ABSTRACT
To succeed in STEM, students need to connect visual representations to domain-relevant concepts, which is a difficult task for them. Prior research shows that physical representations (that students manipulate with their hands) and virtual representations (that they manipulate on a computer) have complementary advantages for conceptual learning. Further, physical and virtual representations are often embedded into different social classroom practices. Thus, to optimally combine these representation modes, we need to understand what social events prompt students to connect representations to concepts, and if different representation modes afford different social prompts. A multiple-case study with 12 high-school students addresses this question. Student pairs worked with physical and virtual representations of chemistry. Frequent patterns obtained from discourse data show that students incrementally co-construct concept-representation connections, and that instructor prompts are key triggers of these connections for both representation modes. Meta-cognitive statements serve as important prompts in the absence of an instructor when students work with virtual representations. I discuss implications for interventions that combine physical and virtual representations.

Keywords
Physical and virtual representations, educational technology, collaboration, conceptual and social learning processes, STEM.

1. INTRODUCTION
Novice students in science, technology, engineering, and math (STEM) domains grapple with a representation dilemma [1]: they have to use visual representations they have never seen before to make sense of concepts they have not yet learned. Educators often take for granted that students can see meaningful concepts in representations [2]. However, much evidence shows that students struggle in connecting concepts to visual representations [3]. Their failure to make such concept-representation connections can impede their learning [4]. For example, in chemistry, difficulties in making concept-representation connections affect students’ understanding of key concepts related to atomic structure and bonding [5]. This issue applies to most STEM domains: because many key concepts cannot be directly observed, STEM domains heavily rely on visual representations [3]. Thus, STEM instruction typically provides conceptual prompts to help students make concept-representation connections [6].

Research in many STEM domains—including chemistry—shows that different representation modes provide different types of prompts for concept-representation connections [7]. Physical representations are tangible objects that students manipulate with their hands (Figure 1, top). In physical representations, haptic sensory input, experiences of movement, and continuous changes serve as prompts by making concepts intuitively accessible [7, 8]. By contrast, virtual representations are digital visualizations that students manipulate via mouse or text input (Figure 1, bottom). In virtual representations, visualizations and manipulations of invisible processes and immediate feedback can serve as prompts for concept-representation connections [7]. Thus, physical and virtual representations serve complementary roles in prompting for students to make concept-representation connections [7, 9].

Besides providing different types of conceptual prompts for concept-representation connections, physical and virtual representations may provide different types of social prompts. Social prompts are discourse events that elicit collaborative co-construction of such connections [10]. Such events can emerge from student-student or student-instructor interactions. Because physical representations are typically used in collaborative contexts, interactions among students and instructors may prompt concept-representation connections [11]. By contrast, virtual representations are embedded in educational technologies that provide help in making concept-representation connections. In this context, students may work individually or collaboratively, typically with less help from an instructor [12]. Hence, interactions with instructors may be less important in prompting concept-representation connections. Thus, because physical and virtual representations are embedded in different social classroom practices, they may yield different social prompts for concept-representation connections.

Figure 1. Physical representations (top) and virtual representations (bottom) of chemical molecules

Proceedings of the 9th International Conference on Educational Data Mining 478
Considering what social events serve as prompts for concept-representation connections is important for the design of instructional interventions that combine physical and virtual representations. Prior research has not investigated whether different representation modes afford different types of social prompts for concept-representation connections. At a theoretical level, addressing this question will help us understand the mechanisms by which representation modes affect students' ability to make concept-representation connections. It will also help us understand why one representation mode may be more effective than another for a given concept. At a practical level, it will allow us to design instructional activities that take advantage of the social prompts that the different representation modes afford.

The goal of this paper is to take a first step towards identifying social prompts of concept-representation connections for physical and virtual representation modes. To this end, I used a multiple-case study approach; specifically, I observed and recorded collaborative discourse among six student pairs over an extended learning period. Case-study approaches are particularly appropriate for investigating how social processes unfold over a longer learning intervention within the given social classroom practices [13]. The study compared two instructional contexts: (1) student pairs working with physical representations while receiving support from an instructor and (2) student pairs working with virtual representations embedded in an educational technology.

