyad for the task to move forward, although how it unfolded and who benefited was mixed and nuanced. Our findings point to the potential for both teacher and systems to provide supports that help fill gaps in both self-efficacy and conceptual understanding in ways that moves the work forward and builds these two dimensions in students.

Our study was limited in sample size and diverse socio-demographic characteristics and our findings need to be interpreted with respect to these limitations. Future research could utilize this analysis approach with a larger and more diverse sample. Moreover, analyses around students' prior experiences in programming are important to incorporate. Lastly, we were unable to gather complete post-intervention data due to the COVID-19 pandemic; this study would have benefited from a thorough pre-post analysis of both the students' CS Attitudes and their CS conceptual understanding.

Lastly, personality differences may have influenced not only the students' individual performance on the CS conceptual understanding assessment, the ways he or she self-assessed their CS efficacy, and their experience while collaboratively coding, but it may also have affected the pair's ability to engage in problem-solving. In particular, Pietarinen et al. (2019) found that if students report feeling confident, they were far more likely to actively participate, collaborate with, and support their group members than if they were feeling less confident, or insecure. We believe that some of the differences we saw in our groups likely hinge on the individual students' belief in their ability to complete the CS work as it is in tension with their beliefs about their partner's capabilities. In other words, one student's high self-efficacy might not be enough to overcome the lack of support and disinterest in the task a partner may have offered. As such, pairing students by similar collaboration interests or self-efficacy may be a consideration for future research (see, Campe et al., 2019). It is essential for students to appropriately assess their abilities, as inaccurate understanding can prevent students from seeking assistance. Computer science is one subject area that can easily provide such feedback to students as they can run their code and immediately know the accuracy of their work. Future work may consider expanding on Roll et al.'s (2011) finding that a self-assessment tutor improves students' accuracy. Similarly, future efforts in learning analytics may consider the use of on-screen prompts to guide students' collaborative discourse.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This material is based upon work supported by the National Science Foundation under Grant No. DRL-1721160. Additionally, this work was funded in part by the National Science Foundation (DRL-1661036, DRL-1713110), the Wisconsin Alumni Research Foundation, and the Office of the Vice

Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies, cooperating institutions, or other individuals.

## ORCID

Jessica Vandenberg http://orcid.org/0000-0001-6497-1840
Eric Wiebe http://orcid.org/0000-0002-5920-5225

## References

Arastoopour, G., Shaffer, D. W., Swiecki, Z., Ruis, A., & Chesler, N. C. (2016). Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis. *International Journal of Engineering Education*, *32*(3), 1492–1501. https://vtb.engr.wisc.edu/wp-content/uploads/sites/733/2017/01/Teaching-and-Assessing-Engineering-2015.pdf

Baek, Y., Xu, Y., Han, S., & Cho, J. (2015). Exploring effects of intrinsic motivation and prior knowledge on student achievements in game-based learning. *Smart Computing Review*, *5*(5), 368–377. https://doi.org/10.6029/smartcr.2015.10.001

Bandura, A., Freeman, W., & Lightsey, R. (1999). *Self-efficacy: The exercise of control*. Springer.

Beyer, S., & Haller, S. (2006). Gender differences and intragender differences in computer science students: Are female CS majors more similar to male CS majors or female nonmajors? *Journal of Women and Minorities in Science and Engineering*, *12*(4). https://doi.org/10.1615/JWomenMinorScienEng.v12.i4.50

Beyer, S., Rynes, K., Perrault, J., Hay, K., & Haller, S. (2003). Gender differences in computer science students. In *Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education* (pp. 49–53). Reno, Nevada: ACM.

Bradley, R. L., Browne, B. L., & Kelley, H. M. (2017). Examining the influence of self- efficacy and self-regulation in online learning. *College Student Journal*, *51*(4), 518–530. https://www.learntechlib.org/p/191368/

Brennan, S. E. (1998). The grounding problem in conversations with and through computers. In S. R. Fussell & R. J. Kreuz (Eds.), *Social and Cognitive Approaches to Interpersonal Communication* (pp. 201–225). Mahwah, NJ: Erlbaum.

