Hands-on Small Group vs. Whole Class Use of an Interactive Simulation: Qualitative Comparisons

A. Lynn Stephens, John J. Clement, University of Massachusetts-Amherst, Amherst, MA 01003
Email: lstephens@educ.umass.edu, clement@educ.umass.edu

Abstract: Assumptions about the superiority of hands-on use of computer simulations over projecting them in whole class have seldom been tested. Contrary to expectations, preliminary pre-post results from two lesson sequences yielded no evidence for an advantage for students in the hands-on condition. We conduct qualitative analyses of one of the lesson sequences, in which a popular simulation was used in eight high school physics class sections, half in whole class discussion and half in small groups. Videotape and activity sheet analyses of such factors as percentage of time spent on conceptual difficulties and amount of support for using key visual features of the simulation yielded no evidence for an advantage for small group students. No small group students in lower level physics sections showed evidence in written or drawn work for having utilized key visual features. A balance of complementary small group and whole class use is recommended.

Introduction

Studies have suggested that students benefit from control of the pace of animations; the speed of a presentation needs to match the speed of comprehension of a topic (e.g., Mayer and Chandler, 2001). Small group work has been prized for allowing such control as well as for providing students opportunities to interact with others, to create metaphors that other students can readily understand, and to enter engaged exploration of the concepts of a lesson. On the other hand, we have at times noticed scenarios such as the following: Above-average high school physics students were working at a computer with an interactive simulation. When one student raised an important conceptual issue concerning a question on the activity sheet, another student suggested that the first was over-interpreting the question (he wasn't). The forceful comment appeared to shut down a potentially fruitful discussion and the conceptual issue was never discussed. Observations such as this one raise the question of whether a teacher might have been able to encourage discussion of the student's question in whole class discussion, and more generally, whether there are complementary advantages to small and large group formats for discussing simulations. Here we report on a set of comparative case studies in the two formats.

Theoretical Background

A number of studies have investigated the effects of instructional guidance for simulations when the guidance was provided within the learning materials themselves (Perkins, et al., 2006; review by Cook, 2006) and the effectiveness of animations or simulations when teachers provided the verbal information (Russell & Kozma, 2005) or when at least part of the use of the simulations was in the context of whole class discussion (Raghavan, Sartoris, & Glaser, 1998). The use of simulations in small groups and by individual students has been studied (e.g., Adams et al., 2008; Linn, 2003; Reid, Zhang, & Chen, 2003; Williams, Linn, Ammon, & Gearhart, 2004). However, there do not appear to be many studies that investigate how best to support students when simulations are used in a whole class setting. Hands-on activity afforded by small group work would appear to offer students a more active learning experience with simulations than would a whole class format. In the context of think-aloud interviews, Adams, et al. (2008), felt that simulations can be highly effective only if the student's interaction is directed by the student's own questioning. This kind of self-directed interaction with a simulation would seem to require a lesson format with hands-on opportunities. On the other hand, studies have reported a variety of issues concerning the effective use of small group discussions in science classes, such as the fact that students can exhibit a low level of engagement with tasks (Bennett, Hogarth, Lubben, Campbell, & Robinson, 2010). Two studies that each compared a single small group class with a single whole class discussion (Wu & Huang, 2007; Smetana & Bell, 2009) did not find significant differences in pre-post gains. These studies, published after we had begun our project, suggest the importance of conducting a larger study.

Summary of Preliminary Results

In an earlier study (Stephens, 2012), a preliminary analysis was conducted of short answer pre-post results from two high school physics lesson sequences that used simulations and animations in hands-on and whole class contexts. We will briefly review these results. The pre-post tests consisted of transfer questions; these targeted the concepts of the lessons via questions that had not been directly addressed during instruction. Tests were administered immediately before and after instruction.

