J Sci Educ Technol (2009) 18:321–333
DOI 10.1007/s10956-009-9159-1

Initial Structuring of Online Discussions to Improve Learning and Argumentation: Incorporating Students' Own Explanations as Seed Comments Versus an Augmented-Preset Approach to Seeding Discussions

Douglas B. Clark · Cynthia M. D'Angelo · Muhsin Menekse

Published online: 26 March 2009

Abstract Collaboration scripts can facilitate argumentation in online settings by grouping students with other students who have expressed differing perspectives on a discussion topic. This general scripting approach is referred to as a "conflict schema." Prior studies suggest that a specific conflict schema script known as *personally-seeded* discussion is more productive for students than a standard discussion format in terms of the structural quality of the resulting argumentation and participation within the discussions. The purpose of the current study involves comparing the personally-seeded script with a variant augmented-preset script to determine the relative contributions of components of the scripts in terms of (1) increasing personal engagement of the students versus optimizing of the starting seed-comments and (2) grouping students using the conflict schema approach versus random assignment of students to groups. The results suggest that engaging students in the exploration of a diverse set of preset discussion seed-comments coupled with a conflict schema approach leads to the highest gains in learning.

Keywords Argumentation · Online learning environments · Conflict schema · Science inquiry

Any opinions, findings, and conclusions or recommendations expressed in this study are those of the researchers and do not necessarily reflect the views of the National Science Foundation.

D. B. Clark (✉)
Department of Teaching and Learning, Vanderbilt University, 230 Appleton Place, Box 230 GPC, Nashville, TN 37203, USA
e-mail: clark.learning@gmail.com

C. M. D'Angelo · M. Menekse
Arizona State University, Tempe, AZ, USA

Introduction

Online learning environments designed to engage and support students in dialogic argumentation provide excellent opportunities for students to propose, support, evaluate, critique, and refine ideas. Over the last decade, a number of sophisticated environments have been developed to support students engaging in this type of knowledge-building discourse. Examples, among others, include *CONNECT* (e.g., de Vries et al. 2002), TC3 (e.g., Erkens et al. 2003), *DUNES* (e.g., Schwarz and Glassner in press), Virtual Collaborative Research Institute (e.g., Janssen et al. 2006), *ArgueGraph* (e.g., Jermann and Dillenbourg 2003), and the *personally-seeded discussions* within the Web-based Inquiry Science Environment (e.g., Clark 2004; Clark and Sampson 2005, 2007, 2008; Cuthbert et al. 2002).

The design of many of these environments can be thought of in terms of "scripts" that orchestrate and control students' interactions with each other and the environments (e.g., Hesse 2007; King 2007; Kollar et al. 2007; Stahl 2007; Weinberger et al. 2007). As Weinberger and colleagues explain, "collaboration scripts provide more or less explicit and detailed instructions for small groups of learners on what activities need to be executed, when they need to be executed, and by whom they need to be executed in order to foster individual knowledge acquisition" (Weinberger et al. 2007, p. 195).

One particular class of scripts focuses on grouping students together with other students who have expressed differing perspectives or stances. This general scripting approach can be referred to as a "conflict schema" (Dillenbourg and Jermann 2007, p 292). *ArgueGraph* (e.g., Jermann and Dillenbourg 2003), and personally-seeded discussions represent two different examples of this class of scripts that have proven successful in supporting

argumentation (e.g., Clark 2004; Clark and Sampson 2005, 2007, 2008; Cuthbert et al. 2002; Dillenbourg and Jermann 2007; Jermann and Dillenbourg 2003). While the success of the general conflict schema approach has been demonstrated, less is known about optimal specific components for conflict schema scripts. The purpose of the current study involves comparing the personally-seeded script with a variant augmented-preset script to determine the relative contributions of components of the scripts to student learning within the discussions, both when a conflict schema approach is used to organize the discussion groups as well as when students are randomly assigned to discussion groups.

The Original Personally-Seeded Discussion Script

The first component of the original personally-seeded script engages students in constructing an explanation for the phenomenon to be debated in the subsequent online discussion. In order to help students focus on the salient issues and articulate their ideas clearly, students use an interface of pull-down menus to construct their explanation from sentence fragments (see Fig. 1). The predefined phrases and elements in the current study include components of inaccurate ideas that students typically use to describe heat, thermal equilibrium, and thermal conductivity that were identified through the misconceptions and conceptual change literature (Albert 1978; Clough and Driver 1985; Erickson and Tiberghien 1985; Erickson 1979; Harrison et al. 1999; Hewson and Hamlyn 1984; Jones et al. 2000; Lewis and Linn 1994; Rogan 1988; Shayer and Wylam 1981; Slone et al. 1996) and an earlier thermodynamics curriculum development project (Clark 2000, 2001; Lewis 1996; Linn and Hsi 2000). In summary, this first component of the script is intended to (a) focus students on the salient issues and (b) highlight distinctions between ideas for students.

Once the students have submitted their explanations, the second component of the personally-seeded script organizes students into discussion groups with other students who have created explanations different from one another. Discussion groups consist of three to five students. Organizing students with different perspectives together is intended as a "pedagogical strategy that will both initiate and support argumentation" (Osborne et al. 2004, p. 997). This is the core mechanism of the conflict schema approach and increases students' exposure to alternative interpretations of the phenomenon under discussion.

The third component of the personally-seeded script includes the students' own specific explanations as the initial seed comments in the ensuing online discussions. The goal of including students' own explanations as the seed-comments focuses on increasing personal relevance and investment in the discussions. Students then participate in an asynchronous online discussion of their explanations where they are encouraged to propose, support, critique, evaluate, and revise ideas (see Fig. 2).

The design rationale for the original personally-seeded script is therefore more elaborate than a less scripted discussion plan that simply directs students to evaluate one another's ideas. By asking the students to construct preliminary explanations before joining the discussion, the script attempts to familiarize students with the ideas and distinctions at the heart of the upcoming debate. By grouping students with students who created different explanations, the script attempts to increase diversity of perspectives in the discussion. By including students' own explanations as the seed-comments, the script attempts to increase personal relevance and students' ownership over their ideas.