To identify social prompts of concept-representation connections, I applied frequent pattern mining to discourse data. This analysis identified social prompts that are successful for both representation modes and social prompts that were specific to a particular representation mode. I discuss implications for blending interventions that combine physical and virtual representations.

2. METHODS

2.1 Multiple-Case Study

Participants were 12 students from a small charter high school in the Midwestern U.S. The study was conducted as part of a chemistry workshop. Students had very limited prior knowledge about the concepts and the visual representations. The study took place as part of an in-school workshop on 3 days spread across 4 weeks. Each study day was 3h long. Prior to day 1, the teacher gave an introduction to chemical bonding. On day 1, students received an introduction into collaborative strategies and then worked on the chemistry workshop materials for the remaining study days.

All students were randomly assigned to pairs for the duration of the study. For each study day, the pairs were randomly assigned to a sequence of representation mode (i.e., physical-then-virtual, virtual-then-physical). For example, a pair might be assigned to the physical-then-virtual order for day 1. This pair would work with physical representations for the first half of day 1 and then switch to virtual representations for the second half of day 1. On day 2, the pair was randomly assigned to a new sequence.

The workshop covered basic concepts related to the polarity of chemical bonds. Students were presented with the visual representations shown in Figure 1: Lewis structures, ball-and-stick models, space-filling models, and electrostatic potential maps. Each was presented in the physical and virtual mode. When working with physical representations, students received a worksheet that asked them to construct a physical representation of a molecule, answer questions about the target concepts (e.g., about electronegativity) and about how the representation depicts these concepts. Each student pair was teamed up with an instructor—a research assistant who was trained on facilitating student collaboration and on the chemistry concepts covered. Instructors provided feedback and assistance as students solved the problems.

Virtual representations were integrated in an educational technology for chemistry: Chem Tutor [14]: a type of intelligent tutoring system designed specifically to help students make concept-representation connections. To this end, Chem Tutor provides interactive virtual representations that students manipulate to solve problems about bonding. Chem Tutor prompts students to reflect on how each visual representation depicts particular concepts. Chem Tutor provides error-specific feedback and hints on demand. Chem Tutor was shown to significantly enhance learning of chemistry knowledge and conceptual understanding of representations [14]. While working with Chem Tutor, students could request help from an instructor who circulated the classroom.

2.2 Analysis

The goal of the analysis was to identify social events that prompt students’ concept-representation connections and to investigate whether these prompts differ between representation modes.

The first step in the analysis was to code discourse data. All interactions among students and instructors were video-taped and transcribed. To develop a coding scheme, we used a grounded, bottom-up approach: we summarized discourse utterance-by-utterance to discover emerging themes. Next, we formalized these themes as codes, and then applied the codes to the discourse data. The coding scheme comprises 45 codes (see Table 1 for examples). Inter-rater reliability was substantial with kappa = .77.

The second step in the analysis was to identify discourse segments in which students succeed in making a concept-representation connection, defined as establishing the relation between a visual feature in a representation and the domain-relevant concept it illustrates [6]. Hence, a concept-representation connection was operationalized as an utterance made by a student that correctly refers to a concept and a representation (e.g., Table 2, #5).

The third step in the analysis was to operationalize social events that may prompt students to make concept-representation connections. In principle, any aspect of student-student or instructor-student discourse could serve as a social prompt: mentioning a concept, encouragement, evaluating, a meta-cognitive statement, a mistake, etc. Hence, I considered any code as a potential prompt.

The fourth step in the analysis was to specify the unit of analysis. Because I was interested in social events as prompts, I defined two consecutive discourse turns as the unit of analysis (i.e., utterances by two different speakers). I segmented the discourse data in the following way. First, I identified turns with concept-representation connections (e.g., Table 2, row 5). Second, I identified the two prior turns and considered them as a case (e.g., rows 3-4 in Table 2). This case was labeled as ‘connection present’ (i.e., a concept-representation connection occurs in the next turn). Third, I segmented the remaining discourse data such that two consecutive turns serve as a case (e.g., rows 1-2 in Table 2), labeled as ‘connection absent’ (i.e., no concept-representation connection in the next turn). Thus, each case was composed of two consecutive turns, labeled as connection-present/absent, annotated with codes, speaker (student or instructor) and mode (physical or virtual). Table 3 shows an overview of the dataset.