In Tables 1 and 2, College Prep (CP) was the least advanced physics level included in the study, Honors Physics (HP) was an intermediate level, and Advanced Placement (AP) the most advanced. WC and SG refer to lessons that used interactive simulations in Whole Class and Small Group formats, respectively. Given the nature and sizes of the samples, we did not attempt to extrapolate to a larger population or even to compare one matched set with another; however, we found the results intriguing enough to motivate the case study analyses that constitute the main study.

Table 1: Gravitational PE short answer transfer question pre-post gains.

	WC Gains			SG Gains					
	N	Mean	SD	N	Mean	SD	*t*-value	Sig.	Cohen's *d*
CP	11	**0.26**	0.20	14	0.25	0.24	0.097	0.924	0.04
HP	20	**0.22**	0.21	19	0.09	0.15	2.221	0.033**	0.71
AP	23	**0.10**	0.12	21	0.02	0.11	2.368	0.023*	0.71
AP	21	**0.09**	0.16	21	0.07	0.10	0.506	0.616	0.16

Boldface indicates the larger mean gain within each matched set. *Significant difference in gains in favor of the whole class condition. **Significant difference in gains in favor of the whole class condition; however, unanticipated events may have had a disproportionate effect on the small group condition.

Table 2: Projectile Motion short answer transfer question pre-post gains.

	WC Gains			SG Gains					
	N	Mean	SD	N	Mean	SD	*t*-value	Sig.	Cohen's *d*
CP	14	**0.31**	0.29	9	0.27	0.29	0.336	0.741	0.14
HP	21	0.35	0.35	25	**0.36**	0.34	-0.087	0.931	0.03
HP	34	**0.35**	0.29	19	0.32	0.24	0.294	0.770	0.08
HP	15	**0.41**	0.22	22	0.37	0.33	0.471	0.640	0.16
AP	20	0.22	0.22	21	0.22	0.23	0.036	0.971	0.01

Boldface indicates the larger mean gain within each matched set.

To the surprise of the teachers in the study, in no comparison was there a significant advantage for the small group condition. If anything, in the Gravitational PE lesson sequence, there appeared to be a slight trend in favor of the whole class condition. Throughout the two years in which these sequences were conducted, the teachers continued to predict superiority for small group work over whole class work. After the second year, one of the teachers stated in an interview, "When at all possible, most of the time simulations are better done with students working in (small) groups." Thus, the results raised questions for us and for the teachers. The present study uses qualitative analyses of classroom videotapes from the Gravitational PE sequence to look for possible differences between small group and whole class interactions. Analyses of selected answers on student activity sheets shed further light on differences in student learning. (Qualitative analyses of the Projectile Motion lesson sequence will be discussed elsewhere; Stephens & Clement, in preparation.)

Qualitative Study Methods

We use comparative case study analysis of the Gravitational PE lessons, which had been conducted in whole class and small group formats, to shed light on the following overall question: *Why did the whole class format produce gains as strong as those of the hands-on small group format for these classes?*

Research Questions – For Both Conditions

1. To what extent did students and teachers engage in discussion about certain key concepts while working with the simulation?
2. To what extent did teachers and students respond to conceptual difficulties and misconceptions exhibited during work with the simulation?
3. To what extent did teachers and students support the recognition, use, and interpretation of certain key visual features of the simulation?
4. Did students recognize and use key visual features of the simulation?
5. Cutting across the above four themes: Is there a difference in the way the above issues played out in the whole class and small group formats used in these class sections?

