Investigating Collaborative Problem Solving Behaviors during STEM+C Learning in Groups with Different Prior Knowledge Distributions

Caitlin Snyder, Vanderbilt University, caitlin.r.snyder@vanderbilt.edu
Cai-Ting Wen, Soochow University, leafhere@gmail.com
Nicole M. Hutchins, University of Florida, nicole.hutchins@ufl.edu
Caleb Vatral, Vanderbilt University, caleb.m.vatral@vanderbilt.edu
Chen-Chung Liu, National Central University, ccliu@cl.ncu.edu.tw
Gautam Biswas, Vanderbilt University, gautam.biswas@vanderbilt.edu

Abstract: In collaborative problem-solving (CPS), students work together to solve problems using their collective knowledge and social interactions to understand the problem and progress towards a solution. This study focuses on how students engage in CPS while working in pairs in a STEM+C (Science, Technology, Engineering, Mathematics, and Computing) environment that involves open-ended computational modeling tasks. Specifically, we study how groups with different prior knowledge in physics and computing concepts differ in their information pooling and consensus building behaviors. In addition, we examine how these differences impact the development of their shared understanding and learning. Our study consisted of a high school kinematics curriculum with 1D and 2D modeling tasks. Using an exploratory approach, we performed in-depth case studies to analyze the behaviors of groups with different prior knowledge distributions across these tasks. We identify effective information pooling and consensus building behaviors in addition to difficulties students faced when developing a shared understanding of physics and computing concepts.

Introduction

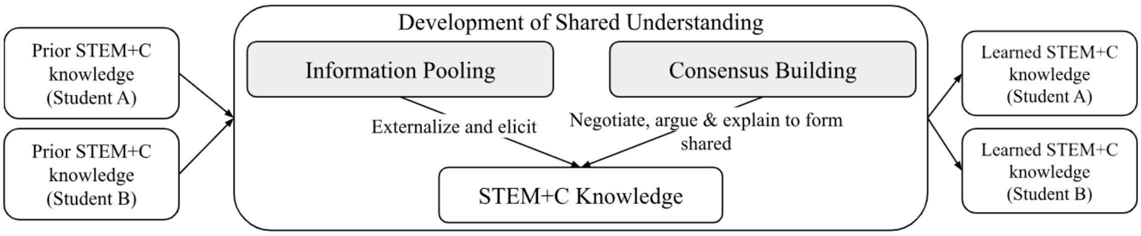
In collaborative problem solving, students jointly construct knowledge through conversations to reach a *shared understanding* and apply it to problem-solving tasks (OECD, 2015). Our study focuses on groups working collaboratively to build computational models in an open-ended learning environment (OELE). Effective computational modeling and problem solving necessitate integrating the STEM and computing domains (NRC, 2012). Building on prior research that highlights the influence of prior knowledge on learning in single domains (e.g., Zambrano et al., 2019), we investigate how groups' prior knowledge in science and computing influenced their Collaborative Problem Solving (CPS) behaviors as they constructed computational models in kinematics.

Our analysis of students' CPS builds on research that has identified shared understanding as a key component of effective collaborative problem solving (e.g., Baker, 2015; OECD, 2015). Specifically, we study how students develop their shared understanding through: (1) *information pooling*, where students externalize and elicit domain-specific knowledge from members in the group; and (2) *consensus building*, where students use arguments and explanations to negotiate and create shared knowledge and apply it to their problem solving tasks (Meier et al., 2007). Clearly, these conversations combine domain-specific information and social interactions (e.g., elicit knowledge from their partners by asking questions and negotiate differences to form a consensus; Weinberger & Fischer, 2006) to discuss STEM and computing concepts and construct their computational models (Snyder et al., 2019). We predict that these interactions will vary depending on students' prior knowledge within groups, as their existing understanding of the problem impacts the information they seek and share during collaborative efforts. Similarly, research indicates that differences in initial knowledge significantly affect argumentation skills and consensus building (Yang et al., 2015). We utilize this framework (Figure 1) to examine students' collaborative problem-solving behaviors in STEM+C learning.

