Wong, L.-H. et al. (Eds.) (2013). Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education

Program Visualization: Effect Of Viewing Vs. Responding On Student Learning

Gargi BANERJEE*, Sahana MURTHY & Sridhar IYER
Indian Institute of Technology, IIT Bombay, India
*gargi@it.iitb.ac.in

Abstract: Visualizations in computer science topics are known to have several benefits such as promoting conceptual and procedural understanding, improving prediction and reasoning abilities and helping learners construct mental models. This learning effectiveness has been found to be a function of students' engagement level with visualization. In the current study, we did a controlled field experiment to determine the effect of two different instructional strategies with visualization on procedural understanding of the topic of pointers in a 1st year programming classroom. These instructional strategies, operationalizing different engagement levels, were: prediction activity interleaved with instructor feedback using visualization (experimental), and simply viewing the visualization with parallel instructor commentary (control). We found significant difference in the relative rate of correct solution of the procedural questions on the post-test. However, there was no significant difference on the post-test scores. We also found a significant difference in classroom behavioral engagement between the two groups. We propose that there may be conditions, other than engagement level with visualization, such as learner characteristics or challenge level of assessment questions that may play a role in the determining learning effectiveness of visualizations.

Keywords: Visualization, instructional strategy, engagement levels, behavioral engagement

1. Introduction

Computer-based visualizations such as videos, animations and simulations have been shown to be effective learning resources (Linn & Eylon, 2006). They are known to be useful in making the invisible visible (Rutten et.al. 2012), constructing mental models (Buckley, 2005), and improving prediction and reasoning abilities (Riess & Mischo, 2010). In Computer Science (CS) topics, visualizations have been found to promote conceptual and procedural understanding (Byrne et. al., 1999; Hansen et.al., 2000; Laasko et.al., 2009) and improve vocabulary of programming terms (Ben-Basset Levy et.al, 2003). Visualizations used in CS for teaching and learning fall into two categories - Program visualization (PV) (Ben-Ari et.al, 2011; Urquiza-Fuentes et.al, 2013) and Algorithm visualization (AV) (Byrne et.al., 1999; Hansen et.al., 2000; Grissom et.al., 2003; Laakso et.al., 2009).

Prior research has developed guidelines for design of visualizations to promote learning (Ilomäki et.al, 2009). However, even a well designed visualization can get reduced to a visual textbook if the instructional strategy used is to simply play the visualization in classroom (Lindgren & Schwartz, 2009). Thus, the instructional strategy used with visualization is an important determinant of learning effectiveness of visualizations. The instructional strategies with visualization that have been reported to be successful are: prediction worksheets with visualization (Ben-Bassett Levy et.al, 2003), exercise sheets (Laakso et.al, 2009), integrated prediction activity (Hansen et.al, 2000) and online quiz (Hansen et.al, 2000). Naps et.al. (2002) hypothesized that learning outcome from visualization will increase with increasing level of student engagement with visualization. Numerous studies have been done to test these hypotheses by contrasting learning at multiple levels of student engagement with visualization. We report some of these studies in Section 2 as part of Related Work. (For the rest of this paper we will refer to Naps' engagement levels as 'engagement level with visualization' which is different from students' behavioral engagement).

In CS domain, procedural thinking skills, that is, an understanding of 'how to' carry out stepwise execution of a program for a given set of input data, are crucial in the learning of CS topics (Hundhausen et.al, 2002). Multiple empirical studies have been conducted with visualizations at different engagement levels for improving procedural thinking skills but the results have been mixed.

For example, for the strategy of prediction activity with visualization, Byrne et al. (1999) found that groups that viewed animation and/or made predictions did significantly better in procedural understanding on challenging questions than the No viewing-No prediction group where students were made to do oral prediction on the topic of binomial heap. But, Jarc et.al. (2000) found no difference in procedural understanding between groups that only viewed the visualization versus the group where prediction activity was integrated into the visualization for a set of sorting algorithms. The topic in both these studies was algorithms. For programming topics such as if-while constructs, significant learning gain in terms of procedural understanding was found in a field study with tenth grade students using Jeliot program visualization tool (Ben-Bassett Levy et.al., 2003).

