Does the Teaching/Learning Interview Provide an Accurate Snapshot of Classroom Learning?

Jacquelyn J. Chini[1], Adrian Carmichael[1], N. Sanjay Rebello[1] and
Sadhana Puntambekar[2]

[1]Department of Physics, 116 Cardwell Hall, Kansas State University, Manhattan, KS 66506-2601
[2]Dept. of Educational Psychology, 693 Educational Sciences, University of Wisconsin, Madison, WI 53706-1796

Abstract. The teaching/learning interview has been used to investigate student learning. The aim of the teaching/learning interview is to model a natural learning environment while allowing more direct access to a student's or group's thinking and reasoning. The interview typically involves one to four students working with a researcher/interviewer while being audio and video recorded. It has previously been reported [1] that the data collected in a teaching/learning interview is richer in detail than data collected in an actual classroom. We investigated the possibility that there were also other differences between these formats. We used the same instructional materials as well as pre-, mid- and post-tests in a teaching/learning interview and in a classroom laboratory setting. We will describe how the data collected in these two settings compare.

INTRODUCTION

The teaching/learning interview has been used to investigate student learning. Information about students' behavior during the teaching/learning interview is used to inform decisions about curricula that will eventually be used in classrooms. Thus, it is sensible to question whether students act the same in the teaching/learning interview environment as they would in a normal classroom.

In this mixed methods study, we examined how the data collected from students completing the CoMPASS (Concept Map Project-based Activity Scaffolding System) pulley curriculum in a teaching/learning interview setting compare with data collected from students completing the same curriculum in a laboratory class setting.

METHODOLOGY

The Teaching/Learning Interview

The teaching/learning interview is adapted from the teaching experiment [2, 3]. The teaching experiment involves teaching episodes, individual or group interviews and modeling of students' responses. The teaching experiment differs from the traditional clinical interview because it is acceptable for the teaching experiment to change students' thinking. Similarly, in the teaching/learning interview the researcher aims to understand how interventions affect students' thinking. Thus, it is acceptable and even expected that students' thinking will change during the teaching/learning interview.

While the teaching/learning interview is intended to model a natural learning environment, it differs from the natural setting in several ways. Our teaching/learning interviews are typically conducted in a room that more closely resembles an office than a classroom. Unlike a usual class, the interview is video and audio recorded and the camera and microphones used are visible to the students. Students sign an informed consent form before they begin the interview. The interview typically involves one to four students working with a facilitator/interviewer, which is a far smaller "class size" than a typical classroom. The interviewer typically asks the students to explain their ideas about the interventions and will ask follow-up questions to get a clearer picture of the students' thinking processes. Thus, the students themselves may be considering their ideas and thinking processes more carefully than they typically would.

This study diverged from our typical teaching/learning interview format to try to control for some of these differences. The interview still took place in our interview room, was audio and video recorded and involved only one or two students. However, the interviewer/researcher intervention was altered to mimic a more typical laboratory teaching assistant role. The students worked through the intervention worksheet, and the interviewer only intervened when the students asked for assistance.

For this study, students completed the CoMPASS pulley curriculum in about two hours and were compensated $25 for their participation. Participants were recruited from a conceptual-based introductory physics class on the basis that they were not enrolled in the laboratory section of the course, which was to be used for the classroom study. Some students were also recruited from an algebra-based physics course. Pulleys were not explicitly covered outside this activity in any of the classes involved.

The Classroom Setting

Our classroom study took place in the regular laboratory sections of a conceptual-based introductory physics class. The pulley curriculum was completed in a two-hour lab section. Students received a class grade that was based on both participation and correctness on the post-test. The regular teaching assistant and one researcher were present as facilitators. The comparison between the interview and classroom setting are summarized in Table 1.

TABLE 1. Teaching/Learning Interview vs. Classroom

Teaching Interview	Classroom Study
N= 12	N=132
Paid $25 for participation	Part of normal laboratory
Two hour intervention	Two hour intervention
Interview Room	Laboratory
Alone or with partner	Groups of 3 or 4 students
Researcher facilitates	Researcher & TA facilitate
Audio& video recorded	No audio/video recording
Worksheets collected	Worksheets collected

The Curriculum

In both the teaching/learning interview and classroom study students completed the CoMPASS pulley curriculum. The CoMPASS (Concept Map Project-based Activity Scaffolding System) curriculum combines design-based and project-based activities with an interactive hypertext system [4]. Thus, the designers have merged the benefits of hands-on activities with the CoMPASS hypertext system, which combines interactive concept maps and text for students to navigate through the information. The

hands-on activities included manipulating physical pulleys as well as a pulley simulation. Students were guided to make predictions about successful pulley setups, to use the hypertext system to learn about pulleys, to take and record data from both the physical and virtual pulleys and to answer analysis questions. Additionally, students took a pre-test before starting the curriculum, a mid-test after completing either the physical or virtual pulley experiment and a post-test after completing the entire curriculum.