Thus, this activity structure is intended to go beyond typical small group work. Rather then viewing small group work as an opportunity to divide up the labor in order to finish a task quicker or as an opportunity to rely on a more

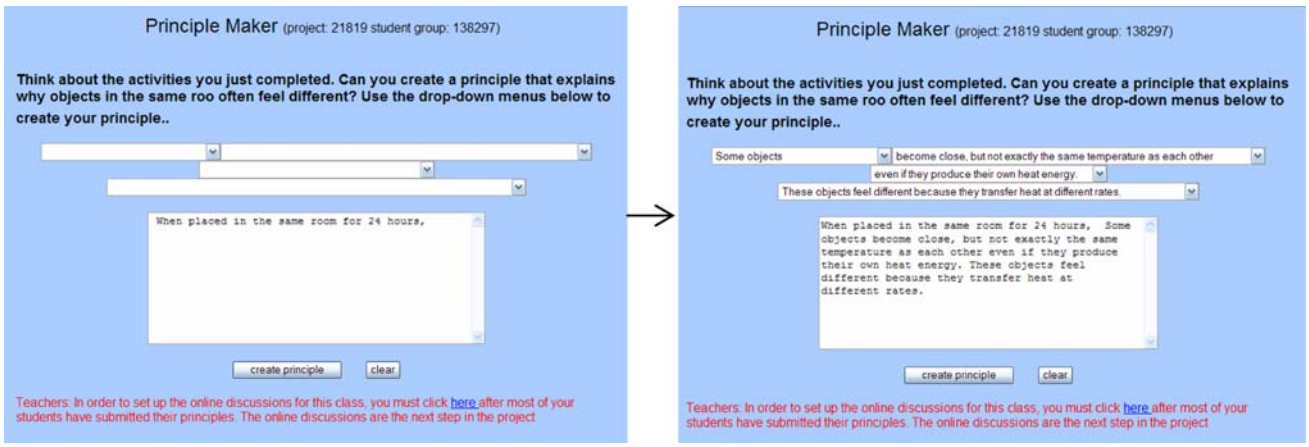

Fig. 1 The explanation construction interface. Students use a pull-down menu to construct an explanation from four sentence fragments that include common misconceptions

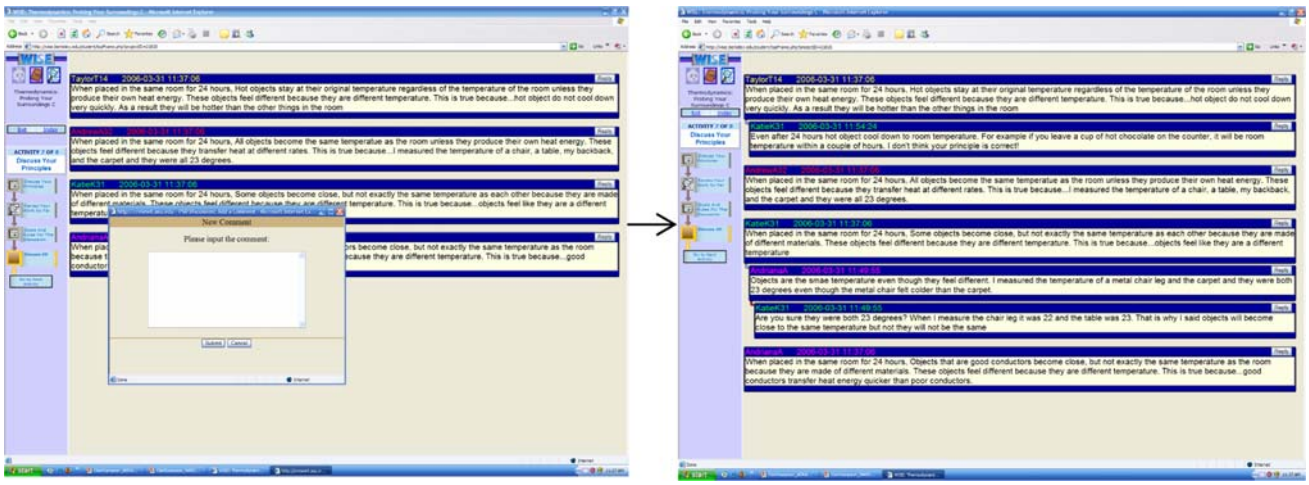

Fig. 2 Threaded discussions with initial seed comments on the *left* (personally-seeded condition) and the early stages of a discussion on the *right*

knowledgeable peer (Cohen 1994; Linn and Burbules 1993), students must engage in genuine collaboration and consensus building. Pilot work that measured the structural quality of argumentation and participation in the ensuing discussion showed that this original personally-seeded script is superior to standard online discussions that involved (1) no pre-exploration of the explanation fragments that constitute the preset explanations, (2) random group assignment, and (3) preset explanations as seeds (Clark 2004; Clark and Sampson 2005; Cuthbert et al. 2002).

Purpose of the Current Study

As discussed above, the personally-seeded script includes three main components: (1) initial scaffolding and exploration of the sentence fragments from which the seed comments will be constructed, (2) sorting the students into groups with students who have constructed different explanations, and (3) including students' own explanations as the seed comments for the ensuing discussion. The purpose of this study involves comparing the second and third components of this script with potential alternatives to better understand the relative contributions of each component to student learning and participation in the subsequent discussions.

The first trial of the current study compares the seed-comment selection component of the original personally-seeded script with a variant augmented-preset script. Essentially, the augmented-preset groups received a predetermined set of seed-comments constructed by the researchers using the fragments from the interface shown in Fig. 1, while the students in personally-seeded groups received the explanations they constructed with the interface shown in Fig. 1 as their seed comments. Table 1 shows two sets of seed comments. The first set is the set received by the four-person augmented preset groups. The second set is an example from a four-person group in the personally-seeded condition. The table also includes scoring information that will be explained later in this document. Both conditions, however, include (1) the initial scaffolding in terms of allowing students to explore the fragments that will constitute the initial seed comments using the interface depicted in Fig. 1 prior to the discussions, and (2) the conflict schema approach of grouping students together for the discussions with students who have created different explanations.

The two conditions diverge solely in terms of the third component (i.e., the nature of the initial seed comments). The augmented-preset condition includes the preset list of initial seed comments for the discussions tailored from the original sentence fragments that the students explored during the first component of the script. The preset seed comments were constructed to represent an optimized range of student misconceptions instead of including students' own explanations as the seed comments. Table 2 includes all of the fragments available through the interface along with scoring information. The second trial repeats this comparison in terms of seed-comment selection, but involves random assignment of group members to discussion groups rather than implementing the "conflict schema" approach of grouping students with opposing perspectives.

Through these comparisons, this study clarifies the relative contributions of the second and third components of the original collaboration script in terms of (1) increasing personal engagement of the students versus optimizing of the starting seed-comments and (2) grouping students using the conflict schema approach versus random assignment of students to groups.