In this paper, we report a study of students' procedural understanding in the programming topic of Pointers, under two different conditions of classroom instructional strategies, operationalizing two different engagement levels with visualizations. The topic of Pointers is often difficult to grasp for beginners because of its abstract nature. We did a controlled field experiment where one instructional strategy was to show the visualization with parallel commentary by the instructor, whereas the other strategy was to make students do prediction activity with instructor feedback interleaved with visualization demonstration. After the treatment, both groups solved the same post-test that tested their procedural understanding of pointers. We also administered a survey to capture student perception of learning from the respective instructional strategies. To triangulate our results, we performed classroom observations for behavioral engagement based on the BOSS protocol (Shapiro, 2003). The sample consisted of 230 students in a first year CS1 programming course.

We found no significant difference in the post-test scores between the two groups. However, the experimental group exhibited a significantly higher rate of correctly solving the post-test problems than the control group. From the student perception survey, we found each group highly favored the respective instructional strategy used with visualization. The classroom observation study revealed that the experimental group was at a significantly higher active engagement level than the control group. To identify possible reasons for these results, we noted that students of both groups had prior exposure to prediction activity with program visualization and were highly trained in procedural thinking. We concluded that there may be conditions such as learner characteristics, topic complexity, and challenge level of questions, along with engagement level with visualization which may play a role in the learning outcome.

2. Theoretical Background and Related Work

In this section we give an overview of the work done to test learning outcome from visualization in response to change in engagement level with visualization, effected by different instructional strategies with visualization. We describe key theories on effect of engagement level with visualization on learning outcome followed by literature survey of positive and negative empirical studies with focus on those that measured procedural understanding in CS topics. We also report studies on the students' behavioral engagement while viewing visualizations. We conclude the section by outlining the need for our work.

From their meta-analysis of learning effectiveness studies for visualization in CS, Hundhausen et.al. (2002) postulated that how students interact with visualization has a significant impact on their learning from visualization. Based on this, Naps et.al (2002) proposed a taxonomy of six engagement levels for algorithm visualizations - No Viewing, Viewing, Responding, Changing, Constructing and Presenting - hypothesizing that learning will increase as the engagement level with visualization proceeds from No Viewing to Presenting, such as, Responding level will lead to better learning outcome with visualization than Viewing. In the 'No Viewing' level, no visualization is involved. In the 'Viewing' level students simply watch the visualization. In the 'Responding' level students not only watch but interact with the visualization by responding to the visual cues presented like answering exercise or prediction questions. In the 'Changing' level, students interact with visualization by changing variable parameters. In the 'Constructing' level, students create their own visualization whereas in 'Presenting' level, they present their created visualizations to the class. The engagement levels with visualization have historically been one of the most explored conditions while measuring learning from visualization. Naps' hypotheses have been tested by multiple studies but the results are mixed.

In one of the successful studies, Grissom et. al. (2003) found learning gain increased with increasing student engagement for simple sorting algorithms (insertion and bubble sort) across no viewing, viewing and responding through online quiz. Similar result was reported by Hansen et.al (2000) where instructional strategy used for responding level was interactive prediction and question-answering. Byrne et.al, (1999) did a controlled experiment with CS majors who were aware of algorithm analysis but had no prior knowledge of the topic binomial heap. These students did better in procedural understanding in post-test when at Responding level (viewing with oral prediction) or Viewing level compared to No viewing – No prediction group. However, the effect of visualization and prediction could not be isolated in this study. Ben-Bassett Levy et.al. (2003) did a field study at school level on programming topics like if – while statements with the post-test containing questions on predicting output of a program code using Jeliot. They found all groups of students – Strong, Average and Weak – showed significant learning gain with average students gaining the most. Laakso (2009) found learning gain for conceptual understanding at both Viewing and Changing levels but gain was statistically significant for Changing level on the topic Binary heap. However, this result was obtained only after correction for behavioral engagement of student pairs since all students did not perform at the expected level of engagement with the visualization. In contrast Myller et.al (2009) found visualization led to increased behavioral engagement in terms of amount of collaborative activity. Thus in the current study we measured the behavioral engagement of both the groups to confirm the attainment of the intended engagement level with visualization.