Data Sources and Collection

The intention in our classroom observations was not to train teachers to use these tools in a particular way, but to study how these teachers naturally used the tools in two common classroom formats. The Gravitational PE lesson sequence involved two teachers at a high school in a suburban college town. The teachers were purposefully selected; they had to be willing to teach model-based lessons, to foster discussions in both whole class and small group settings, and to use computer simulations as part of their lesson plans. Class sections taught by each teacher were purposefully selected for analysis according to whether they fit criteria for matched sets, as follows. The teacher must have been teaching at least two comparable sections in a given semester and been willing to conduct the lesson sequence in at least one section in a whole class format and in at least one other section in a small group format. Teachers' evaluations and records were relied upon to determine that the sections within a set had students comparable in terms of age and demonstrated levels of aptitude for the content as evidenced by their prior work in the course. In addition, the classes in each section must have provided similar levels of preparedness for the lesson as indicated by the teachers' records of their lesson plans. Finally, the sequence as taught in the two formats must have been similar (see Materials and Procedure below) and the class sections must have been allowed similar amounts of time on the lessons and pre- and post-tests. Fifteen lesson sequences were observed; seven sequences and one teacher were dropped from analysis because they did not meet the above criteria, leaving eight sequences from two teachers to be subjected to comparative analysis. These comprised four matched sets of class sections, N = 150. Once it was determined that sections were matched, they were assigned to the whole class (WC) or small group (SG) condition according to practical logistical considerations. Class times rotated; on some days the teacher taught the whole class condition first and on other days the small group condition first. Teacher A taught this as a two-day sequence while Teacher B taught it as a one-day lesson. Each lesson was videotaped and one or both authors observed all lessons.

Materials and Procedure

Although materials varied slightly for each level of physics, for the two conditions within each matched set, the teacher used the identical simulation and other materials but varied the way in which the simulation was used. In the whole class condition, each teacher used a single computer to project the visuals onto a screen in front of the class and facilitated a whole class discussion as students worked through the activity sheet. In the small group condition, multiple computer stations were used with 2-4 students to a computer; they engaged in hands-on exploration and small group discussion supported by the activity sheet while the teacher circulated among the groups. In both conditions, the teacher introduced the computer activity to the whole class. In both conditions, the teacher was available for questions the entire time the simulation was in use. Other than the constraints provided by the technological set-up, the activity sheets, the simulation mode (whole class or small group) and the data-collection needs of the study, teachers were free to conduct their classes as they saw fit and were encouraged to use the best teaching strategies they could devise. Time-on-task was controlled within each matched set by using the same activity sheets and number of class periods. Though early versions of the materials were inspired by sample lesson plans from the PhET website, the final lesson plans and activity sheets, designed to work with both whole class and small group formats, were largely the construction of the teachers.

The teachers selected a simulation ahead of time from freely available online sources. They chose a sophisticated simulation developed by a research group, *Energy Skate Park* at http://PhET.colorado.edu (Perkins, et al., 2006). See Figure 1. The track can be added to or reshaped, the skater placed anywhere in the scene and released, and the simulation run to see how the skater would respond under the influence of gravity. Activity sheets guided students through an exploration of the skater's motion, the changes in the skater's potential, kinetic, thermal, and total energy with time, and the relationships between those changes. The sheets explicitly instructed students to turn on the Gravitational Potential Energy (GPE) Reference Line (the dotted line in Figure 1) and to move it around. It also instructed students to turn on the animated Energy Bar Graph (to the right in Figure 1), which showed clearly when the potential energy of the skater took on negative values.

Method of Videotape Analysis

Videotape analysis was used to develop a picture of what an individual hypothetical student could have been exposed to in a given lesson. In this analysis the video camera can be viewed as a proxy for an individual student; that is, it took the viewpoint of a hypothetical student in that classroom and recorded what she might

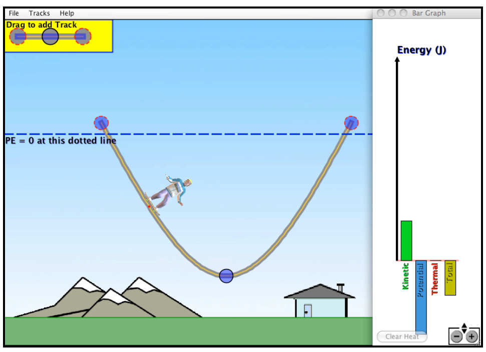

Figure 1. PhET Energy Skate Park with two key features turned on.