Table 1 Example sets of seed comments from discussion groups

	Score
Preselected set of seed comments from the augmented-preset groups ($M = 7.50$, SD = 3.70)	
When placed in the same room for 24 hours, all objects become the same temperature as the room unless they produce their own heat energy. These objects feel different because they transfer heat at different rates	12
When placed in the same room for 24 hours, all objects become the same temperature as the room but only on their surface not inside them. These objects feel different because they transfer heat at different rates	9
When placed in the same room for 24 hours, objects that are good insulators stay at their original temperature regardless of the temperature of the room unless air can get inside them. These objects feel different because they are different temperature	5
When placed in the same room for 24 hours, some objects become close, but not exactly the same temperature as each other because they are made of different materials. These objects feel different because they are different temperature	4
Example set of seed comments from a personally-seeded group ($M = 8.00$, SD = 2.16)	
When placed in the same room for 24 hours, objects that are good conductors become the same temperature as the room unless they produce their own heat energy. These objects feel different because they are different temperature	10
When placed in the same room for 24 hours, hot objects become the same temperature as the room even if they produce their own heat energy. These objects feel different because they transfer heat at different rates	9
When placed in the same room for 24 hours, metal and glass objects become close, but not exactly the same temperature as the room unless they produce their own heat energy. These objects feel different because they transfer heat at different rates	8
When placed in the same room for 24 hours, all objects are at a different temperature than other objects in the same room because they are made of different materials. These objects feel different because they transfer heat at different rates	5

Methods

As discussed above, the first trial in this study compares the personally-seeded and augmented-preset conditions when implementing the "conflict schema" approach of grouping students together for the discussions with students who have created different explanations. The second trial in this study compares the personally-seeded and augmented-preset conditions when assigning students randomly to discussion groups.

Methods: Participants

The current study was conducted with nine ninth-grade integrated science classes (five classes for Trial 1 in the first year and four classes for Trial 2 in the second year) taught by the same teacher at a public high school in a large metropolitan area in the southwestern United States. The classes in Trial 1 and Trial 2 were the same course at the same school with the same teacher and were therefore generally analogous, although the teacher may have adjusted his curriculum somewhat in the second year as outlined in the "Results and Discussion" section. The teacher is an experienced teacher, but he had not worked with the online environment employed in this study or our research group before the study began. The classes are typical ninth-grade integrated science classes, labeled neither "honors" nor "remedial." Prior to this study, the students had conducted various typical high school inquiry projects, but had not explicitly studied dialogic argumentation within the curriculum of the class. The students in each trial worked on this project for approximately six class periods. The public school is located in a diverse city and has a roughly even distribution of boys and girls. The district is 58% Non-Hispanic White, 29% Hispanic, 6% Black, 6% Asian/Pacific Islander, and 1.4% American Indian/Alaska Native. The district categorizes 27% of the student population as economically disadvantaged.

As described above, Trial 1 analyzes five ninth-grade science classes during the first year. In total, there were 147 students, 38 discussion groups, and 2,157 discussion comments. After students were assigned to discussion groups using the conflict-schema mechanism, the discussion groups were randomly assigned to conditions within each classroom (19 personally-seeded groups and 19 augmented-preset groups). Most groups had four students, but some had five or three. Trial 2 analyzes four ninth-grade science classes during the following year. In total, there were 111 students, 28 discussion groups, and 1,430 comments. After students were randomly assigned to discussion groups, the discussion groups were randomly assigned to conditions within each classroom (14 personally-seeded groups and 14 augmented-preset groups). Most groups had four students, but some had three.

Methods: Online Inquiry Project Context

We embedded our discussion scripts for the current study within an online thermodynamics project in the web-based inquiry science environment (see WISE at http://wise.berkeley.edu). The *Thermodynamics: Probing your Surroundings* project consists of eight activities (see Fig. 3). During the first five activities, students make predictions and collect real time data about the temperatures of objects

Table 2 Scoring of explanation fragments

Discussion explanation	Score
Fragment #1	
All objects	2
Some objects	0
Hot objects	1
Metal and glass objects	1
Wood objects	1
Cold objects	1
Objects that are good conductors	1
Objects that are good insulators	1
Fragment #2[a]	
Stay at their original temperature regardless of the temperature of the room	0
Become the same temperature as each other but not the room	0
Become close, but not exactly the same temperature as each other	0
Become the same temperature as the room	2
Become close, but not exactly the same temperature as the room	1
Are at a different temperature than other objects in the same room	0
Fragment #3	
Even if they produce their own heat energy	0
Unless they produce their own heat energy	2
Because they are made of different materials	1
Unless air can get inside them	1
But only on their surface not inside them	1
Fragment #4	
These objects feel the same as each other because they are the same temperature	1
These objects feel different because they are different temperature	1
These objects feel different because they transfer heat at different rates	2
The objects feel different because they transfer heat at the same rate	0
The objects feel the same because they transfer heat at the same rate	1(2)[b]

[a] The score for the second fragment is multiplied by a factor of three in the overall weighted score

[b] If the student chooses the fourth, fifth, seventh, or eighth phrase for the first fragment (the materials to which the explanation will apply), then the student receives a score of "2" instead of "1" for fifth phrase of the fourth fragment. We make this modification because these combinations make the fifth phrase of the fourth fragment a normative selection. For the first trial, this applied to three students in the pre-explanation and two students in the post-explanation. For the second trial, this did not apply to any students

found inside the classroom and explore interactive simulations dealing with such ideas as heat transfer, thermal conductivity, and thermal sensation. Students record the data they gather and describe the observations they make using the WISE note feature. Together, these five activities are intended to provide students with the empirical data and other scientific ideas necessary to stimulate productive argumentation during the subsequent discussions. The sixth activity engages students in the first component of the script in terms of the initial scaffolding of exploring the sentence fragments for constructing explanations. The seventh activity involves the actual online discussions at the heart of the current study. The eighth activity engages students in constructing a revised explanation after the discussions and reflecting on how their ideas have changed.

Methods: Analysis

The personally-seeded script and the augmented-preset script for assigning seed comments were compared in each trial. Whereas earlier research on the personally-seeded script focused on the structural quality of argumentation and participation in the ensuing discussion, this study focuses both on student learning as well as on the structural quality of argumentation and participation in the ensuing discussion. Analyses of student learning focused on the actual explanations the students constructed in the sixth activity prior to the discussions (pre-explanation) and in the eighth activity following the discussions (post-explanation). Students also completed a pre-test and post-test before and after the full project, but this analysis does not