Bressler, D. M., Bodzin, A. M., Eagan, B., & Tabatabai, S. (2019). Using epistemic network analysis to examine discourse and scientific practice during a collabora- tive game. *Journal of Science Education and Technology*, *28*(5), 553–566. https://doi.org/10.1007/s10956-019-09786-8

Brosnan, M. J. (1998). The impact of computer anxiety and self-efficacy upon perfor- mance. *Journal of Computer Assisted Learning*, *14*(3), 223–234. https://doi.org/10.1046/j.1365-2729.1998.143059.x

Campe, S., Green, E., & Denner, J. (2019). *k-12 pair programming toolkit*. ETR.

Canfora, G., Cimitile, A., Garcia, F., Piattini, M., & Visaggio, C. A. (2007). Evaluating performances of pair designing in industry. *Journal of Systems and Software*, *80*(8), 1317–1327. https://doi.org/10.1016/j.jss.2006.11.004

Cassel, L., McGettrick, A., Guzdial, M., & Roberts, E. (2007). The current crisis in computing: What are the real issues? *ACM SIGCSE Bu l letin*, *39*(1), 329–330. https://doi.org/10.1145/1227504.1227426

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*(6), 1045. https://doi.org/10.1037/a0016239

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.

Crippen, K. J., & Earl, B. L. (2007). The impact of web-based worked examples and self-explanation on performance, problem solving, and self-efficacy. *Computers & Education*, *49*(3), 809–821. https://doi.org/10.1016/j.compedu.2005.11.018

de Jong, F., Kollöffel, B., van der Meijden, H., Staarman, J. K., & Janssen, J. (2005). Regulative processes in individual, 3d and computer supported cooperative learning contexts. *Computers in Human Behavior*, *21*(4), 645–670. https://doi.org/10.1016/j.chb.2004.10.023

Denner, J., Werner, L., Campe, S., & Ortiz, E. (2014). Pair programming: Under what conditions is it advantageous for middle school students? *Journal of Research on Technology in Education*, *46*(3), 277–296. https://doi.org/10.1080/15391523.2014.888272

Erkens, G., Prangsma, M., & Jaspers, J. (2006). Planning and coordinating activities in collaborative learning. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative Learning, Reasoning, and Technology* (pp. 233–263). Mahwah: Erlbaum.

Gillies, R. M., & Boyle, M. (2010). Teachers' reflections on cooperative learning: Issues of implementation. *Teaching and Teacher Education*, *26*(4), 933–940. https://doi.org/10.1016/j.tate.2009.10.034

Goode, J. (2010). Connecting K-16 curriculum & policy: Making computer science engaging, accessible, and hospitable for underrepresented students. In *Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, Milwaukee, Wisconsin (pp. 22–26). ACM.

Grigg, S., Perera, H. N., McIlveen, P., & Svetleff, Z. (2018). Relations among math self efficacy, interest, intentions, and achievement: A social cognitive perspective. *Contemporary Educational Psychology*, *53*(53), 73–86. https://doi.org/10.1016/j.cedpsych.2018.01.007

Hadwin, A., & Oshige, M. (2011). Self-regulation, coregulation, and socially shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record*, *113*(2), 240–264. https://doi.org/10.1177/016146811111300204

Hadwin, A. F., Oshige, M., Gress, C. L., & Winne, P. H. (2010). Innovative ways for using gstudy to orchestrate and research social aspects of self-regulated learning. *Computers in Human Behavior*, *26*(5), 794–805. https://doi.org/10.1016/j.chb.2007.06.007

Hadwin, A. F., Wozney, L., & Pontin, O. (2005). Scaffolding the appropriation of self-regulatory activity: A socio-cultural analysis of changes in teacher–student discourse about a graduate research portfolio. *Instructional Science*, *33*(5), 413–450. https://doi.org/10.1007/s11251-005-1274-7

He, J., & Freeman, L. A. (2010). Are men more technology-oriented than women? The role of gender on the development of general computer self-efficacy of college students. *Journal of Information Systems Education*, *21*(2), 203–212. https://www.proquest.com/docview/745597250?pq-origsite=gscholar&fromopenview=true

Hinckle, M., Rachmatullah, A., Mott, B., Boyer, K. E., Lester, J., & Wiebe, E. (2020). The relationship of gender, experiential, and psychological factors to achievement in computer science. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, Trondheim, Norway (pp. 225–231). ACM.

Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2012). Task-related and social regulation during online collaborative learning. *Metacognition and Learning*, *7*(1), 25–43. https://doi.org/10.1007/s11409-010-9061-5

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in cscl. *Educational Psychologist*, *48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006

Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported col- laborative learning: How to support collaborative learning? How can technologies help? *Educational Psychologist*, *51*(2), 247–265. https://doi.org/10.1080/00461520.2016.1158654

Jepson, A., & Perl, T. (2002). Priming the pipeline. *ACM SIGCSE Bu l letin*, *34*(2), 36–39. https://doi.org/10.1145/543812.543826

Johnson, D. W., Johnson, R., Dudley, B., & Acikgoz, K. (1994). Effects of conflict resolution training on elementary school students. *The Journal of Social Psychology*, *134*(6), 803–817. https://doi.org/10.1080/00224545.1994.9923015

Johnstone, B. (2017). *Discourse analysis*. John Wiley & Sons.

Kallia, M., & Sentance, S. (2018). Are boys more confident than girls? The role of calibration and students' self-efficacy in programming tasks and computer science. In *Proceedings of the 13th workshop in primary and secondary computing education*, Potsdam, Germany (pp. 1–4). ACM.

Kong, S.-C., Chiu, M. M., & Lai, M. (2018). A study of primary school students' in- terest, collaboration attitude, and programming empowerment in computational thinking education. *Computers & Education*, *127*, 178–189. https://doi.org/10.1016/j.compedu.2018.08.026

Kumpulainen, K., & Mutanen, M. (1999). The situated dynamics of peer group interaction: An introduction to an analytic framework. *Learning and Instruction*, *9*(5), 449–473. https://doi.org/10.1016/S0959-4752(98)00038-3

Lishinski, A., Yadav, A., Good, J., & Enbody, R. (2016). Learning to program: Gender differences and interactive effects of students' motivation, goals, and self-efficacy on performance. In *Proceedings of the 2016 acm conference on international computing education research*, Melbourne, Australia (pp. 211–220). ACM.

Mason, S. L., & Rich, P. J. (2020). Development and analysis of the elementary student coding attitudes survey. *Computers & Education*, *153*, 103898. https://doi.org/10.1016/j.compedu.2020.103898

Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. N. (2017). Programming experience promotes higher stem motivation among first-grade girls. *Journal of Experimental Child Psychology*, *160*, 92–106. https://doi.org/10.1016/j.jecp.2017.03.013

Mazman, S., & Altun, A. (2013). The effect of introductory to programming course on programming self-efficacy of ceit students. *Journal of Instructional Technologies & Teacher Education*, *2* (3), 24–29. https://dergipark.org.tr/en/pub/jitte/issue/25082/264710

McDowell, C., Werner, L., Bullock, H. E., & Fernald, J. (2006). Pair programming improves student retention, confidence, and program quality. *Communications of the ACM*, *49*(8), 90–95. https://doi.org/10.1145/1145287.1145293

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://doi.org/10.11613/BM.2012.031

Mejias, M., Jean-Pierre, K., Washington, G., & Burge, L. (2019). Underrepresented groups threats to belonging in computing. In *Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, Minneapolis, Minnesota (pp. 1–4). IEEE.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal*, *25*(1), 95–111. https://doi.org/10.1080/0141192990250107

Missiroli, M., Russo, D., & Ciancarini, P. (2016). Learning agile software development in high school: An investigation. In *Proceedings of the 38th international conference on software engineering companion*, Austin, Texas (pp. 293–302). ACM.

Mitchell, A., Purchase, H. C., & Hamer, J. (2009). Computing science: What do pupils think? In *Proceedings of the 14th annual acm sigcse conference on innovation and technology in computer science education*, Paris, France (p. 353). ACM.

Mohamed, A. (2019). Designing a cs1 programming course for a mixed-ability class. In *Proceedings of the western Canadian conference on computing education*, Calgary, Alberta, Canada (pp. 1–6). ACM.

Moos, D. C., & Azevedo, R. (2008). Monitoring, planning, and self-efficacy during learning with hypermedia: The impact of conceptual scaffolds. *Computers in Human Behavior*, *24*(4), 1686–1706. https://doi.org/10.1016/j.chb.2007.07.001

Moyer, L., Klopfer, M., & Ernst, J. V. (2018). Bridging the arts and computer science: Engaging at-risk students through the integration of music. *Technology and Engineering Teacher*, *77*(6), 8–12. https://proxying.lib.ncsu.edu/index.php/login?url=https://www.proquest.com/scholarly-journals/bridging-arts-computer-science-engaging-at-risk/docview/2036401417/se-2?accountid=12725