The final step in the analysis was to search for social events that trigger concept-representation connections. Given the focus on social mechanisms, I was interested in discovering which codes co-occur in collaborative discourse. To this end, I used frequent pattern mining to identify undirected patterns that describe which
Table 1. Subset of codes in the coding scheme with examples from the transcripts.

<table>
<thead>
<tr>
<th>Code</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concept</td>
<td>Utterances that relate something to a scientific concept</td>
<td>“They want to be able to make a complete number, a complete number of the eight on the outside”</td>
</tr>
<tr>
<td>Concept-request</td>
<td>Suggesting / prompting utterances that relate something to a concept</td>
<td>“What’s the rule for the bonding?”</td>
</tr>
<tr>
<td>Representation</td>
<td>Utterances that relate something to the representation; utterances that explain information shown by a representation</td>
<td>[pointing at a representation] “So, one, two, three, four, five. He have five.”; [pointing at a representation] “So, wait, that’s carbon?”</td>
</tr>
<tr>
<td>Representation-request</td>
<td>Suggesting / prompting utterances that relate something to the representation; utterances that explain information shown by a representation</td>
<td>“By looking at the Lewis structure, can you answer the question about electronegativity?”; “What are these things [points at dots in Lewis structure]?”</td>
</tr>
<tr>
<td>Assent</td>
<td>Expression of approval or agreement</td>
<td>“yeah”; “ok”; “I know.”; “Mhmhm.”</td>
</tr>
<tr>
<td>Meta-confusion</td>
<td>Utterances about oneself that describe confusion about how to proceed or about a concept, or about not knowing a concept</td>
<td>I don’t know.”; “this is very confusing.”; “Maybe.”; “This is hard.”; “So, now we’re stuck.”; “I don’t get it why it’s lines.”</td>
</tr>
<tr>
<td>Meta-understanding</td>
<td>Utterances about oneself that describe a novel insights or understanding of how to proceed or of a concept</td>
<td>“Got it”; “Well, I know that part”; “I like this explanation.”; “then I was like, well, duh”; “We’ve been making this so much harder than it is!”</td>
</tr>
<tr>
<td>Reading</td>
<td>Reading the problems statement or instructions or hints / feedback from Chem Tutor</td>
<td>“well it says right here that, “Choose the letters that show each atom,”</td>
</tr>
<tr>
<td>Explanation</td>
<td>Utterances that explain / elaborate a concept</td>
<td>“But when they say dinitrogen, means they bonded.”; “I’ll give a little bit more help.”; “So, carbon has more electrons than hydrogen.”</td>
</tr>
<tr>
<td>Explanation-request</td>
<td>Suggesting / prompting utterances that explain / elaborate a concept</td>
<td>“So what do you think that is?”; “Could you try, try to put as a complete sentence”; “But why?”; “How did you know?”</td>
</tr>
<tr>
<td>Metaphor</td>
<td>Utterances that use a metaphor, intuitive example, embellished language to describe an abstract concept</td>
<td>“To make it lock on kind of.”; “can I borrow your electrons”; “It’s the same pulling forces.”; “So, like magnetic, plus and minus.”</td>
</tr>
</tbody>
</table>

Table 2. Excerpt transcript showing 4 turns before a concept-representation connection (turn #5), with codes assigned to each turn. All student names are fake.

<table>
<thead>
<tr>
<th>#</th>
<th>Speaker</th>
<th>Utterance</th>
<th>Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bridget</td>
<td>Electronegativity are the same so makes it covalent which is no difference.</td>
<td>Concept</td>
</tr>
<tr>
<td>2</td>
<td>Adriana</td>
<td>[reads] Does the Lewis structure show the polarity? Why or why not? Um. I’d say I feel like no, be- Well, yeah. I don’t know.</td>
<td>Reading; metaphor; representation-request; concept-request</td>
</tr>
<tr>
<td>3</td>
<td>Bridget</td>
<td>What does polarity mean?</td>
<td>Explanation; representation-request; concept-request</td>
</tr>
<tr>
<td>4</td>
<td>Instructor</td>
<td>Polarity means plus and minus. This [points at representation] By looking at this one, can you see it has like electronegativity or stuff. Polarity means that-</td>
<td>Explanation; metaphor; representation-request; concept-request</td>
</tr>
<tr>
<td>5</td>
<td>Adriana</td>
<td>I mean, like yeah, it doesn’t like show really like the pulling or the not pulling or the same.</td>
<td>Explanation; representation-request; concept-request; metaphor</td>
</tr>
</tbody>
</table>