have seen and heard. In small group classes, at the point that the students moved into small groups, the camera moved to one of the groups also. Although fewer students were visible on camera than in the whole class condition, the videotape recorded what an individual student in that group could have seen and heard. We began analysis by using a constant comparative method to identify key behaviors observable in videotapes and transcripts of the first four classroom discussions that occurred during use of the PhET Energy Skate Park simulation. Observation categories developed from this procedure were honed in an iterative process along with coding criteria for assigning video segments to categories. This honing process constituted a major part of the effort involved in this study; criteria were developed, applied to fresh transcript sections, critiqued by a second researcher, then refined until the observation categories and their coding criteria stabilized. Finally, the criteria for the stable categories were used to code the entire simulation portion of the lesson in all eight transcripts.

Method of Analysis for Selected Questions on the Activity Sheet

A different lens is provided by activity worksheet analysis. This analysis has the strength of including work from almost all of the students in the classes and is not restricted to students who spoke on camera. However, student drawing and writing abilities varied widely and some activity sheets were difficult to interpret or were not completed. This analysis provided an estimate of how many students actually used certain visual features in their own thinking, as evidenced by their written and drawn answers to selected activity sheet questions. We began analysis by using a constant comparative method to code student responses to selected questions in a stratified sample of 30 activity sheets. Questions were selected that 1) addressed key concepts that the key visual features were thought to support; 2) asked for open-ended written and drawn answers. Coding categories developed from this procedure were honed in an iterative process with critiques from a second researcher. These were used to code student responses to the selected questions on all activity sheets for which there were legible answers (135 out of 150 sheets). Activity sheet coding was done blind to whole class or small group condition.

Results

Videotape Analysis

To what extent did students and teachers engage in discussion about certain key concepts while working with the simulation?

From pilot lessons, a fundamental concept that we observed posing difficulty for the pilot students was the concept of the existence of negative energy quantities, especially negative total energy.

Code: Student or teacher mentions possibility of total energy of some system being zero.
Code: Student or teacher mentions possibility of some kind of energy value being negative.

Percentage of discussion time spent on these concepts was determined. The results provide an estimate of what an individual student in the position of the camera could have been exposed to during the discussion.

Table 3: Gravitational PE lesson pre-post gains: short answer transfer questions.

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 CP	Teacher B	4.32 min / 42.42 min = **0.10**	0.40 min / 23.90 min = 0.02
Yr 1 HP	Teacher A	2.85 min / 62.03 min = **0.05**	0.75 min / 29.23 min = 0.03
Yr 1 AP	Teacher B	0.92 min / 41.10 min = 0.02	0.99 min / 32.32 min = **0.03**
Yr 2 AP	Teacher B	2.58 min / 41.71 min = **0.06**	1.16 min / 28.95 min = 0.04

Results expressed in minutes, *not* in minutes and seconds. Boldface indicates the larger percentage in each matched set. HP=Honors Physics; CP=College Prep; AP=Advanced Placement

The percentage of discussion time spent on these concepts is shown in Table 3. Notably:

- Discussion about these two concepts ranged from 2% to 4% of discussion time in the small groups on camera and from 2% to 10% in the whole class discussions.
- Small groups spent less time on the lesson, not because less time was allowed but because they chose to finish early, thinking they were done with the activity. Therefore, the *total* amount of time spent on these key concepts was substantially less in the small group discussions than in the matched whole class discussions, ranging from less than half a minute to a little over a minute in the small group discussions and from a minute to over four minutes in the whole class discussions.

In each condition there was little discussion time devoted to these concepts. This was surprising, given that the animation provided important potential affordances for developing the concepts, including the two features shown in Figure 1, and the fact that students occasionally expressed frustration concerning these ideas. These are not the only important concepts necessary for students to understand the material; however, they were of particular interest because their lack appeared to constitute a block to acquiring other concepts of the lesson. The evidence described here does not suggest an advantage for the students in the small group condition regarding a chance to address these stumbling blocks. Even if the quality of discussion had been much higher in the small groups than in the whole class discussions, it is doubtful that less than half a minute of discussion, as in the lower level small group, would have been sufficient to explore the concept of zero or negative energy.