Our computer based OELE targets synergistic learning in the science and computing domains. In this work, we used a 1D and 2D high school kinematics curriculum that combines inquiry activities, instructional tasks, formative assessments, and model building activities. Instructional and inquiry activities, along with formative assessments, help students learn the primary physics and computing concepts and relations between these concepts. At the end of each unit, students are given a challenge task, which requires them to build a comprehensive computational model. By combining students' conversations and their model building activities in the environment, we adopt an exploratory case study approach to analyze students' information pooling and consensus building behaviors in groups with different types of prior knowledge distributions. Our data comes from students' work in the 1D and 2D challenge tasks and pre-post assessments in science and computing.

Figure 1

CPS Framing: Group's Development of a Shared Understanding of STEM+C Knowledge

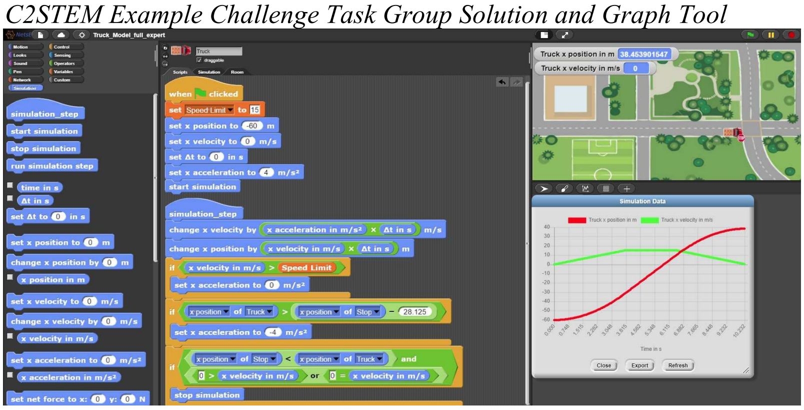


STEM+C Learning Environment

Students work collaboratively in Collaborative Computation STEM (C2STEM; see Figure 2), our open-ended computational modeling environment that adopts a modular approach to help students progressively learn complex science and computing concepts in specific curricular domains, e.g., kinematics (Hutchins et al., 2020). Within C2STEM, students create partial or complete models to study the movement of the objects. Along with animation and variable inspection functions that are displayed on the simulation stage, students have access to graphing and table tools that are updated dynamically at each simulation step to help them debug their evolving models. Students create these models by developing and leveraging their understanding of kinematics (e.g. relationship between position, velocity, and acceleration) and computing knowledge (e.g., initializing and updating variables and applying conditional constructs). S*ynergistic learning*, i.e., the simultaneous learning of science and computing, has been shown to be effective in developing successful solutions (Hutchins et al., 2020) but students may also have difficulties, such as transferring their STEM knowledge to computing constructs to build their computational model (Basu et al., 2016) that can be mitigated through collaboration.