In contrast to the above studies, there are studies that did not find a difference in learning outcome at different engagement levels with visualization. Jarc et. al. (2000) found no difference in learning outcome (conceptual and procedural understanding) between Viewing and Responding where Responding level was operationalized through automated prediction questions for a set of eleven algorithms. Again, Stasko et. al. (1993) did not get any significant difference in procedural understanding between no viewing group and group that could run the visualization on their input data sets (Active viewing) on the topic of pairing heap. A possible reason cited was the visualization design was not suited to novice learners. Hundhausen & Douglas (2000) did a similar experiment comparing two groups at Constructing and Active viewing levels for procedural understanding. They also did not get any significant difference between the groups on the topic Quick select algorithm. Urquiza-Fuentes et. al. (2013) found no difference in learning outcome between three groups at No viewing, Viewing and Constructing levels when the topic is simple like in-fix operators. But visualization showed an effect when topic was of medium complexity like user-defined data types. However, for medium complex topic there was no significant difference in learning outcome between Viewing and Constructing levels. A significant difference was however obtained in favour of Viewing level rather than Constructing on Analysis and Synthesis level questions when the topic was of high complexity like recursive data types.

The overview above shows that most of the above studies have been for algorithm visualizations, and fewer for program visualizations (this was also noted by Ben-Ari et.al, 2011). The studies related to program visualizations either do not emphasize topics for procedural understanding or do not try to isolate visualization engagement levels through controlled experiments. Analyzing the successful studies for the instructional strategies employed, we found that the successful studies involved active learning techniques like prediction activity and integrated question-answering activity. In the current study, we vary the engagement level between Viewing (watching visualization only) and Responding (prediction activity interleaved with watching visualization) levels using a quasi-experimental research design, for a programming topic involving procedural understanding.

3. Research Questions & Hypotheses

The current study explores three research questions on outcomes with visualization using two different instructional strategies with visualization at Viewing and Responding engagement levels in a large classroom setting. The topic of the study was of medium complexity and students in both groups were exposed to active learning strategies and trained in procedural thinking. Under these conditions, we explored three research questions.

- RQ1: Does prediction activity with visualization (Responding) lead to higher levels of learning outcome than simply viewing the visualization (Viewing) for programming topics?

- RQ 2: Do prediction activity with visualization lead to higher behavioral engagement than just watching visualization for programming topics?
- RQ3: What are student perceptions about learning from visualization through the strategies used?

The substantive hypothesis flowing from RQ1 can be stated as: Prediction activity interleaved with visualization will lead to higher learning outcome than simply viewing the visualization with parallel instructor commentary in a programming class (H1_1). The null hypothesis corresponding to this is, Prediction activity with visualization leads to same learning outcome in a programming topic as watching visualization alone (H0_1). To test this hypothesis, we conducted a post-test measuring students' procedural understanding of the topic of Pointers. We compared the post-test scores of the two groups as also the rate of correct solving of the post-test questions by the two groups.

The substantive hypothesis corresponding to RQ2 is formulated as: Prediction activity with visualization will lead to higher behavioral engagement in a programming class than simply viewing the visualization (H1_2). Thus the null hypothesis here is Prediction activity with visualization will lead to same amount of behavioral engagement as watching visualization alone in a programming lecture (H0_2). To test this hypothesis, we did a classroom observation of student behavior based on the BOSS protocol.

RQ3 leads to the substantive hypothesis that students will perceive the classroom strategy of prediction activity with visualization to be more effective for learning than students who were taught through the strategy of viewing alone (H3_3). The null hypothesis that follows is students who only saw the visualization (Viewing level) will perceive the instructional strategy to be as useful for learning as the students who did prediction activity with visualization (H0_3). To test this hypothesis, we executed a student survey on the instructional strategies used with visualization.