To what extent did teachers and students attempt to respond to conceptual difficulties and misconceptions during work with the simulation?

Students sometimes expressed frustration, confusion, or puzzlement in connection with ideas presented within the animation, the activity sheet, or the class discussion, including (but not limited to) the key concepts described above. At other times, students appeared to try to address each other's misconceptions.

Code: Response to conceptual difficulty: Classroom activity following a student expression of conceptual difficulty was considered a response if it bore some relationship to the expressed difficulty.
Code: Response to misconception: Classroom activity was considered a response to a misconception if it appeared to be an attempt by teacher or student to address a misconception.

Total time spent on such discussion was noted. The results provide an estimate of what an individual student in the position of the camera could have been exposed to during the discussion.

Table 4: Response to conceptual difficulties (expressed as percentage of discussion time).

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 CP	Teacher B	6.15 min / 42.42 min = **0.14**	0.89 min / 23.90 min = 0.04
Yr 1 HP	Teacher A	14.05 min / 62.03 min = **0.23**	3.35 min / 29.23 min = 0.11
Yr 1 AP	Teacher B	3.72 min / 41.10 min = **0.09**	1.58 min / 32.32 min = 0.05
Yr 2 AP	Teacher B	1.79 min / 41.71 min = 0.04	3.12 min / 28.95 min = **0.11**

Results expressed in minutes, *not* in minutes and seconds. Boldface indicates the larger percentage in each matched set. HP=Honors Physics; CP=College Prep; AP=Advanced Placement.

The percentage of discussion time spent on responding to conceptual difficulties is shown in Table 4. Notably:

- The whole class discussions spent a *greater percentage* of time responding to student difficulties than the matched small group discussions in 3 out of 4 comparisons.
- The *total* amount of time spent on such discussion was longer in those whole class discussions also. In the medium level (HP) and lower level (CP) classes, the whole class discussions spent 4x and 6x as long, respectively, on addressing student conceptual difficulties as did the matched small groups.

Even for the most capable students observed, the AP classes, there did not appear to be any overall advantage for the small groups in having their conceptual difficulties addressed by discussion.

To what extent did teachers and students support the recognition, use, and interpretation of key visual features of the simulation?

Visual features identified in a pilot study as having played a key role were the movable Gravitational Potential Energy Reference Line and the Animated Energy Bar Graph (Figure 1). The two key concepts, the possibility that energy could take on negative values and the possibility that the total energy of a system could equal zero, could be explored by coordinated use of these two features. However, at times, we observed students experiencing difficulties with those two features—difficulties that appeared to be more perceptual than conceptual, where they misinterpreted the meaning of a feature or failed to find it at all. On the other hand, we observed small group students in some groups helping each other identify and use these interactive features.

Code: Student or teacher supports use and/or interpretation of a key visual feature or relationship in the simulation.*
*Here, by "interpretation of a feature," we mean the interpretation of its meaning, the development of some degree of understanding, as opposed to attaining rote knowledge or the ability to re-create a visual aspect through mimicry.

This was coded when the student or teacher was observed engaged in one or more of the following to indicate or interpret a key visual feature or relationship: 1) **Selectively pointing out** some aspect of the visual feature or relationship; 2) **Giving a hint** to encourage use or interpretation of the feature; 3) **Gesturing in the air or over the display** to indicate this; 4) **Asking a question to prompt its use or interpretation;** 5) **Suggesting a**

manipulation of the simulation to expose it; 6) **Pointing out a limitation** to interpreting its meaning. Individual visual support 'moves' were identified and counted. The results below provide an estimate of what an individual student in the position of the camera could have been exposed to during each discussion.

Table 5: Support for key visual features (expressed as episodes / hour).