Figure 2

C2STEM Example Challenge Task Group Solution and Graph Tool

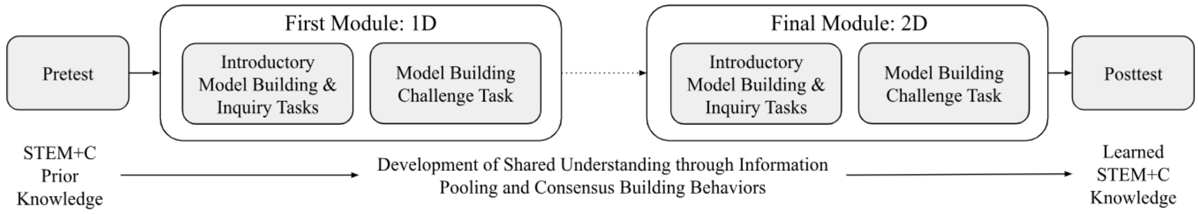


Methods

Each module in our curriculum (Figure 3) comprises inquiry, computational modeling, and formative assessment tasks, developed using a systematic evidence-centered design (ECD) approach (Mislevy & Haertel, 2006). Inquiry tasks utilize the CoSci platform (https://cosci.tw), that provides students a scaffolded simulation-based learning environment to explore physics variables and dynamic processes (Wen et al., 2018). In the 1D challenge, students model a truck's motion speeding up to a speed limit, cruising at that speed limit, and then slowing to a stop at a STOP sign. To construct the model, students had to translate their physics understanding of the relations between position, velocity and acceleration to a computational form that included initializing the necessary variables and modeling the dynamics of the truck movement by updating variables under different conditions. In the final 2D challenge, they simulate a drone dropping packages onto specific targets, applying physics principles like gravity to construct computational models.

Our analysis targets the research question: *How do groups prior knowledge distributions affect their information pooling and consensus building behaviors during computational modeling tasks and how do these behaviors relate to students' STEM+C learning?* We answer this research question adopting an exploratory case study approach where we analyze the differences between groups that were (1) *balanced*, i.e., one student in the pair had high prior knowledge in physics and low prior knowledge in computing, whereas the second student had high prior knowledge in computing and low prior knowledge in physics; and (2) *unbalanced* where one student in the pair had high prior knowledge in physics and computing, and the second student had low prior knowledge in both domains. Specifically, we studied each groups' information pooling and consensus building behaviors across two model-building tasks (1D and 2D acceleration) to understand how their behaviors evolved over the course of the curriculum. We also related these behaviors to students' STEM+C learning gains measured by their pre- to post-test learning gains.

Figure 3
Curriculum Trajectory



Students and Their Data

Our research team conducted a two-month-long study, working with 10th grade high school students, aged 14-15, for two hours a week in a classroom in the United States. None of the students had taken a high school physics course, but some had been introduced to basic kinematics in introductory science classes. Their background in computing varied. 27 students were divided into 13 groups (12 dyads and one triad) assigned based on their pretest scores. The student with the highest total pretest score (i.e., the sum of their pretest scores in kinematics and computing) was matched with the student with the lowest pretest score and so on. The study was approved by the university Institutional Review Board. This included analyzing summative assessment data and computational models, and video and audio data collected using the OBS software on each group's shared laptop. Student conversations were transcribed using Otter.ai and then edited by two researchers. One student in the triad did not consent to data collection so we did not analyze the data for that group.

Each student was categorized as having high or low prior knowledge in physics and computing based on their pretest scores relative to the median. The pre- and post-tests contained four physics questions (17 points) and three computing questions (16 points) in multiple choice and constructed response formats. Students' pretest scores in physics ranged between [5,15] with a median score of 11 (SD = 2.37). Their computing scores ranged between [3,13] with a median score of 10.5 (SD = 2.80). When we looked at group scores, we had a total of 6 *unbalanced* groups, 3 *balanced* groups and 3 groups in which both students had low prior knowledge in physics. In this paper, we adopt an exploratory contrasting case study approach to compare the behaviors of *unbalanced* and *balanced* dyads as these groups had similar overall prior knowledge in physics and computing but the distributions of knowledge across the students differed. We chose two *unbalanced* and two *balanced* groups for in-depth analysis by considering the quality of the collected video and audio data.

Analysis

To evaluate groups' social interactions during information pooling and consensus building, we coded students' utterances using the Weinberger & Fischer (2006) social modes framework (see Table 1). Additionally, we included an off-task label to code discourse that was not related to the computational modeling task. Two members of the research team coded the dialogue with good agreement (Cohen's inter-rater reliability kappa value of 0.77). From these codes, we extracted information pooling and consensus building segments of student dialog. To evaluate groups' domain-specific STEM+C knowledge integration during information pooling and consensus building, we coded each utterance based on the physics or computing concepts discussed in the utterance (Cohen's kappa value of 0.83). The physics concept codes included conversations about position, velocity, acceleration, displacement, and time, and the computing concept codes included Δt, control structures, initializing variables, updating variables, and conditional structures. Leveraging these coded utterances, we calculated a *synergistic score* for each information pooling and consensus building segment using following formula: $SYN = \#\,Utterances_{C} - \#\,Utterances_{PHY}$. This computed value was then normalized to the range [-1,1]; a value closer to 0 indicated high synergistic discourse (i.e., conversations in this segment included concepts in both domains), a value closer to -1 indicated more physics-focused conversation and a value closer to 1 indicated more computing-focused conversation. To evaluate groups' application of STEM+C knowledge, we scored each group's final computational models using a predefined rubric. Learning gains were calculated based on a summative post-test in science and computing that was identical to the pre-test.