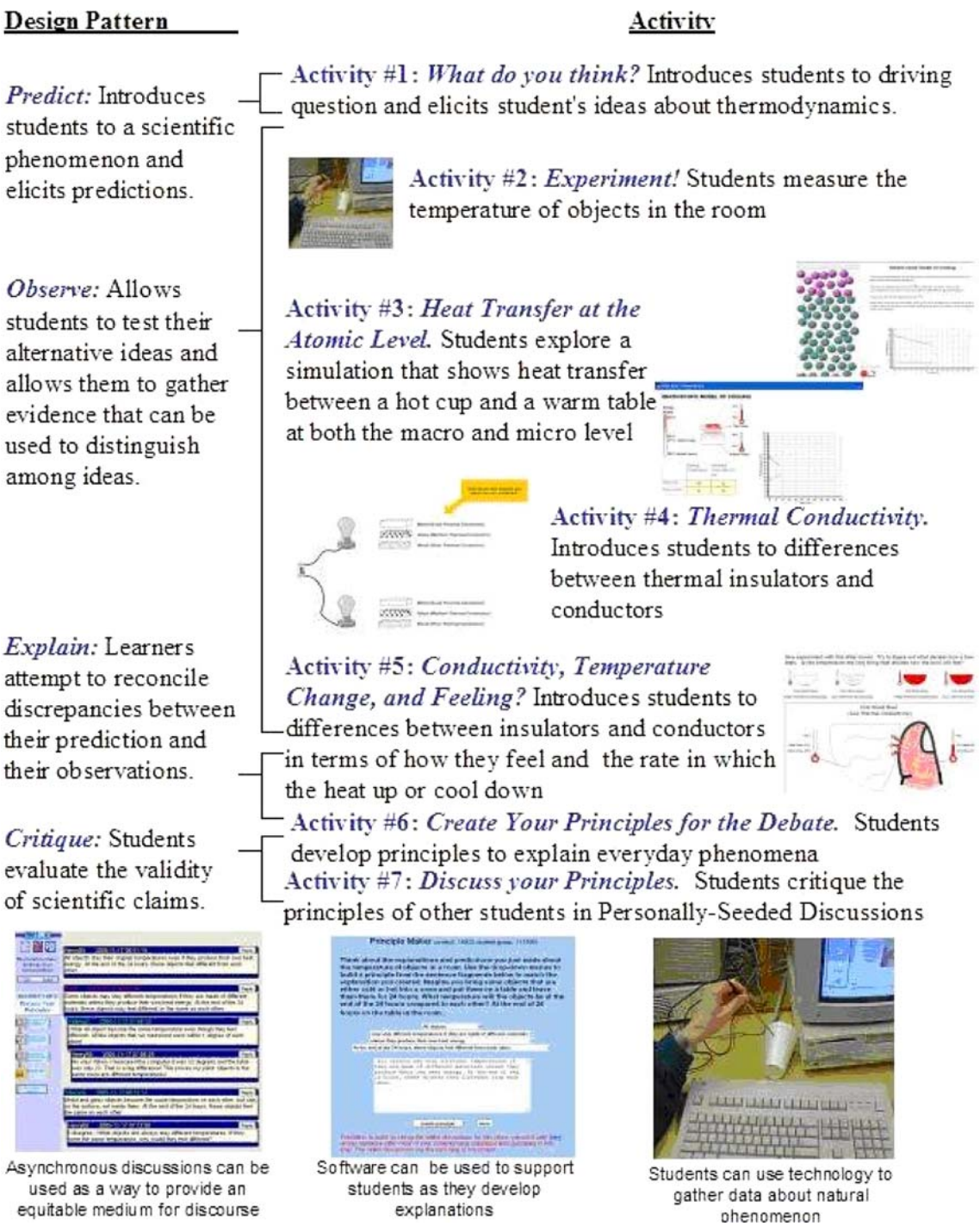

Fig. 3 The *Thermodynamics: Probing Your Surrounding* project

focus on these tests because they encompass the overall project rather than the discussions specifically.

Analyses of the pre-explanations and post-explanations were conducted in terms of raw gains and normalized gains by individual and by group. Differences between conditions were calculated using standard ANOVA tests. Specific analyses are discussed in greater detail in the "Results and Discussion" section. The rubric for scoring the students' constructed explanations is outlined in Table 2 from the previous section. Each explanation consists of four separate fragments. Students choose one phrase for each fragment to construct their explanation. Each fragment phrase choice has a score describing its scientific accuracy: 0 points (non-normative), 1 point

(transitional), or 2 points (normative). For example, the second fragment has a normative option of "become the same temperature as the room," a transitional option of "become close, but not exactly the same temperature as the room," and many non-normative options, including "stay their original temperature regardless of the temperature of the room" (see Table 2 for full list of fragments and scoring and Table 1 for example seed comments constructed from these fragments). The weighted overall score for a student multiplies the second fragment, which is the most critical from the perspective of thermal equilibrium (the science concept at the heart of the curriculum), by a factor of three. Weighted scores therefore range from "0" to "12" for each student.

In addition to our analysis of the pre-explanations and post-explanations, we also analyzed argumentation quality and participation in the ensuing discussions in terms of discourse move types, conceptual quality, and grounds quality of the 2,157 individual comments contributed by students in the first trial as well as the overall structural level of argumentation evidenced in their discussions. The methods for this analysis are explained in abbreviated form in the "Results and Discussion" section. The full description of this analytical approach is detailed in Clark and Sampson (2008).

Results and Discussions

We present the results in terms of (1) the gains on the post-explanations for each trial and a comparison of the gains across trials, (2) an examination of the relationship between seed comment diversity and post-explanation gains, and (3) an analysis of argumentation structural quality, comment types, conceptual quality, grounds quality, and participation levels within the actual discussions.

Gains on Post-Explanation Scores

All students created an explanation before and after their discussions about the thermodynamics concepts from the activities. The explanations that students created from the pull-down menus were scored using the point values listed in Table 2 to determine the conceptual quality, with a total possible weighted score of 12. Students who did not complete all four fragments of their explanation or who typed in their own responses were not included in the analysis (17 students in Trial 1 and 6 students in Trial 2 were thrown out).

Trial 1 Results

When analyzing the explanation scores, we looked only at the groups that had at least three members with complete pre- and post-explanations. This left us with 19 personally-seeded groups and 15 augmented-preset groups. Looking across conditions, there was no significant difference between pre-explanation scores ($F = 0.19$, $p = .66$), so we can say that the two conditions were equivalent on this measure at the beginning of the study. As reported in Table 3 above, there was an overall gain for groups in both conditions, but only the augmented-preset groups exhibited a significant difference.

Table 3 Pre- and post-explanation scores

	Trial 1		Trial 2	
	Pre-explanation	Post-explanation	Pre-explanation	Post-explanation
Augmented-preset	6.63	7.73	7.76	8.24
Personally-seeded	6.90	6.94	7.40	7.52
All	6.79	7.28	7.58	7.88

Trial 2 Results

For Trial 2 we analyzed the explanations from individuals in 13 personally-seeded groups and 13 augmented-preset groups (one group in each condition did not have three students with complete explanations). Across conditions, there was no significant difference between the pre-explanations ($F = 0.29$, $p = .59$). There was a slight gain for the groups in the personally-seeded condition and a larger (but not statistically significant) gain for the groups in the augmented-preset condition (see Table 3).