Class	Teacher	Whole Class Format	Small Group Format
Yr 1 CP	Teacher B	37 / 42.42 min = **52** per hour	4 / 23.90 min = 10 per hour
Yr 1 HP	Teacher A	26 / 62.03 min = **25** per hour	8 / 29.23 min = 16 per hour
Yr 1 AP	Teacher B	19 / 41.10 min = **28** per hour	10 / 32.32 min = 19 per hour
Yr 2 AP	Teacher B	19 / 41.71 min = **27** per hour	10 / 28.95 min = 21 per hour

Boldface indicates the larger percentage in each matched set. HP=Honors Physics; CP=College Prep; AP=Advanced Placement.

Frequencies of visual support episodes are given in Table 5.

- Rates of visual support episodes ranged from 10 to 21 per hour for the small group discussions and from 25 to 52 per hour for the whole class discussions.
- Total numbers of episodes ranged from 4 to 10 per small group discussion and from 19 to 37 episodes per whole class discussion.
- In no comparison did the small group discussion show an advantage.

Episodes of student-student support were included; it was not required that the person engaging in support be correct, only that the move appeared intended to help other students in addition to the supporter.

Activity Sheet Analysis

Did students recognize and use key visual features of the simulation?

We asked whether the activity sheets would give evidence for student recognition and use of key features. As students tried to describe their understanding of the concepts, they frequently mentioned the features or indicated them in drawings. The written and drawn answers to relevant open-ended questions were coded.

Code: Answer refers to GPE reference line in a way that implies that it is movable (as per a rubric).
Code: Answer contains evidence (as per a rubric) for use of at least one of 3 concepts supported by the Animated Energy Bar Graph having to do with changing energy and negative energy quantities.
Code: Answer contains evidence (as per a rubric) for use of a key relationship supported by coordinated use of the two key features (that TE and/or PE depend on position of the reference line).

Each student's answers were assigned a 1 or 0 for each code and an average was tabulated for the class section. Table 6 and Figure 2 show the percentage of students in each section whose work was assigned a 1 for a given code. In Table 6, whole class discussion data are listed above the matched small group data. Because the same data were scored along all three dimensions, the results are not added across dimensions.

Each group of 3 bars in Figure 2 represents a single class section analyzed along 3 binary dimensions listed in Table 6; each bar represents the percentage of students in that class exhibiting one of the three types of evidence. For instance, the first bar in each set represents the percentage of students who referred to the reference line in a way that implied it was movable. The light bars that would represent the CP and HP small group dimensions are not visible because they are zero for each of the three types of evidence. Notably:

- In each instance where the teacher facilitated whole class discussion about the relevant questions (did not inadvertently skip them; see Table 6), greater percentages of whole class students showed evidence in written and drawn work for having used the GPE reference line, for having used the Energy Bar graph, and for having used a key visual relationship between these two features.
- The only small group students who showed evidence on their activity sheets for having used the key features were Advanced Placement students. No student in the Honors Physics or College Preparatory small groups showed written or drawn evidence for having used either of the key features or the relationship between them.

These results, representing all students who answered the relevant questions on the activity sheets, suggest no advantage for the small group students over the whole class discussion students in making use of concepts supported by the key visual features, as evidenced in their written and drawn responses.

Table 6: Percentage of Students Showing Activity Sheet Evidence for Use of Key Visual Features

Class	Teacher	N	Lesson Format	Evidence for use of GPE ref line	Evidence for use of bar graph	Evidence for use of relationship
Yr 1 CP	Teacher B	11	WC	**0.36**	**0.27**	**0.18**
Yr 1 CP	Teacher B	13	SG	0.00	0.00	0.00
Yr 1 HP	Teacher A	20	WC	**0.10**	**0.05**	**0.05**
Yr 1 HP	Teacher A	18	SG	0.00	0.00	0.00
Yr 1 AP	Teacher B	13	WC	0.15*	0.23*	0.08*
Yr 1 AP	Teacher B	18	SG	**0.33**	**0.44**	**0.22**
Yr 2 AP	Teacher B	21	WC	**0.95**	**1.00**	**0.95**
Yr 2 AP	Teacher B	21	SG	0.81	0.95	0.48

Boldface indicates the larger mean scores within each matched set. HP=Honors Physics; CP=College Prep; AP=Advanced Placement. *Teacher inadvertently skipped the relevant questions on the activity sheet during whole class discussion.