Table 1
Information Pooling and Consensus Building Coding Scheme

Social Mode	Description
Information Pooling: "*eliciting information and giving appropriate explanations" (Meier et al., 2007)*	
Elicitation	One student is questioning another about information relating to the task
Externalization	Student(s) are articulating to the other by stating facts, observations and/or narrating their actions in the system
Consensus Building: *"discussing and critically evaluating information in order to make a joint decision" (Meier et al., 2007)*	
Conflict-oriented	During discussion, the students are disagreeing over their interpretation of a concept, model component, or what to do next.
Integration-oriented	During discussion, one student adds a new component to the discussion, integrates a concept/perspective, and/or applies the perspective proposed by the other student
Quick	During discussion, one student makes a suggestion, and their partner accepts it with no further discussion

Results

Table 2 lists the groups' task scores, STEM+C synergistic integration scores, and social dimension metrics for information pooling and consensus building behaviors for the 1D and 2D challenge tasks. Note that most of the groups scored lower in the 2D task and there was a drop in computing performance in G1, G2, and G3. We argue this may be due to the increased computational complexity of this task as G2 and G3's physics performance stayed the same across the tasks while G4 raised their physics score, and subsequently their total score, in the final task. However, as discussed below, G4 received help from another group in the physics component of the final task. In the following subsections, we first contrast groups' CPS behaviors and conclude by analyzing individual students' STEM+C learning in the context of these behaviors and groups' different knowledge distributions. Students' prior knowledge categories were based on their pretest scores in Table 3.

Contrasting Information Pooling and Consensus Building Behaviors

Balanced Groups (G1 and G2)

The information pooling behaviors in G1 and G2 are relatively synergistic ($|\text{Avg SYN}| \leq 0.08$) for both tasks. We hypothesize that the *balanced* groups leveraged synergistic information pooling because each partner externalized their knowledge in their high prior knowledge domain. For example, in the 1D task, S3 often leveraged his physics knowledge to externalize how he believed the model (the motion of the truck) should behave, while S4 often leveraged her computing knowledge to suggest the use of specific computing blocks to effectively simulate the behavior.

However, there were differences in G1 and G2's consensus building behaviors. In G2, consensus building differed between tasks. In the 1D task, conversations were more integrated-oriented (18%) compared to conflict-oriented (12%). Students tended to defer to their partner's ideas rather than challenge them. For example, in the 1D modeling task, S4 (who had high prior knowledge in computing) took the lead by identifying when they needed to use a *change block* (to update a variable based on the previous simulation value) or a *set block* (to set a variable to a specific value) and S3 (who had low prior knowledge in computing) followed along. This reliance on their partner's knowledge did not negatively impact task performance, since the group had the highest score on the 1D model (0.95). In the 2D task, conflict-oriented consensus building increased (22%), with instances of both students challenging each other, such as a disagreement over using a change block, where S3 (who had low prior knowledge in computing) correctly, disagreed saying "*no we just need to do set. Set y velocity"*, resulting in a discussion that concluded with S4 agreeing, *"Oh it has to be at 0, yeah you're right."*

Table 2
Students' Task Scores, STEM+C Knowledge Synergistic Scores and Percentage of Social Interactions