codes often occur together [15, 16]. I ran this algorithm separately for cases with connections present and absent and for physical and virtual representations. Essentially, this analysis discovered:

1. Frequent patterns for cases with concept-representation connections present for physical representations
2. Frequent patterns for cases with concept-representation connections absent for physical representations
3. Frequent patterns for cases with concept-representation connections present for virtual representations
4. Frequent patterns for cases with concept-representation connections absent for virtual representations

Comparing findings 1 and 2 identified prompts of concept-representation connections for physical representations. Comparing findings 3 and 4 identified prompts of concept-representation connections for virtual representations. Comparing findings 1 and 3 identified differences between representation modes.

3. RESULTS

In the following, I first discuss which discourse patterns were found to prompt concept-representation connections with physical representations or with virtual representations. Then, I compare the physical and virtual representation modes.

3.1 Physical models

To identify prompts of concept-representation connections with physical representations, I considered patterns found only for cases with a present concept-representation connection (i.e., cases that correspond to two turns followed by a concept-representation connection). Table 4 shows statistics for the patterns.

Several results are worth noting. First, it stands out that all patterns involve either a reference to a concept or to a representation.
Table 3. Number of cases by representation mode and speaker.

<table>
<thead>
<tr>
<th>Representation mode</th>
<th>Label</th>
<th>Connection present</th>
<th>Connection absent</th>
<th>Student</th>
<th>Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical</td>
<td></td>
<td>229 (7.33%)</td>
<td>2,895 (92.67%)</td>
<td>2,115 (67.70%)</td>
<td>1,009 (32.30%)</td>
</tr>
<tr>
<td>Virtual</td>
<td></td>
<td>67 (3.28%)</td>
<td>1,976 (96.72%)</td>
<td>1,780 (86.13%)</td>
<td>263 (12.87%)</td>
</tr>
</tbody>
</table>

Table 4. Frequent patterns for physical representations (underlined: instructor utterances, italics: patterns that overlap with virtual representations).

<table>
<thead>
<tr>
<th>Frequent pattern</th>
<th>Support</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. instructor-assent; student-concept</td>
<td>0.100</td>
<td>0.410</td>
</tr>
<tr>
<td>2. instructor-assent; student-representation</td>
<td>0.087</td>
<td>0.377</td>
</tr>
<tr>
<td>3. instructor-representation-request; instructor-concept-request</td>
<td>0.074</td>
<td>0.684</td>
</tr>
<tr>
<td>4. student-representation; student-concept</td>
<td>0.201</td>
<td>0.803</td>
</tr>
<tr>
<td>5. instructor-assent; student-representation; student-concept</td>
<td>0.083</td>
<td>0.536</td>
</tr>
</tbody>
</table>

This finding suggests that it may be easiest for students to make a concept-representation connection if discourse is already focused on the concept or representation. A related finding is that 3 of 5 patterns include references to both concepts and representations—either as a request to relate to concepts and representations by the instructor (#3 in Table 4) or by the students themselves (#4 and #5). These patterns have the highest support and confidence. Hence, students may be particularly likely to make a concept-representation connection if already occurs in previous discourse.

Second, 4 of 5 patterns involve instructor utterances. This finding suggests that instructors may be better than students at prompting concept-representation connections.

Finally, 3 of 5 patterns include assent by the instructor. Assent is defined as agreement with a previous statement (see Table 1), often in the form of encouragement (e.g., “mhm”). In the identified patterns, such encouragement co-occurs with references to a concept or to a representation (or both) provided by one of the students or by the instructor. This finding suggests that encouragement by the instructor—when discourse is already focused on a concept or representation—prompts students to elaborate by making a concept-representation connection.