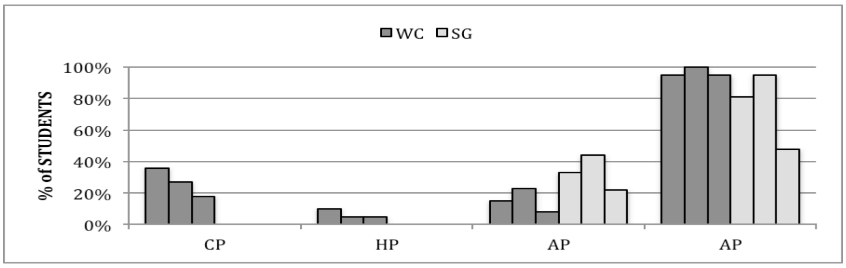

Figure 2. Activity Sheet Evidence for Use of Visual Features (Dark = Whole Class; Light = Small Group)

Discussion

Although almost all classes showed significant gains on the pre-post short answer questions, the teachers were surprised that there appeared to have been no pre-post advantage for students in the small group condition. This was true even though small group participants had had the advantage of hands-on experience with the simulations, opportunity for every student to raise questions with group-mates and with the teacher, opportunity for shyer students to speak up, and the engagement supported by small group work. Why did the small groups not do better than the whole class students? Our qualitative analyses suggest several plausible hypotheses.

First, even though the small groups had the advantage of hands-on experience with the simulations, our videotape analysis showed smaller percentage of *time spent on certain key concepts* than in the whole class discussions in 3 of 4 comparisons. Second, we identified a smaller percentage of *time spent on addressing student difficulties* in 3 of 4 comparisons. At times we observed student conceptual questions being ignored or only very briefly attended to in small groups, possibly due to a goal-oriented "complete the worksheet" mind-set (Conlin, et al., 2007). Third, 4 of 4 whole class discussions had more episodes (and greater frequency of episodes) where a teacher or student provided support for using the visual features of the simulations. As discussed in Stephens (2012), in small groups, even the episodes where students supported each other appeared to cluster around teacher visits to the group. Fourth, activity sheet analyses revealed that in 3 of 4 matched classes, students in the whole class discussions exhibited more evidence for actually having used the visual features in their own thinking. The only small group students who showed any evidence for this were in the highest-level classes in the study; none of the Honors or College prep small group students in these classes exhibited any written or drawn evidence for use of the visual features along any of the three dimensions examined. These results suggest the possibility, consistent with Wu and Huang (2007), that in certain situations there could be a disproportionate advantage for lower level students to participate in whole class discussion.

Conclusions and Implications

Small group students chose, on average, to spend less time on the activity sheets than was spent in whole class. We know this may have contributed to the preliminary quantitative pre post results, but we believe we have also identified other factors that can contribute. In this study, although each pair of classes was matched, the four pairs were not matched with each other in terms of student level and other factors. This means we are limited to within-pair comparisons as opposed to summing over the groups. Despite these small samples, we believe that the four pairwise comparison studies are sufficient to raise questions about the common assumption that small groups are always a better configuration, and suggest hypotheses as to why they may not be.