Groups		G1 (*Balanced*)		G2 (*Balanced*)		G3 (*Unbalanced*)		G4 (*Unbalanced*)	
Challenge Task		1D	2D	1D	2D	1D	2D	1D	2D
Total Task Score		0.95	0.63	0.97	0.84	0.89	0.79	0.95	0.97
PHY Task Score		0.94	0.67	1	1	0.94	0.94	0.94	1
C Task Score		0.95	0.6	0.95	0.7	0.85	0.65	0.95	0.95
STEM+C Knowledge during Information Pooling and Consensus Building									
Avg SYN (SD) - Information Pooling		0.02 (0.12)	-0.05 (0.16)	-0.08 (0.15)	-0.08 (0.16)	-0.20 (0.25)	-0.19 (0.30)	-0.10 (0.19)	-0.11 (0.16)
Avg SYN (SD) - Consensus Building		-0.06 (0.23)	0.07 (0.17)	-0.06 (0.17)	0.00 (0.16)	-0.05 (0.10)	-0.01 (0.32)	0.04 (0.15)	-0.09 (0.09)
Social Interactions during Information Pooling and Consensus Building									
Information Pooling	**Elicitation**	11%	12%	12%	22%	11%	11%	16%	5%
	Externalization	**51%**	**44%**	**58%**	**45%**	**66%**	**67%**	**49%**	**77%**
	Total	**62%**	**57%**	**70%**	**67%**	**77%**	**78%**	**65%**	**82%**
Consensus Building	**Conflict-oriented**	5%	9%	12%	**22%**	**11%**	**11%**	16%	5%
	Integration-oriented	**25%**	**24%**	**18%**	10%	10%	10%	**19%**	**12%**
	Quick	8%	10%	1%	1%	2%	1%	0%	0%
	Total	38%	43%	30%	33%	23%	22%	35%	18%

G1 also had an increase in conflict-oriented consensus building (5% to 9%) from the 1D to 2D task but they still favored integration-oriented consensus building behaviors in both tasks (25% and 29%). During these consensus-building segments, S1 often led with new ideas while S2 contributed. For example, when they were initializing the position of the packages, with S2 controlling the laptop, S1 made a suggestion about changing the value saying, *"A little bit less, like 4.5 for now… we want it on top of the other packages"* and S2 agreed, saying *"Yeah we can move the other packages to put them [below]"*. Interestingly, although their discussions remained synergistic, they shifted slightly from a physics focus in the 1D task to a computing focus in the 2D task (Avg SYN = -0.06 to Avg SYN = 0.07). We hypothesize that this may be attributed to the group struggling with the increased computational complexity in the 2D task. They had the lowest score out of all the groups (0.63), and their consensus building centered more around the computing component of the model. Overall, both G1 and G2 had synergistic information pooling and consensus building behaviors in both tasks. They primarily used integration-oriented consensus building behaviors on the 1D task. However, while G1 continued to have similar consensus building behaviors in the 2D task, G2's behaviors evolved into more conflict-oriented discussions.

Unbalanced Groups (G3 and G4)

The information pooling behaviors for G3 and G4 were less synergistic and more physics-focused ($-0.19 \leq$ Avg SYN ≤ -0.10). G3's discourse was primarily information pooling focused on both tasks (77% and 78%, respectively) and was characterized by the high prior knowledge student (S6) narrating actions. For example, in G3, S6 primarily narrated the model construction actions with very little contributions made by S5 (low prior knowledge student). During one segment, while debugging the truck slowing down segment, possibly to elicit collaboration from her partner, S6 expressed a lack of understanding, saying "*I don't know what to do next because I'm confused"* but S5 did not respond and S6 resorted to an ineffective trial and error strategy. This lack of collaboration increased in the 2D task, where there was a 14-minute segment in which S6 tinkered with the model and made only five utterances whereas S5 did not make any. In fact, S6 attempted to collaborate with another group in the 2D task when her partner would not engage. Despite S6 having higher prior knowledge, the complexity of the model building tasks necessitated collaboration (Kirschner, et al., 2011). We hypothesize this lack of collaboration contributed to G3's poor consensus building behaviors and low task performance, with the group scoring the lowest in the 1D task and second lowest in the 2D task.