3.2 Virtual models

To identify triggers of concept-representation connections with virtual representations, I considered patterns found only for cases with a present concept-representation connection. Table 5 shows statistics for these patterns.

The following findings stand out. First, all patterns include a reference to a concept or to a representation. Hence, students may be likely to make a concept-representation connection if discourse is already focused on a concept or on a representation. A related result is that 7 of 16 patterns include a reference to both concept and representation (either as request by the instructor, or a direct reference to both by the instructor or the student). These patterns...
had the highest support and confidence. Hence, students may be particularly likely to deepen their discussion about a connection if prior discourse already focuses on the connection.

Second, 7 of 16 patterns involve instructor utterances. This ratio seems surprisingly high, given that students worked without the instructor for most of the time. Recall that when working with virtual representations, instructor support was available only upon request, and that when students worked with virtual representations, they generated 86.13% of the utterances—instructors only 12.87% (see Table 2). Thus, this finding may indicate that students need help from an instructor to make concept-representation connections, even if they receive technology support.

Third, 6 of 16 patterns include a student (4 of 6) or a student (2 of 6). Recall that assent is defined as agreement with a previous statement (see Table 1), often in the form of encouragement. Again, assent always co-occurs with a reference to a concept or representation. Hence, this finding suggests that encouragement can prompt a concept-representation connection—regardless of whether it is provided by a student or a tutor.

Fourth, 4 of the 7 patterns that involve instructor utterances involve explicit requests for the student to relate to a concept or a representation. This request is always combined with an instructor reference to a concept or to a representation. This finding suggests that prompts to elaborate on a previously mentioned concept or representation yields concept-representation connections.

Finally, 6 of 16 patterns include a meta-cognitive utterance by the student about understanding (3 of 6) or confusion (3 of 6). All of these meta-cognitive utterances co-occur with a reference to a concept and/or a representation. None of these meta-cognitive utterances co-occur with instructor utterances. This finding suggests that meta-cognitive statements about one’s own understanding can prompt concept-representation connections; for example, after a student voices confusion about a concept, the partner may use a representation to explain the concept.

3.3 Comparing physical and virtual modes
Finally, I investigated whether prompts of concept-representation connections differ by representation mode. The following commonalities stand out. First, all patterns found for physical representations were also found for virtual representations. Hence, prompts that help students connect concepts to physical representations are also successful prompts for virtual representations.

Second, patterns with highest support and confidence for both representation modes involved relations to concepts and/or representations, indicating that students co-construct concept-representational competencies incrementally, over the course of consecutive social exchanges.

Third, the instructor plays a prominent role in prompting concept-representation connections both for physical and virtual representations: instructor utterances were involved in 4 of 5 patterns for virtual representations and in 7 of 16 patterns for physical representations. This result suggests that the role of an instructor is critical to students’ success in making concept-representation connections, regardless of representation mode.

Fourth, assent that co-occurs with a reference to concepts or representations plays an important role for both representation modes. Hence, encouraging students to elaborate by agreeing with prior utterances may prompt concept-representation connections.

Several differences between representation modes stand out. First, students made fewer concept-representation connections with virtual representations (3.28%; see Table 2) than with physical representations (7.33%). Given the finding that instructors play a critical role for concept-representation connections, it may be that the lower involvement of an instructor when students work with virtual representations accounts for this difference.

Second, when students work with physical representations, assent seems to prompt concept-representation connections only when it is provided by the instructor. By contrast, when students work with virtual representations, assent provided by the student partner also prompts concept-representation connections. Hence, this type of prompt may be one that students can take responsibility for when working collaboratively without instructor support.

Finally, meta-cognitive utterances of confusion or understanding of concepts or representations were important prompts only for virtual representations. Given that none of the patterns that included meta-cognitive utterances included instructor utterances, it seems that meta-cognitive utterances are a major mechanism by which students can prompt concept-representation connections in the absence of instructor support.

4. DISCUSSION
My goal was to investigate the representation dilemma: how novice students make connections between new concepts and new representations. I investigated which social events in collaborative classroom practices prompt students’ concept-representation connections. Using frequent pattern mining, I identified such prompts for physical and virtual representations.