Okada, Y., & Matsuda, T. (2019). Development of a social skills education game for elementary school students. *Simulation & Gaming*, *50*(5), 598–620. https://doi.org/10.1177/1046878119880228

Okal, G., Yildirim, B., & Timur, S. (2020). The effect of coding education on 5th, 6th and 7th grade students' programming self-efficacy and attitudes about technol- ogy. *Educational Policy Analysis and Strategic Research*, *15*(2), 143–165. https://doi.org/10.29329/epasr.2020.251.8

Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice*, *41*(2), 116–125. https://doi.org/10.1207/s15430421tip4102_8

Panadero, E., Kirschner, P. A., Järvelä, S., Malmberg, J., & Järvenoja, H. (2015). How individual self-regulation affects group regulation and performance: A shared regulation intervention. *Small Group Research*, *46*(4), 431–454. https://doi.org/10.1177/1046496415591219

Pietarinen, T., Vauras, M., Laakkonen, E., Kinnunen, R., & Volet, S. (2019). High school students' perceptions of affect and collaboration during virtual science inquiry learning. *Journal of Computer Assisted Learning*, *35*(3), 334–348. https://doi.org/10.1111/jcal.12334

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. H. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Elsevier.

Porter, L., Guzdial, M., McDowell, C., & Simon, B. (2013). Success in introductory programming: What works? *Communications of the ACM*, *56*(8), 34–36. https://doi.org/10.1145/2492007.2492020

Potter, J. (1998). Discursive social psychology: From attitudes to evaluative practices. *European Review of Social Psychology*, *9*(1), 233–266. https://doi.org/10.1080/14792779843000090

Rachmatullah, A., Akram, B., Boulden, D., Mott, B., Boyer, K., Lester, J., & Wiebe, E. (2020). Development and validation of the middle grades computer science concept inventory (mg-csci) assessment. *EURASIA Journal of Mathematics, Science and Technology Education*, *16*(5), em1841. https://doi.org/10.29333/ejmste/116600

Rachmatullah, A., Wiebe, E., Boulden, D., Mott, B., Boyer, K., & Lester, J. (2020). Development and validation of the computer science attitudes scale for middle school students (mg-cs attitudes). *Computers in Human Behavior Reports*, *2*(2), 100018. https://doi.org/10.1016/j.chbr.2020.100018

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for all. *Communications of the ACM*, *52*(11), 60–67. https://doi.org/10.1145/1592761.1592779

Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011). Metacognitive practice makes perfect: Improving students' self-assessment skills with an intel- ligent tutoring system. In *International conference on artificial intelligence ineducation* (pp. 288–295). Verlag, Berlin, Heidelberg: Springer.

Román-González, M., Pérez-González, J.-C., & Jiménez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? Criterion validity of the com- putational thinking test. *Computers in Human Behavior*, *72*, 678–691. https://doi.org/10.1016/j.chb.2016.08.047

Ruis, A., Rosser, A. A., Quandt-Walle, C., Nathwani, J. N., Shaffer, D. W., & Pugh, M. (2018). The hands and head of a surgeon: Modeling operative competency with multimodal epistemic network analysis. *The American Journal of Surgery*, *216*(5), 835–840. https://doi.org/10.1016/j.amjsurg.2017.11.027

Saab, N., van Joolingen, W., & van Hout-wolters, B. (2012). Support of the collaborative inquiry learning process: Influence of support on task and team regulation. *Metacognition and Learning*, *7*(1), 7–23. https://doi.org/10.1007/s11409-011-9068-6

Schunk, D. H. (1995). Self-efficacy, motivation, and performance. *Journal of Applied Sport Psychology*, *7*(2), 112–137. https://doi.org/10.1080/10413209508406961

Schunk, D. H., & Zimmerman, B. J. (2012). Self-regulation and learning. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of Psychology,* (Vol. 7, pp. 59–78). John Wiley.

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and inter- action data. *Journal of Learning Analytics*, *3*(3), 9–45. https://doi.org/10.18608/jla.2016.33.3

Shah, N., Lewis, C., & Caires, R. (2014). *Analyzing equity in collaborative learning situations: A comparative case study in elementary computer science*. International Society of the Learning Sciences.