Comparing Trials 1 and 2: The impact of the Conflict Schema

The pre-explanation scores for the two trials differ somewhat. The average pre-explanation score in Trial 1 was 6.79 and the average pre-explanation score in Trial 2 was 7.58. All classes in both trials were the same ninth-grade integrated science course taught by the same teacher at the same school. It therefore seems likely that the teacher augmented his curriculum leading up to the project in year 2. We compared the two trials and conditions in terms of gain scores. Overall, students in Trial 1 (conflict schema) achieved higher standard gain scores than the students in Trial 2 (random assignment to groups), gains of 0.49 for Trial 1 and 0.30 for Trial 2. When comparing within conditions between trials, however, the result is more dramatic. For the augmented-preset condition, Trial 1 students had an average gain of 1.10 while Trial 2 students had an average gain of 0.47. This magnitude of difference

is not seen with the personally-seeded condition, however, students in Trial 1 had an average gain of 0.04 while Trial 2 students had an average gain of 0.12. Normalized gain scores follow this same pattern with .21 for augmented-preset groups in Trial 1 and .10 in Trial 2, while normalized gain scores for the personally-seeded groups went from .04 in Trail 1 to .07 in Trial 2. Therefore, the augmented-preset condition benefitted substantially from the conflict schema approach. One possible explanation for these findings could be the range of ideas presented to the students in each condition, which we will explore in the next section.

Diversity of Ideas

The results show that the students in the augmented-preset condition had higher gains on their explanations than the students in the personally-seeded condition. Why did this happen? One hypothesis is that the explanations created by the personally-seeded groups (which became the seed comments for their discussions) did not have the same diversity of ideas as the preset seed comments (which were chosen to represent an optimized range of student conceptions). For Trial 1, the computer sorted students into groups to maximize differences within each group (the conflict schema), but it is possible that there was not enough diversity in the class's explanations to match the diversity in the preset seed comments.

To test this hypothesis, we compared the standard deviation of the pre-selected seed comments in the augmented-preset condition with the standard deviation of the seed comments in the personally-seed condition (which were the students' pre-explanations) in terms of each group's normalized gain from the pre- to post-explanations. Table 4 reports the average standard deviation and normalized gain for groups in each condition overall for both trials. In Trial 1, the augmented-preset group has a much higher normalized gain than the personally-seeded group. In Trial 2, we do not see such a dramatic difference between the two groups in terms of normalized gain.

In order to determine what relationship the average standard deviation has with the normalized gain, we also sub-divided the personally-seeded groups by thirds into three categories in terms of low, medium, and high standard deviations in their seed comment (pre-explanation) scores. Table 1 provides an example of a group's set of seed-comments with a high standard deviation in scores (scores of 12, 9, 5, and 4 with a standard deviation of 3.70) and an example of a group's set of seed comments with a low standard deviation in scores (scores of 10, 9, 8, and 5 with a standard deviation of 2.16). In general, many groups with a high standard deviation of seed-comment scores had a couple very low scores and a couple very high scores.

As shown in Table 4, the personally-seeded groups in Trial 1 with high standard deviations in their groups' seed-comment scores have a larger normalized gain than the other groups and are on par with the augmented-preset groups. This supports our hypothesis that a larger diversity of initial seed comments leads to higher gains on the post explanations. In Trial 2, we see a similar but less pronounced pattern. The personally-seeded groups with high standard deviations and middle standard deviations in terms of their seed comments (pre-explanations) had higher normalized gains than the low standard deviation groups and the augmented-preset groups.

Table 4 Standard deviation and normalized gain of groups

	Trial 1		Trial 2	
	Average SD	Normalized gain	Average SD	Normalized gain
Augmented-preset	3.75	.21	3.78	.10
Personally-seeded	3.30	.04	2.91	.07
Personally-seeded: low SD	2.45	−.07	1.80	−.30
Personally-seeded: mid SD	3.27	.00	2.83	.20
Personally-seeded: high SD	4.20	.18	4.11	.15

Group Members' Improvement

These higher gains in the high standard deviation groups might result from a single student in each group with a very low pre-explanation score improving significantly. Therefore, in addition to incorporating normalized gains into our analysis, we also investigated how many members in each group improved in their explanation scores from pre to post (Fig. 4).

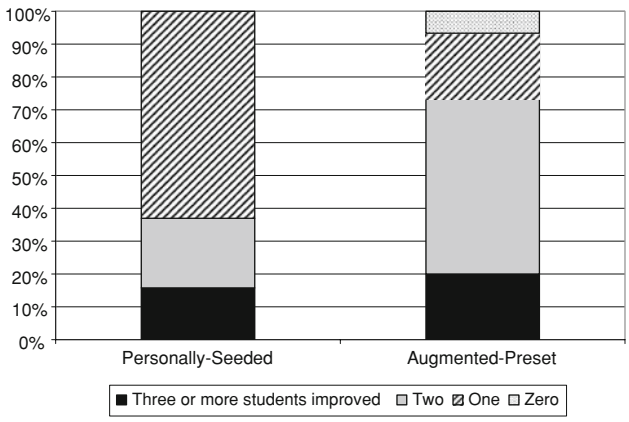

Fig. 4 Percentage of groups with members who improved their score in Trial 1

Trial 1 Results

Not only did the augmented-preset groups on average have a higher gain than the personally-seeded groups, but they also had more group members who improved their scores. Specifically, 73.3% of the augmented-preset groups had two or more people with a gain from pre- to post-explanations, while only 36.8% of the personally-seeded groups had two or more people with a gain from pre- to post-explanations. Furthermore, 53.8% of the students in the augmented-preset condition improved their score on the post-explanation compared to only 42.9% of the students in the personally-seeded condition.

Trial 2 Results

We also performed these analyses for Trial 2. As with Trial 1, the augmented-preset condition also had more groups with two or more students who exhibited gains in their explanations (46.2% of groups in the augmented preset groups versus 38.5% for the personally-seeded groups, see Fig. 5). Overall, 37.3% of the students in the augmented-preset condition and 32.0% of the students in the personally-seeded condition had gained from pre- to post-explanation in Trial 2. These differences are not as large as the differences seen in Trial 1, and that could be connected to the fact that overall there were higher scores and lower gains in Trial 2.

Analysis of Discussions

After analyzing students' explanations, we next analyzed the nature of students' participation in the actual discussions to provide some formative data for future studies. This analysis focused on the discourse moves, grounds quality, and conceptual normativity of the 2,157 student comments in the discussions from Trial 1. We employed the analytic methods detailed in Clark and Sampson (2008) for these analyses.