RESULTS & DISCUSSION

Quantitative Analysis

The median score for the teaching interview and classroom group for the pre-, mid- and post-test is displayed in Fig 1. The Mann-Whitney test was used to compare scores as the data violated the parametric assumptions of normality and homogeneity of variances. There was no statistically significant difference between the pre-test scores for the teaching interview (Median=38) and the class study (Median=31), U=9546.5, p=.15, r=.119. However, the teaching interview group scored significantly higher on the mid-test (Median=62) than the class study group (Median= 50), U=9222.5, p<.001, r=.289. They also scored significantly higher on the post-test (Median= 85) than students in the class study group (Median= 69), U= 9380.0, p=.013, r=.206.
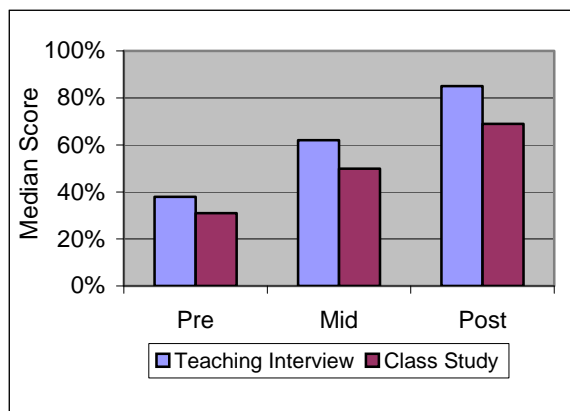

FIGURE 1. Comparison of median test scores of groups.

The percentage of correct responses for each group on each post-test question is shown in Fig. 2. Setting a threshold for significant performance difference around 20% identifies three questions for further study: Question 9, 2.1 and 13. We discuss each separately.
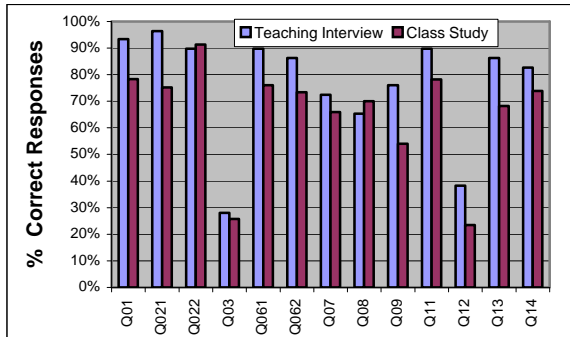
FIGURE 2. Percentage correct responses on each question.

In Q9, which had the largest performance spread, students were asked to identify which of three pictured pulley setups (single fixed, single compound and double compound) would require the most work to lift the same load to the same height. The choices were: A - single fixed, B - single compound, C - double compound and D (correct) - all setups would require the same work. Choice D was the most popular, followed by A (Fig. 3) for both groups. However, a few of the class study students chose B and C as well, while none of the teaching interview students did.
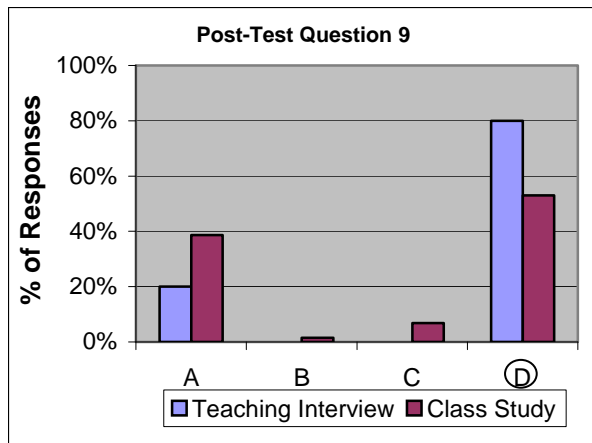

FIGURE 3. Distribution of post-test responses on Q9.

In Q2.1 which had the second largest performance spread, students were asked what would happen to the distance the rope needed to be pulled to lift an object to a certain height if the pulley used was changed from a single fixed to a single movable. The choices were: A (correct) - increase, B - decrease, C - stay the same, and D - not enough information. Figure 4 shows that in the teaching interview as well as the class study, most students selected choice A. However, nearly 20% of students in the class study selected Choice C while no students from the teaching interview chose that option.
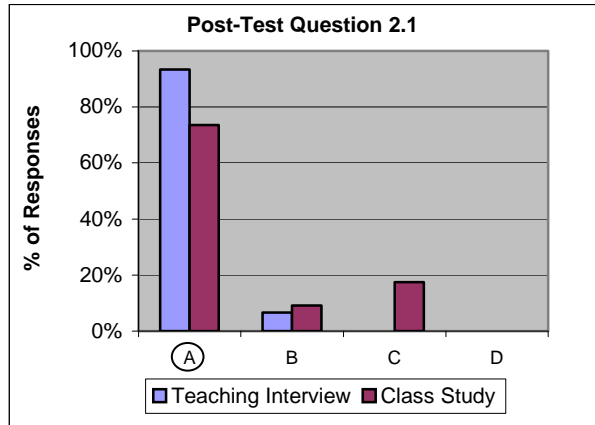

FIGURE 4. Distribution of post-test responses on Q2.1.