G4 had comparable amounts of consensus building behaviors as G1 and G2 in the 1D task (35%). At the beginning of the task the low prior knowledge student, S7, took the lead in controlling the laptop mouse while the high prior knowledge student, S8, gave suggestions on what actions to take. They switched between information pooling behaviors (where S8 was narrating) and synergistic consensus building behaviors (Avg SYN = 0.04) as the two students often discussed specific suggestions made by S8. When the group switched control of the laptop, S7 stayed involved and elicited information from S8 as they performed actions. For example, when modeling the truck motion transitioning from cruising to slowing down, S7 asked, *"Question… What are we looking for here?"* with S8 clarifying the current goal, *"looking for how long we need to get this to cruise for because… you don't know where to start decelerating..."*. During the 2D task their behaviors changed considerably as they had more information pooling behaviors (82%) with much less collaboration between the partners. Their struggles with a physics component of the 2D task, resulted in physics-focused information pooling and consensus building behaviors (Avg SYN = -.11 and -0.09, respectively). Eventually, they got another group to give them the answer.

In summary, across the two tasks, students exhibited three information pooling behavior types: (1) synergistic information pooling (G1 and G2, both students externalized knowledge); (2) information pooling externalized by one student after prompting from the other (G4-1D); and (3) non-collaborative information pooling in which one student primarily narrated actions (G3, G4-2D). There were also three consensus building behaviors: (1) primarily integration-oriented in which groups formed a consensus primarily through deferring to their partner but also added different ideas (G1, G2-1D, G4-1D); (2) primarily conflict-oriented consensus building in which groups developed enough shared understanding and individual knowledge to more easily challenge their partner (G1-2D); and (3) minimal consensus building overall (G3, G4-2D).

Table 3
Summative Physics (PHY) and Computing (C) Learning by Group

Group	Student	Pretest			Posttest			Learning Gains (LG)		
		TOTAL	PHY	C	TOTAL	PHY	C	TOTAL	PHY	C
G1 (*Balanced*)	S1	0.67	0.65	0.69	0.77	0.76	0.78	0.10	0.11	0.09
	S2	0.61	0.71	0.50	0.64	0.71	0.56	0.03	0.00	0.06
G2 (*Balanced*)	S3	0.61	0.71	0.50	0.71	0.65	0.78	0.10	-0.06	0.28
	S4	0.65	0.59	0.72	0.76	0.71	0.81	0.11	0.12	0.09
G3 (*Unbalanced*)	S5	0.58	0.65	0.50	0.59	0.65	0.53	0.01	0.00	0.03
	S6	0.74	0.71	0.78	0.83	0.88	0.78	0.09	0.17	0.00
G4 (*Unbalanced*)	S7	0.39	0.53	0.25	0.42	0.35	0.50	0.03	-0.18	0.25
	S8	0.80	0.88	0.72	0.72	0.82	0.63	-0.08	-0.06	-0.09

STEM+C Learning
Table 3 presents the summative assessment results. Students in the *balanced* groups, G1 and G2, had higher overall learning gains (Avg = 0.09) compared to the students in the *unbalanced* groups, G3 and G4 (Avg = 0.01). This result holds across all the *balanced* and *unbalanced* groups in the study as all the *balanced* groups had an average overall learning gains of 0.11 (SD = 0.04, n=3) and all the *unbalanced* groups had an average overall learning gains of 0.04 (SD = 0.12, n=6).