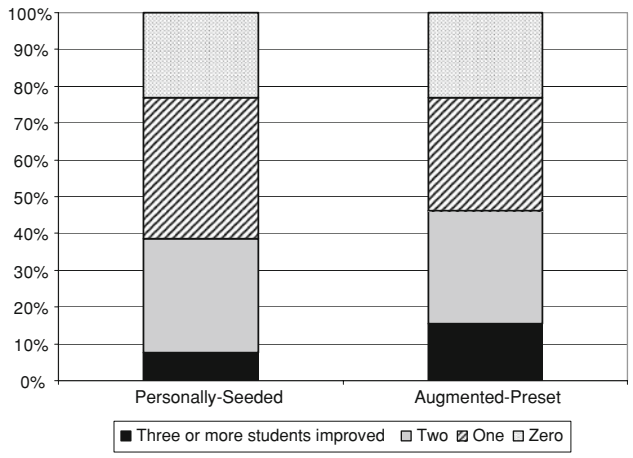

Fig. 5 Percentage of groups with members who improved their score in Trial 2

An independent samples *t*-test shows that the mean conceptual normativity level per episode of the comments in the augmented-preset condition ($M = 1.38$, SD $= 1.04$) is significantly higher than the mean in the personally-seeded condition ($M = 1.21$, SD $= 0.77$), $t(422) = 1.94$, $p < .05$. In examining the individual comments, we hypothesize that this may have resulted from the fact that students in the augmented-preset condition were guaranteed to have a fully normative explanation as one of the seed comments in their group while students in the personally-seeded condition have their own explanations as seed comments for their group and therefore may or may not have had a fully normative explanation included depending on their group. Having fully normative ideas introduced into discussions may have led to a higher frequency of normative ideas propagating throughout those discussions.

Table 5 provides an example of how normative ideas may have been introduced into the discussions by the preset seed comments. Table 5 shows a sequence of three comments added by a student early in his discussion. The other comments have been removed from the example to focus on the students' initial contributions. The student initially responds to a non-normative preset seed comment with a transitional comment. The student disagrees with the non-normative preset seed comment, but the student does not explicitly challenge critical non-normative facets of the preset seed comment. The student next responds to a normative preset seed-comment. The student's response adopts some of the normative ideas from the seed comment and the response is normative. The student next responds to a non-normative preset seed comment using some of the normative ideas appropriated from the previous normative preset seed-comment. Essentially, the student disagrees with non-normative preset seed-comment by providing an explanation to support the student's normative argument. In this example, it therefore appears that the provided fully normative preset seed-comment may have contributed to the higher mean conceptual quality of comments in the augmented-preset groups by introducing normative ideas to the discussion that were adopted by the student.

A few other differences between the conditions in terms of discussion quality also suggest advantages in terms of the augmented-preset approach. These differences were not statistically significant, but followed trends from earlier studies and thus invite speculation in terms of our formative analyses. The mean grounds quality level of comments in the augmented-preset condition is higher (but not significantly higher), for example, than the mean in the

Table 5 Example of augmented-preset seed-comments introduce normative ideas

Comment type	Time stamp	Comment	Conceptual quality
Preset seed comment 2		*When placed in the same room for 24 hours, objects that are good insulators stay at their original temperature regardless of the temperature of the room unless air can get inside them. These objects feel different because they are different temperature.*	*Non-normative*
Student *Z*'s response	08:34:09	I believe with you a little but it's not airgoing through the object it is heat energy, and also they can feel different at the same temperature	Transitional
Preset seed comment 3		*When placed in the same room for 24 hours, all objects become the same temperature as the room unless they produce their own heat energy. These objects feel different because they transfer heat at different rates.*	*Normative*
Student *Z*'s response	08:38:26	You are exactly correct because the objects do obtain the temperature of the room and the objects do feel different because they have different rates.	Normative
Preset seed comment 4		*When placed in the same room for 24 hours, some objects become close, but not exactly the same temperature as each other because they are made of different materials. These objects feel different because they are different temperature.*	*Non-Normative*
Student *Z*'s Response	08:41:50	Well when placed in a room for a day the material are going to get the temperature of the room the can still feel the same at the same temperature	Normative

Note Student *Z* adopts the normative ideas in Explanation 3 and carries them into the subsequent post. (The rest of the discussion comments have been removed from this example to highlight Student Z's initial responses to the seed comments)

personally-seeded condition (Fig. 6). If future studies determine that there is indeed a significant difference of this nature, this may be a function of how discussions change when the students become personally involved and connect the ideas to one another. Qualitative analysis of the comments suggests that students in the personally-seeded condition are more likely to make comments such as "I don't agree with your principle" or "I think [this other person's] principle is more accurate" instead of "I disagree because…" or "I think this principle can work" as in the augmented-preset condition. Thus, the students in the augmented-preset condition appear potentially more likely to include grounds for their statements rather than focusing on the connections to the individual participants.

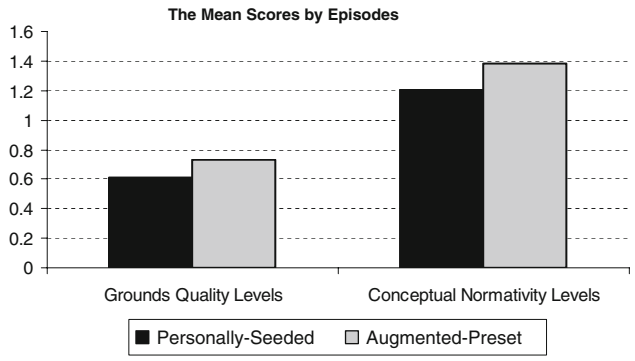

Fig. 6 Comparing the mean scores of grounds quality and conceptual normativity of the student comments by condition per episode

Similarly, the frequency of rebuttals in the augmented-preset condition is higher (although also not significantly) than the frequency of rebuttals in the personally-seeded condition. This may again be another function of the personal connections in the sense that students may be less willing to rebut or contradict an explanation when it is "owned" by another person, in comparison to when the explanation is attributed to a non-present third party. Along the same lines, structural quality was assessed for each episode. The *level of opposition* framework (Clark and Sampson 2008), which focuses on frequency of rebuttals, shows that there is a higher mean structural quality of the episodes in the augmented-preset condition compared to the personally-seeded condition.

Another suggestive, but not statistically significant, difference between the augmented-preset and personally-seeded conditions involves the frequency of comments that are off-task. Approximately 28% of all comments in the personally-seeded condition were coded as *off-task* behaviors compared to 21% of the comments in the augmented-preset condition (Fig. 7). In addition, further

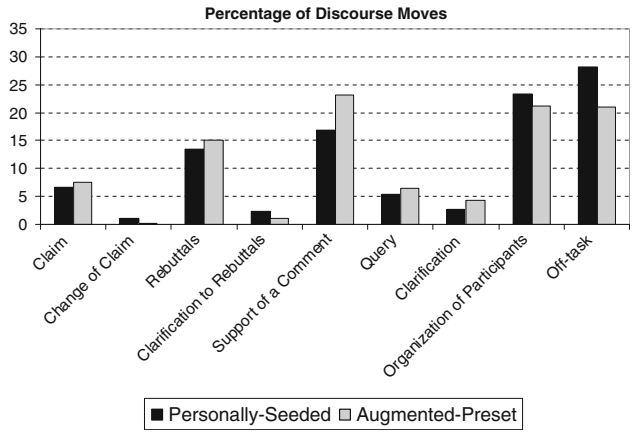

Fig. 7 Comparing the percentage of discourse moves by condition during student discussions about their explanations

analysis of the average number of off-task comments per student in the discussion groups reveals that the number of off-task comments per student in the personally-seeded condition is higher (but not significantly) than the number of off-task comments in the augmented-preset condition.