4. Research Methods

A field experiment was conducted using mixed methods research design. The quantitative part involved a 2-group post-test only design to determine students' learning, along with a short perception survey to identify their perceptions of the instructional strategies implemented. The qualitative part consisted of in-class observations using a structured protocol to determine students' classroom behavioral engagement.

4.1 Sample

The sample consisted of 230 first-year undergraduate students from different branches of engineering (Electrical, Mechanical, Aerospace and Chemical) enrolled for an introductory course in computer programming at the Indian Institute of Technology Bombay, India. These students were among the highest ranked in an extremely competitive exam testing analytical skills in mathematics, physics and chemistry. They were trained in procedural thinking through their preparation for the exam and other courses in their first year. These students had been exposed to learning programming topics from visualizations in the programming course. However, only those students who had self-declared no prior knowledge of pointers (the programming topic in this study) were considered for the study.

Students in the programming course were divided among two sections for scheduling reasons. Thus two groups were accessible to the researchers to conduct a controlled study. The first section was the responding group (N=135; Male = 120; Female = 15) and second section was the viewing group (N= 95; Male = 85; Female = 10). Assignment of the treatment to the groups was done on a random basis. The groups were tested for equivalence on basis of quiz marks conducted prior to the study. The marks in each group were found to be normally distributed. We compared the means of the quiz marks for the two groups using independent samples t-test and found them to be equivalent ($M_{experimental}$ = 16.91 (SD = 5.83); $M_{control}$ = 15.72 (SD = 6.09); $p > 0.05$).

4.2 Learning materials used

The topic chosen for the study with visualization was Pointers. Pointers are variables that store computer memory addresses. The topic was deemed suitable for learning with visualization since it involved making the invisible memory address manipulations visible to the students. The visualization chosen was a non-interactive program visualization covering basic pointers and pointer arithmetic (Student project, University of Pittsburg, 2012). The reason for the choice was it satisfied the requirements of visualizations at the responding level as specified by previous research studies (Urquiza-Fuentes et.al., 2009) like presence of explicit feedback and additional narrative or text explanations of what is happening. This visualization displayed the change in memory map in response to execution of each line of code with explicit explanation as also its output, if any (Figure 1).

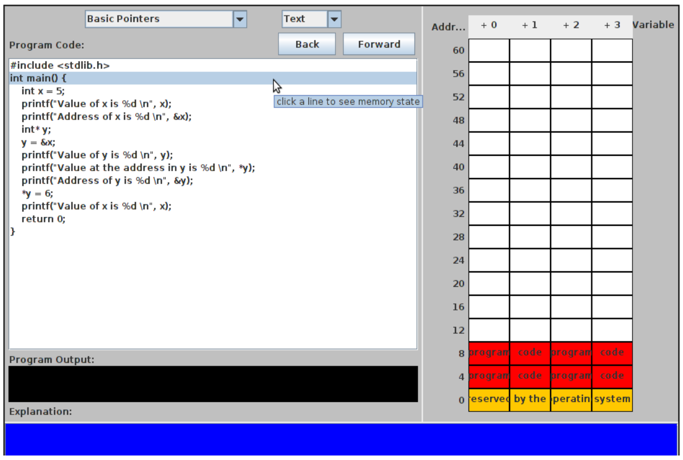

Figure 1: Screenshot of Pointer animation

4.3 Instruments

To measure students' understanding of the topic, achievement scores on a post-test were recorded. The post-test scores were used to investigate RQ1 on students' learning outcome. There were four post-test questions measuring procedural understanding the questions were created by the instructor who was also an educational technology (ET) expert and validated by another ET expert. One sample post-test question is given below.

Predict the output of the following program -

```
int main () {
int A[4], *p;
for (int i = 0; i < 4; i++) A[i] = i;
p = &A[0];
printf ( " %d  %d %d /n" , *p , *(p +=2) ,*(p+1) + *(p-1));
return 0}
```

The full marks of the four questions awarded for correct answers were 5, 2, 3, 3, respectively, making a grand total of 13 marks. Partial marking were done if the question contained multiple procedural understanding testing points.