Sheldrake, R. (2016). Differential predictors of under-confidence and over-confidence for mathematics and science students in england. *Learning and Individual Differences*, *49*, 305–313. https://doi.org/10.1016/j.lindif.2016.05.009

Siebert-Evenstone, A. L., Irgens, G. A., Collier, W., Swiecki, Z., Ruis, A. R., & Shaffer, D. W. (2017). In search of conversational grain size: Modelling semantic structure using moving stanza windows. *Journal of Learning Analytics*, *4*(3), 123–139. https://doi.org/10.18608/jla.2017.43.7

Syed, M., & Nelson, S. C. (2015). Guidelines for establishing reliability when coding narrative data. *Emerging Adulthood*, *3*(6), 375–387. https://doi.org/10.1177/2167696815587648

Tran, Y. (2018). Computer programming effects in elementary: Perceptions and career aspirations in stem. *Technology, Knowledge and Learning*, *23*(2), 273–299. https://doi.org/10.1007/s10758-018-9358-z

Tsan, J., Vandenberg, J., Zakaria, Z., Wiggins, J. B., Webber, A. R., Bradbury, A., and Boyer, K. E. (2020). A comparison of two pair programming configurations for upper elementary students. In *Proceedings of the 51st acm technical symposium on computer science education*, Portland, Oregon (pp. 346–352). ACM.

Vandenberg, J., Rachmatullah, A., Lynch, C., Boyer, K. E., & Wiebe, E. (2021a). Interaction effects of race and gender in elementary cs attitudes: A validation and cross-sectional study. *International Journal of Child-Computer Interaction*, *29*, 100293. https://doi.org/10.1016/j.ijcci.2021.100293

Vandenberg, J., Rachmatullah, A., Lynch, C., Boyer, K. E., & Wiebe, E. (2021b). The relationship of cs attitudes, perceptions of collaboration, and pair programming strategies on upper elementary students' cs learning. In *Proceedings of the 26th acm conference on innovation and technology in computer science education v. 1*, Virtual Event, Germany (pp. 46–52). ACM.

Vandenberg, J., Tsan, J., Boulden, D., Zakaria, Z., Lynch, C., Boyer, K. E., & Wiebe, E. (2020). Elementary students' understanding of cs terms. *ACM Transactions on Computing Education (TOCE)*, *20*(3), 1–19. https://doi.org/10.1145/3386364

Vandenberg, J., Tsan, J., Hinckle, M., Lynch, C., Boyer, K. E., & Wiebe, E. (2021). *Discursive gender differences in upper elementary students' pair programming*. American Educational Research Association).

Vandenberg, J., Tsan, J., Zakaria, Z., Boulden, D., Lynch, C., Boyer, K. E., & Wiebe, E. (2020). *Elementary learners' regulation in computer-supported collaborative learning environments*. American Educational Research Association).

Wang, H., Wang, C., & Wu, F. (2020). How multimedia influence group interaction in stem education an epistemic network analysis for online synchronous collaborative learning. In *2020 International symposium on educational technology (iset)*, Bangkok, Thailand (pp. 303–306). IEEE.

Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology*, *87*(3), 406. https://doi.org/10.1037/0022-0663.87.3.406

Weinstein, C. E., Husman, J., & Dierking, D. R. (2000). Self-regulation interventions with a focus on learning strategies. In M. Boekaerts, P. Pintrich, & M. Seidner (Eds.), *Self-regulation: Theory, research, and applications* (pp. 727–747). Academic Press.

Werner, L., Denner, J., Campe, S., & Kawamoto, D. C. (2012). The fairy performance assessment: Measuring computational thinking in middle school. In *Proceedings of the 43rd acm technical symposium on computer science education*, Raleigh, NC (pp. 215–220). ACM.

Werner, L. L., Hanks, B., & McDowell, C. (2004). Pair-programming helps female computer science students. *Journal on Educational Resources in Computing(JERIC)*, *4*(1), 4–es. https://doi.org/10.1145/1060071.1060075

Williams, L., Kessler, R. R., Cunningham, W., & Jeffries, R. (2000). Strengthening the case for pair programming. *IEEE Software*, *17*(4), 19–25. https://doi.org/10.1109/52.854064

Yardi, S., & Bruckman, A. (2007). What is computing? Bridging the gap between teenagers' perceptions and graduate students' experiences. In *Proceedings of the third international workshop on computing education research*, Atlanta, Georgia (pp. 39–50). ACM.

Zhong, B., Wang, Q., & Chen, J. (2016). The impact of social factors on pair pro- gramming in a primary school. *Computers in Human Behavior*, *64*(64), 423–431. https://doi.org/10.1016/j.chb.2016.07.017

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts & P. R. Pintrich (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.