While not significant in this study, this pattern was observed in a previous study (Clark et al. 2008). One possible reason that the personally-seeded condition has more off-task comments per student is that these students potentially felt more ownership over their own explanations and may have needed to defend or support them more than students in the augmented-preset condition. While both groups had many off-task comments about completely non-related topics (e.g., "nice haircut!"), only the personally-seeded condition included comments that were personally focused in terms of how the group should appraise the explanations. For example, students in the personally-seeded condition said things such as, "we should pick mine" or "don't pick his." Ultimately, our future research should probably subdivide the off-task coding category to distinguish between truly "off-task" comments (e.g., "nice haircut!") and comments aimed at personally based or socially based persuasion (e.g., "we should pick mine") that are not theoretically appropriate from the perspective of scientific argumentation but that are common tools in day-to-day argumentation and persuasion.

We also analyzed the average number of words contributed per student in each discussion group by condition. Before counting the words and characters, we removed the seed claims and off-task comments from the discussion in order to see the actual discussion among students. According to the results, similar to those seen in previous studies, students in the personally-seeded condition contributed higher word totals (although not significantly higher in the current study) than the students in the augmented-preset condition. Although this difference was not as large in the current study, this trend suggests that students in the personally-seeded condition tend to type more than the students in the augmented-preset condition, which might suggest higher levels of engagement. Ultimately, therefore, the personal embeddedness and engagement of the personally-seeded condition appears to offer advantages as well as disadvantages compared to the augmented-preset condition along an "objectivity" versus "engagement" continuum. Overall, however, the quality of the augmented-preset condition appears superior based on these formative analyses of the content of the discussions.

Implications, Limitations, Next Steps, and Conclusions

Overall, this study contributes to our understanding of activity structures for engaging students in argumentation in online environments. This study compares two different scripts (personally-seeded versus augmented-preset) using the basic conflict schema approach as well as the random assignment of group members (non-conflict schema approach).

In terms of the implications for the conflict schema approach, comparisons between the trials in terms of modified gain scores show that students in Trial 1 (conflict schema) outperformed students in Trial 2 (non-conflict schema) overall and particularly in the augmented-preset condition. A future study of ours will conduct this comparison directly. This study, however, confirms other work demonstrating the advantages of the general conflict schema approach (Clark 2004; Clark and Sampson 2005, 2007, 2008; Cuthbert et al. 2002; Dillenbourg and Jermann 2007; Jermann and Dillenbourg 2003).

The major contribution of this study (and the issue that this study was designed to assess) involves clarifying the optimal structuring of the seed comments that form the basis of the online discussions in conflict schema scripts. This study shows that the students in the augmented-preset condition demonstrate significant gains on their explanations in Trial 1 (the conflict schema approach) and gains in Trial 2 (the non-conflict schema approach). Furthermore, the actual discussions of the students in the augmented-preset groups generally demonstrate the same or better overall argumentation quality in terms of structure, discourse moves, and grounds quality. Their participation levels are slightly lower, but the overall outcomes favor the augmented-preset condition in terms of the discussions themselves.

These results in conjunction with the results from our earlier work suggest that much of the added value of the personally-seeded script in comparison to standard online discussions focuses on the first two components of the

script in terms of (1) the initial scaffolding provided through students' exploration of the explanation fragments that encapsulate the key idea facets that will be used in the seed comments and (2) the conflict schema approach of grouping students in discussions with students who have expressed different perspectives.

One hypothesis for the enhanced augmented-preset performance is that the sets of seed comments for the personally-seeded groups (which were their own explanations) did not on average include the same diversity of ideas as the sets of preset seed comments in the augmented-preset groups. The average standard deviation for the sets of seed comment scores in the augmented-preset groups (based on the standard explanation rubric) is higher than the average standard deviation of the seed-comment scores in the personally-seeded groups. This hypothesis is supported by the fact that the augmented-preset groups not only showed a higher average gain and normalized gain than the personally-seeded groups, but they also included a higher proportion of group members who improved their scores than the personally-seeded groups. Furthermore, personally-seeded groups with high standard deviations in their seed comments performed similarly to the augmented-preset groups. The augmented-preset condition potentially results in more productive learning than the personally-seeded condition as a function of the higher standard deviation of pre-explanation scores in the augmented-preset condition, which might expose students to a wider range of ideas. Another possibility is that students in the augmented-preset condition were guaranteed to have a fully normative explanation as one of the seed comments in their group while students in the personally-seeded condition have their own explanations as seed comments for their group and therefore may or may not have had a fully normative explanation included depending on their group.

We are currently designing a follow-up study using a 2×2 experimental design to test these hypotheses directly by sorting students into groups that contrast a high diversity of seed comments in each discussion group with a low diversity of seed comments in each discussion group and contrast the augmented-preset and personally-seeded approaches to seed comments. Essentially, the planned study will expand the current study by including a low diversity set of preset seed comments and a high diversity set of preset comments (instead of simply adding the high diversity set of preset seed comments from the current study). We will also attempt to control for the normativity of the highest scoring seed comments in each group. Through this design, we hope to clarify further the role of each of these mechanisms in the value of the conflict schema scripts for online argumentation. Future research will therefore clarify strategies for integrating these components more effectively and investigate the role of diversity in initial seed comments versus the role of normativity of seed comments on student content learning outcomes.

Future research should also investigate the extension of these approaches beyond online argumentation environments to other online contexts involving discussion groups. It seems likely that many environments that focus on discussion of ideas as a means of supporting student learning could benefit from grouping students based on differing opinions (a conflict schema approach) to ensure that multiple different perspectives are present within the discussions. This could potentially apply beyond text-based environments to virtual worlds. A challenge in massively multiplayer virtual worlds, for example, involves scaffolding player interactions with other players when so many individuals participate simultaneously. A conflict-schema mechanism could provide or contribute to a matchmaking functionality to optimize team formation. A conflict-schema mechanism could thus provide a focal point for conducting research on team formation within multiplayer worlds and research on students' interaction with one another. Ultimately, the pre-focusing approach and the conflict schema approach both suggest high potential applicability to a wide range of other online and offline contexts where groups of students need to think critically about new concepts.