When considering the *balanced* groups' STEM+C learning, the students had higher learning gains in the domains they started with low prior knowledge (i.e., low physics prior knowledge students S1 and S4 had learning gains of 0.11 and 0.12 in physics, respectively, while low computing prior knowledge students S2 and S3 had learning gains of 0.06 and 0.28, respectively) suggesting that these groups successfully leveraged their partners' knowledge to develop their own individual knowledge. Interestingly, when considering the physics domain, the students with low prior physics knowledge (S1 and S4) ended up surpassing their partners' physics knowledge by the end of the study: in G1, S1 had a final PHY score of 0.76 while S2's score was 0.71 and in G2, S4 had a final PHY score of 0.71 while S3 had a final PHY score of 0.65. This is partly because S2 and S3 had minimal, if any, learning gains in physics. When investigating the pre-posttests in more detail, S2 in G1 was partially incorrect on a physics-focused 2D question on the pre- and post-test (S1's answer was incorrect on the pre but was partially correct on the post). Since this group struggled with the 2D challenge task, we hypothesize this may be why S2 had no positive learning gains in physics. In G2, S3 correctly answered a 1D graph question in the pre but incorrectly in the post, and S4 answered the same question incorrectly in the pre and post, suggesting that the group had a misunderstanding about graphs. In contrast, all groups improved in computing. Overall, these results suggest that while the balanced groups had overall learning gains, and particularly in their low prior knowledge domains, all students generally gained more computing knowledge.

When considering the *unbalanced* groups' STEM+C learning, the groups had difficulties developing knowledge in both domains. In the computing domain S5, S6, and S8 had minimal (0.03), no, and negative learning gains (-0.09), respectively. The low prior knowledge student in G4, S7, is the only one who showed learning gains in computing (LG.C = 0.25) but they also had the lowest pretest score (0.25) and ended with the lowest posttest score (0.50) in computing. Similarly, the high prior knowledge student in G3, S6, is the only student who gained physics knowledge (LG.PHY = 0.17). Overall, in the *unbalanced* groups knowledge development at the end of the curriculum was also unbalanced as in G3, the high prior knowledge student (S6) had learning gains in physics and their partner's learning gains were minimal, while in G4, the low prior knowledge student (S7) had learning gains in computing and S8 had negative overall learning gains (-0.08).

The *balanced* groups' summative assessments showed consistency in the pre-post answers (the same question incorrectly on the pre and post-test, suggesting a knowledge gap that was not addressed by the intervention or a question that was answered incorrectly on the pre but correctly answered on the post, suggesting that knowledge was gained through the intervention). But the *unbalanced* group (e.g., S8) showed mixed results (i.e., both correct to incorrect and incorrect to correct answers). Overall, the *unbalanced* pairing seems to have helped the low prior knowledge student gain knowledge in computing but negatively impacted the high prior knowledge student through the introduction of new knowledge misunderstandings.

Finally, when considering these STEM+C learning results in the context of their information pooling and consensus building behaviors, the results imply that the collaborative, synergistic information pooling and consensus building behaviors G1 and G2 exhibited led to effective learning overall. Like prior research that has identified disagreements are an important component to individual learning during CPS activities (e.g., Roschelle & Teasley, 1995), we hypothesize that the transition from integrated-oriented consensus building to more conflict-oriented consensus building from the 1D to 2D task helped G2's individual STEM+C learning (as they had more overall knowledge development than G1 who consistently had integration-oriented consensus building behaviors across both tasks). In addition, the integration-oriented consensus building behaviors G4 exhibited in the 1D task seem to be partially responsible for S7's gaining computing knowledge through model construction with suggestions from their partner. However, the results suggest that the less collaborative information pooling and minimal consensus building behaviors that G3 exhibited in both tasks, and G4 in the 2D task, negatively impacted STEM+C learning overall as both G3 and G4 had more difficulty developing knowledge.

Discussion and conclusions

We found that when both partners participated in information pooling, they developed better shared understanding and STEM+C learning. Information pooling with contributions by both partners provides a base of shared facts that students can leverage to develop a shared understanding through consensus building (Baker, 2015. Our results imply that an increase in consensus building behaviors over time linked to pooled information leads to increased shared understanding through critical analyses of relevant STEM+C concepts. While integration-oriented consensus building had a positive impact, our results suggest that conflict-oriented consensus building is a key indicator of increased shared understanding by both partners, resulting in higher individual STEM+C knowledge gains. We also identified difficulties groups had in developing shared understanding, such as a lack of collaboration during information pooling and consensus building, and this negatively impacted STEM+C learning. Sometimes, groups' development of a shared understanding may cause new misunderstandings if students develop and integrate incorrect knowledge. When detected, a teacher (or agent) can intervene to suggest using more effective CPS behaviors.