In Q13, which had the third largest performance spread, students were asked to compare the work needed to lift an object with its potential energy once lifted. The choices were: A - work is more than potential energy, B - work is less, C (correct) - work and potential energy are the same, and D - not enough information. A majority of students in both groups selected choice C (Fig. 5). A larger percentage of students in the class study chose either B or C.
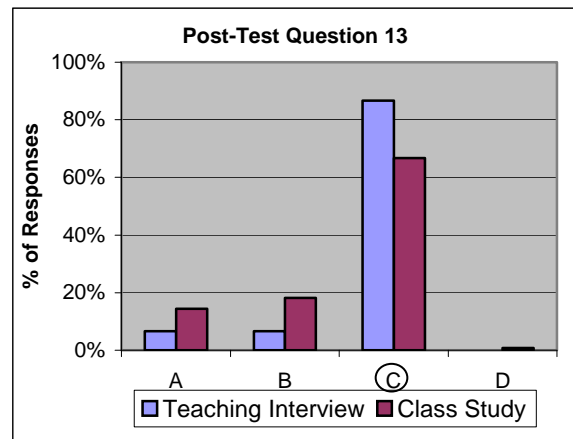

FIGURE 5. Distribution of post-test responses on Q13.

Qualitative Analysis

We used a phenomenographic approach [5] to analyze students' responses to the open-ended worksheet analysis questions in the activity. Codes emerging from the analysis were collapsed into categories. We focus on the analysis questions related to two of the post-test questions discussed above.

In a worksheet question related to post-test Q9, students were asked to explain how the work required to lift an object to a certain height changed when the pulley setup used to lift it was changed. Table 2 shows

the categories of students' responses. A higher percentage of class study (61%) vs. teaching interview (43%) responses was in the 'did not change' category. Conversely, a higher percentage of teaching interview (16%) than the class study (5%) responses were in the 'changed slightly' category.

TABLE 2. Responses to 'work change' question

Categories	Teach. Int.	Class Study
Did not change	43%	61%
Changed slightly	16%	5%
Changed	35%	29%
Changed for some setup	0%	3%
Other	6%	2%

In a worksheet question related to post-test Q13, students were asked to explain how the work done to lift an object compared to its potential energy (PE) once lifted. Table 3 shows that the highest percentage of responses (53%) for both groups was in the 'work = PE', category, which is true in a frictionless case such as simulation. A higher percentage of teaching interview (36%) vs. class study (14%) responses was either in the 'work > PE', 'work < PE' or 'work, PE different' categories. Finally, a higher percentage of class study (15%) vs. teaching interview (3%) responses was in the 'one constant' category, i.e either work or PE were constant. Overall, there was a greater variability in the responses in the class study compared to the teaching interviews.

TABLE 3. Responses to work - potential energy question

Categories	Teach. Int.	Class Study
Work=PE	53%	53%
Work almost PE	6%	9%
Work>PE	21%	7%
Work<PE	6%	3%
Work, PE different	9%	4%
Depends on system	0%	3%
One constant	3%	15%
Other	0%	4%

SUMMARY

The results of the qualitative study of the open-ended analysis questions on the worksheet do not neatly overlap with the quantitative analysis of the multiple choice post-test questions.

The post-test analysis identified several questions on which the students in the teaching interview outperformed the students from the class study. However, a qualitative study of the analysis questions related to two of these post-test questions does not indicate that students in the class study had more difficulty identifying the correct relationship while completing the pulley curriculum. The classroom

study students outperformed or did as well as the teaching interview students on the worksheet questions. These anomalous results indicate that some other factors may be at play.

The main implication of our results is that researchers should be careful before assuming that instructional materials tested in a small scale environment such as a teaching/learning interview will work the same way in a larger classroom.

FUTURE WORK

Further study will be needed to investigate the differences in performance between the two groups. Different incentives were offered to the two groups. The class study students' class scores depended on post-test performance, while the teaching interview students were paid for their participation. Perhaps the teaching interview students took the experience more seriously since they were under more direct observation than the class study students due to a smaller student-researcher ratio. Also, the teaching interview students worked individually or with one partner, while the classroom study students worked with up to three partners.

We plan to repeat the experiment by videotaping some students while they complete the CoMPASS curriculum in a classroom setting. Such a study is likely to provide a more in-depth investigation of the factors at play in a classroom environment. Validity and reliability studies of the tests are currently underway.

ACKNOWLEDGMENTS

REFERENCES

1. D. L. McBride, "Concept Categorization Analysis: Comparing Verbal and Written Data" in *American Association of Physics Teachers Winter Meeting*, Chicago, IL, 2009.
2. L. P. Steffe and P.W. Thompson, in Research design in mathematics and science education., edited by R. Lesh and A. E. Kelly (Erlbaum, Hillsdale, NJ, 2000), pp. 267.
3. P.V. Engelhardt, et al. *The Teaching Experiment - What it is and what it isn't.* in *Physics Education Research Conference, 2003.* 2003. Madison, WI.
4. S. Puntambekar and A. Stylianou, *Instructional Science* **33** 451-481 (2005).
5. F. Marton, *Journal of Thought* **21** 29-39 (1986).