Acknowledgments This study was supported by the National Science Foundation, grant 0334199, as part of the center for Technology Enhanced Learning in Science (TELS).

References

Albert E (1978) Development of the concept of heat in children. Sci Educ 62:389–399. doi:10.1002/sce.3730620316

Clark DB (2000) Scaffolding knowledge integration through curricular depth. Unpublished doctoral dissertation, University of California, Berkeley

Clark DB (2001) New representations of student knowledge integration in CLP: theories or repertoires of ideas? Paper presented at the AERA, Seattle

Clark DB (2004) Hands-on investigation in Internet environments: teaching thermal equilibrium. In: Linn MC, Davis EA, Bell P (eds) Internet environments for science education. Lawrence Erlbaum Associates, Mahwah, pp 175–200

Clark DB, Sampson V (2005) Analyzing the quality of argumentation supported by personally-seeded discussions. Paper presented at the annual meeting of the Computer Supported Collaborative Learning (CSCL) Conference, Taipei

Clark DB, Sampson VD (2007) Personally-seeded discussions to scaffold online argumentation. Int J Sci Educ 29(3):253–277. doi:10.1080/09500690600560944

Clark DB, Sampson V (2008) Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. J Res Sci Teach 45(3):6

Clark DB, Schleigh SP, Menekse M, D'Angelo CM, Sampson V (2008) Improving the quality of student argumentation through the initial structuring of online discussions. Paper presented at

the American Educational Research Association 2008 meeting. New York
Clough EE, Driver R (1985) Secondary students' conceptions of the conduction of heat: bringing together scientific and personal views. Phys Educ 20:176–182. doi:10.1088/0031-9120/20/4/309
Cohen EG (1994) Restructuring the classroom: conditions for productive small groups. Rev Educ Res 64:1–35
Cuthbert AJ, Clark DB, Linn MC (2002) WISE learning communities: design considerations. In: Renninger KA, Shumar W (eds) Building virtual communities: learning and change in cyberspace. Cambridge University Press, Cambridge, pp 215–246
de Vries E, Lund K, Baker M (2002) Computer-mediated epistemic dialogue: explanation and argumentation as vehicles for understanding scientific notions. J Learn Sci 11(1):63–103
Dillenbourg P, Jermann P (2007) Designing integrative scripts. In: Fischer F, Mandl H, Haake J, Kollar I (eds) Scripting computer-supported communication of knowledge—scognitive, computational and educational perspectives. Springer, New York, pp 275–301
Erickson GL (1979) Children's conceptions of heat and temperature. Sci Educ 63(2):221–230
Erickson G, Tiberghien A (1985) Heat and temperature. In: Driver R, Guesne E, Tiberghien A (eds) Children's ideas in science. Open University Press, Philadelphia, pp 52–83
Erkens G, Kanselaar G, Prangsma M, Jaspers J (2003) Computer support for collaborative and argumentative writing. In: Corte ED, Verschaffel L, Entwistle N, Merriënboer JV (eds) Powerful learning environments: unravelling basic components and dimensions. Amsterdam, Pergamon, Elsevier Science, pp 157–176
Harrison AG, Grayson DJ, Treagust DF (1999) Investigating a grade 11 student's evolving conceptions of heat and temperature. J Res Sci Teach 36(1):55–87
Hesse FW (2007) Being told to so something or just being aware of something? An alternative approach to scripting in CSCL. In: Fischer F, Mandl H, Haake J, Kollar I (eds) Scripting computer-supported communication of knowledge—cognitive, computational and educational perspectives. Springer, New York, pp 275–301
Hewson MG, Hamlyn J (1984) The influence of intellectual environment on conceptions of heat. Eur J Sci Educ 6:245–262
Janssen J, Erkens G, Jaspers J, Kanselaar G (2006) Visualizing participation to facilitate argumentation. Proceedings of the 7th International Conference of the Learning Sciences, Bloomington
Jermann P, Dillenbourg P (2003) Elaborating new arguments through a CSCL scenario. In: Andriessen G, Baker M, Suthers D (eds) Arguing to learn: confronting cognitions in computer-supported collaborative learning environments., Computer-supported collaborative learning seriesKluwer, Amsterdam
Jones MG, Carter G, Rua MJ (2000) Exploring the development of conceptual ecologies: communities of concepts related to convection and heat. J Res Sci Teach 37(2):139–159
King A (2007) Scripting collaborative learning processes: a cognitive perspective. In: Fischer F, Mandl H, Haake J, Kollar I (eds) Scripting computer-supported communication of knowledge—cognitive, computational and educational perspectives. Springer, New York, pp 275–301
Kollar C, Fischer G, Fischer F (2007) The interplay of internal and external scripts. In: Fischer F, Mandl H, Haake J, Kollar I (eds) Scripting computer-supported communication of knowledge—cognitive, computational and educational perspectives. Springer, New York, pp 303–324
Lewis EL (1996) Conceptual change among middle school students studying elementary thermodynamics. J Sci Educ Technol 5(1):3–31
Lewis EL, Linn MC (1994) Heat energy and temperature concepts of adolescents, adults, and experts: implications for curricular improvements. J Res Sci Teach 31(6):657–677
Linn MC, Burbules N (1993) Construction of knowledge and group learning. In: Tobin K (ed) The practice of constructivism in science education. Washington, DC, American Association for the Advancement of Science, pp 91–119
Linn MC, Hsi S (2000) Computers, teachers, peers: science learning partners. Lawrence Erlbaum Associates, Mahwah
Osborne J, Erduran S, Simon S (2004) Enhancing the quality of argumentation in science classrooms. J Res Sci Teach 41(10): 994–1020
Rogan JH (1988) The development of a conceptual framework of heat. Sci Educ 72(1):103–113
Schwarz BB, Glassner A (in press) The role of CSCL argumentative environments for broadening and deepening understanding of the space of debate. In: Saljo R (ed) Information technologies and transformation of learning practices. Pergamon, Elsevier Science Ltd., Amsterdam
Shayer M, Wylam H (1981) The development of the concept of heat and temperature in 10–13-year-olds. J Res Sci Teach 18(5):419–434
Slone M, Tredoux C, Bokhorst F (1996) Decalage effects for heating and cooling: a cross-cultural study. J Cross-Cult Psychol 27(1):51–66
Stahl G (2007) Scripting group cognition. In: Fischer F, Mandl H, Haake J, Kollar I (eds) Scripting computer-supported communication of knowledge—cognitive, computational and educational perspectives. Springer, New York, pp 327–335
Weinberger A, Stegmann K, Fischer F, Mandl H (2007) Scripting argumentative knowledge construction in computer-supported learning environments. In: Fischer F, Mandl H, Haake J, Kollar I (eds) Scripting computer-supported communication of knowledge—cognitive, computational and educational perspectives. Springer, New York, pp 191–211