Our results are consistent with previous research on prior knowledge distribution in groups. For example, Deiglmayr & Schalk (2015) found that knowledge interdependence among individuals with complementary prior knowledge increases interactive engagement and fosters rich constructive discourse. We found that *balanced* groups overall exhibited more collaborative behaviors compared to the *unbalanced* groups. While previous research has shown that students with low prior knowledge in a domain perform better when working collaboratively as compared to when they work individually (Zambrano, et al. 2019), results on the impact of such unbalanced pairings on the high prior knowledge student are conflicted (Gijlers & De Jong, 2005; Zhang, et al. 2015). Our study confirms the conflicting STEM+C learning we see in the *unbalanced* groups. This exploratory analysis is limited due to its small sample size. In future work, we will extend such analysis to more dyads and include groups who lack prior knowledge overall (e.g., groups with low prior knowledge in physics). Future work will also leverage these results to develop supports to help students employ more effective CPS behaviors and combine their STEM and computing knowledge to construct computational models.

References

Baker, M. J. (2015). Collaboration in collaborative learning. *Interaction studies*, *16*(3), 451-473.

Basu, S., Biswas, G., Sengupta, P., Dickes, A., Kinnebrew, J. S., & Clark,D. (2016). Identifying middle school students' challenges in computational thinking-based science learning. Research and Practice in Technology Enhanced Learning, 11(1), 1–35.

Deiglmayr, A., & Schalk, L. (2015). Weak versus strong knowledge interdependence: A comparison of two rationales for distributing information among learners in collaborative learning settings. Learning and instruction, 40, 69-78.

Gijlers, H., & De Jong, T. (2005). The relation between prior knowledge and students' collaborative discovery learning processes. Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 42(3), 264-282.

Hutchins, N. M., Biswas, G., Maróti, M., Lédeczi, Á., Grover, S., Wolf, R., ... & McElhaney, K. (2020). C2STEM: A system for synergistic learning of physics and computational thinking. Journal of Science Education and Technology, 29(1), 83-100.

Kirschner, F., Paas, F., & Kirschner, P. A. (2011). Task complexity as a driver for collaborative learning efficiency: The collective working-memory effect. *Applied Cognitive Psychology*, *25*(4), 615-624.

Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning*.

Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. Educational Measurement: Issues and Practice, 25(4), 6-20.

NRC - National Research Council. (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: The National Academies Press.

OECD (2015). PISA 2015. Collaborative problem solving.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In Computer supported collaborative learning. Springer, Berlin, Heidelberg.

Snyder, C., Hutchins, N., Biswas, G., Emara, M., Grover, S., & Conlin, L. (2019). Analyzing students' synergistic learning processes in physics and CT by collaborative discourse analysis. In *Computer-supported collaborative learning*.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. Computers & education, 46(1), 71-95.

Wen, C. T., Chang, C. J., Chang, M. H., Fan Chiang, S. H., Liu, C. C., Hwang, F. K., & Tsai, C. C. (2018). The learning analytics of model-based learning facilitated by a problem-solving simulation game. *Instructional Science*, *46*(6), 847-867.

Yang, W. T., Lin, Y. R., She, H. C., & Huang, K. Y. (2015). The effects of prior-knowledge and online learning approaches on students' inquiry and argumentation abilities. *International Journal of Science Education*, *37*(10), 1564-1589.

Zambrano, J., Kirschner, F., Sweller, J., & Kirschner, P. A. (2019). Effects of prior knowledge on collaborative and individual learning. Learning and Instruction, 63, 101214.

Zhang, L., Kalyuga, S., Lee, C. H., Lei, C., & Jiao, J. (2015). Effectiveness of collaborative learning with complex tasks under different learning group formations: A cognitive load perspective. In International Conference on Hybrid Learning and Continuing Education (pp. 149-159). Springer, Cham.