Our classroom observations suggest that teachers may need more guidance provided along with simulations to help them identify what features and relationships are likely to be overlooked by students;

teachers may also need suggestions for making these features explicit. These results appear to offer encouragement to teachers who do not have the resources to allow their classes to engage regularly in small group work at the computer. The argument here is not that the small group work did not have benefits—it clearly did; small group students had pre-post gains almost as, if not as, large as the whole class students. Rather, we argue that the whole class and small group formats could have had compensating strengths and weaknesses when it came to learning from the sophisticated physics simulation used here. The fact that the students in these whole class discussions matched or exceeded the performance of their small group peers implies that whole class strategies evidently exist that can promote at least some of the active thinking and exploration that has been considered to be the strength of hands-on small group work. We suggest that a mixture of the two formats might be optimal; further investigation is warranted to see which might be best used when.

Most of the teachers in our studies believed that simulations are much more effective when used by small groups; however, preliminary pre-post analysis did not appear to bear this out. In order to investigate what was happening in class discussions in the two formats, we conducted qualitative analyses of matched whole class and small group discussions that accompanied use of an interactive physics simulation. These revealed that in the four whole class discussions, there was generally 1) more time spent on key concepts; 2) more time addressing student conceptual difficulties; 3) more episodes providing support for using key visual features of the simulations; 4) more evidence for student use of the visual features in their writing and drawing. These results are consistent with similar analysis of a different lesson sequence (Stephens & Clement, in preparation). Our results suggest the possibility that there may be certain instructional situations where there is an advantage to spending at least part of the time with an interactive simulation in a whole class discussion mode.

References

Adams, W. K., Reid, S., LeMaster, R., McKagan, S. B., Perkins, K. K., Dubson, M., & Wieman, C. E. (2008). A study of educational simulations Part I - Engagement and learning. *Journal of Interactive Learning Research, 19*(3), 397-419.

Bennett, J., Hogarth, S., Lubben, F., Campbell, B., & Robinson, A. (2010). Talking science: The research evidence on the use of small group discussions in science teaching. *International Journal of Science Education, 32* (1), 69-95.

Conlin, L., Gupta, A., Scherr, R., & Hammer, D. (2007). The dynamics of students' behaviors and reasoning during collaborative physics tutorial sessions. *AIP Conference Proceedings*, *951*(1), 69-72.

Cook, M. (2006). Visual Representations in Science Education: The Influence of Prior Knowledge and Cognitive Load Theory on Instructional Design Principles. *Science Education 90,* 1073-1091.

Linn, M. (2003). Technology and science education: Starting points, research programs, and trends. *International Journal of Science Education, 25*(6), 727–758.

Mayer, R. E., and Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*(2), 390-397.

Perkins, K., Adams, W., Dubson, M., Finkelstein, N., Reid, S., Wieman, C., & LeMaster, R. (2006). PhET: Interactive simulations for teaching and learning physics. *The Physics Teacher, 44*, 18-23.

Raghavan, K., Sartoris, M., & Glaser, R. (1998). Why does it go up? The impact of the MARS curriculum as revealed through changes in student explanations of a helium balloon. *Journal of Research in Science Teaching, 35*(5), 547-567.

Reid, D. J., Zhang, J., & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning, 19*, 9– 20.

Russell, J., & Kozma, R. (2005). Assessing learning from the use of multimedia chemical visualization software. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 229-332). Dordrecht, The Netherlands: Springer.

Stephens, L. (2012). *Student recognition of visual affordances: Supporting use of physics simulations in whole class and small group settings*. (Doctoral Dissertation). University of Massachusetts, Amherst. *Open Access Dissertations*. http://scholarworks.umass.edu/open_access_dissertations/668.

Stephens, L. & Clement, J. (in preparation).

Williams, M., Linn, C., Ammon, P., & Gearhart, M. (2004). Learning to teach inquiry science in a technology-based environment: A case study. *Journal of Science Education and Technology, 13*(2), 189-206.

Wu, H.-K., & Huang, Y.-L. (2007). Ninth grade student engagement in teacher-centered and student-centered technology-enhanced learning environments. Science Education, 91, 727-749.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants DRL-1222709 and DRL-0723709 awarded to John J. Clement. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.