A key finding was that prompts with the highest confidence and support contained relations to a previously mentioned concept or representation, regardless of representation mode. This finding suggests that the conceptual process by which students make concept-representation connections is mediated by a gradual, incremental social mechanism. Students may first discuss a concept or a representation separately from one another before they negotiate the connection between the two.

A further finding was that instructors played a crucial role in prompting concept-representation connections, regardless of the representation mode. With respect to physical representations, this finding is not surprising because students have no other way of receiving feedback and assistance. However, with respect to virtual representations, this finding is surprising because the representations were embedded in an educational technology that supported concept-representation connections (and was shown to be successful in doing so [14]). Hence technology support for concept-representation connections may not be able to “replace” instructor support—at least when students have little prior knowledge about the concepts or representations.

Finally, the results showed that meta-cognitive statements can prompt concept-representation connections when students work on virtual representations. Meta-cognitive statements were the only successful prompts when an instructor was not involved. The social mechanism underlying this effect may be that a meta-cognitive statement by one student prompts the other to explain the given concept-representation connection.

5. LIMITATIONS & FUTURE RESEARCH
Several limitations of the present analysis should be considered when interpreting these results. First, the study used a multiple-case design, which focuses on gaining in-depth insights into social processes that unfold over time rather than on generating generalizable evidence for causal effects. Therefore, this paper does not attempt to make causal claims about which prompts are effective, but to generate new hypotheses about social prompts. Based on
the theoretical consideration that instructional support for concept-representation connections may be most effective if it takes advantage of social prompts that different representation modes afford, one may hypothesize that instructional interventions should be designed to maximize instructors’ capacity to assist students, regardless of the representation mode. One might also hypothesize that interventions with virtual representations are particularly effective if students are prompted (or trained) in monitoring their own understanding and communicate their metacognitive assessments to their partner. These hypotheses should be tested with study designs that allow for causal claims.

Another limitation regarding the generalizability stems from the focus on the representation dilemma; that is, how novice students see novel concepts in novel representations. Because students in this study had limited prior knowledge about concepts and representations, we do not know if the results generalize to advanced students. One may speculate that the importance of instructor support decreases as students learn, especially if students receive technology support. One might also speculate that the incremental way in which students focus on a concept or a representation alone before connecting them plays a lesser role if students have prior experience with representations or concepts. Hence, future research should examine social prompts among advanced students. A related limitation is that many utterances did not involve concept-representation connections. Consequently, the overall support and confidence for the discovered patterns is rather low. Concept-representation connections are one of many mechanisms of students’ learning, so future research may apply the present analysis to other social (or conceptual) mechanisms of learning.

A further limitation results from this study’s focus on social mechanisms that may underlie the complementary effects of representation modes on conceptual learning. Consequently, this study did not consider prompts beyond collaborative discourse, such as availability of resources in the classroom, an individual’s bodily experiences with physical representations, etc. Future research could examine the role of such distributed and embodied types of prompts for concept-representation connections.

Finally, an assumption of this study was that concept-representation connections are a “desirable” educational outcome. While much research documents the importance of connecting concepts to representations for students’ learning [1-12], this study did not test whether concept-representation connections correlate with learning outcomes. Future research could assess learning outcomes and test whether concept-representation connections mediate the effectiveness of physical and virtual representations and of interventions that combine both modes.

6. CONCLUSIONS
This study yields new theoretical insights into the representation dilemma by revealing how novice students connect new concepts to new representations. This study identified social events that prompt students to connect concepts to physical and virtual representations. These connections emerge in a co-constructive process that is incremental and requires instructor support. Meta-cognitive statements prompt students to help one another to make connections when an instructor is not always available.

At a practical level, this study yields new hypotheses suggesting that physical and virtual representations are most effective if instructor support is available. If instructor support is not available, interventions with virtual representations may benefit from meta-cognitive support. These hypotheses are empirically testable in studies on combinations of physical and virtual representations.

7. ACKNOWLEDGMENTS
We thank participating teachers and students, Sally Wu, Jamie Schubeth, Ashley Hong, Amber Kim, and Tae Ho Lee.

8. REFERENCES