To answer RQ2 (students' behavioral engagement in the classroom), in-lecture observations of their behavior were done by six researchers. There is considerable debate about the proper definition of the multidimensional construct of student engagement (Parsons & Taylor, 2011). Fredericks, Blumenfeld and Paris (2004) categorized engagement studies into three categories - behavioral, emotional and cognitive. For our study we measured behavioral engagement of students in terms of student participation in classroom. The results of the classroom observations were used to confirm if the intended level of engagement with visualization was indeed attained during the lecture. The observations were based on the standard classroom observation protocol of Behavioral Observation of Students in Schools (BOSS) (Shapiro, 2003). The in-class observations were coded based on BOSS terminology to report three categories of student behavior – Active engagement (AET), Passive engagement (PET) and Non-engagement (NET). For example, behaviors like reading aloud, raising hand or talking about learning material were coded as AET whereas behaviors like listening to lecture/ peer answer or reading silently were coded as PET. Some of the NET codes used were talking at inappropriate times, manipulating non-related objects and looking around the room. The protocol was piloted in this classroom to establish inter-rater agreement which was found to be 90%.

The students were also given a 2-item Likert scale questionnaire to capture their perception of the instructional strategy used with the visualization in their class. The questionnaire asked students whether the instructional strategy used helped them learn and if they would recommend the strategy for rest of the course. The student responses on this 5-point Likert scale survey were used to answer RQ3 on students' perceptions on learning through different instructional strategies with visualization.

4.4 Procedure

The responding group was given short theoretical introduction to pointers and pointer arithmetic subtopics. In each case, the explanation was followed by the visualization. The visualization was run in step-run mode and students were asked to predict and write down the result of the next step before viewing the result in the visualization. They got immediate explicit feedback from the visualization as also explanation provided by the instructor for each step. After this activity, they took the post-test and the perception survey. The viewing group was given a longer verbal introduction for the same two subtopics. The visualization was demonstrated in step-run mode with parallel commentary by the instructor, but students were not explicitly asked to make predictions. Both groups were taught by the same instructor with the same lecture content and same set of visualizations on the same day with responding group going first. The treatment duration was one hour for both. After the treatment, each group took the same post-test that had questions based on the visualization along with the perception survey. Students in each group had to attempt the post-test individually within a time limit of 20mins.

The qualitative part involved in-lecture observation of student behavioral engagement. Each observer observed random sets of 20 students twice during the prediction activity for responding group and the corresponding code segment for viewing group using the BOSS codes. Individual students were observed for a fixed time interval of 5 seconds at a stretch. Student behaviors as defined in BOSS protocol were observed and classified into engagement (active and passive) and non-engagement categories. The total number of students thus observed per group was (20 x 6) = 120 and the total number of observations were 240 per group.

4.5 Data Analysis

To test if the means of post-test scores for the groups were significantly different, we did independent samples t-test using SPSS ver.16. T-test was deemed suitable since the achievement scores adhered to normalization and homogeneity of variance assumption, besides being interval data. The t-test type chosen was independent samples since two groups were mutually exclusive. We also compared the mean rate of correct solutions of the two groups. To calculate the rate of correct solution for each student (R), we divided the number of correct responses of each student in a group (C) by the average time taken by the group to solve the post-test (t) i.e. $R = (C/t)$. Since the distribution of R was found to be non-normal by the Shapiro-Wilk test, the non-parametric Mann-Whitney U test was done to compare the medians of the two groups.

The engagement survey responses were analyzed by the non-parametric test of Mann-Whitney U to check for significant difference between group responses. This test was chosen since the dependent variable (survey responses) is ordinal and independent variable (treatment) is categorical with two levels. Also, both distributions were found to be non-normal from the Shapiro-Wilk test.

The in-class observations of net engaged and non-engaged observations for each group, based on BOSS protocol, were tested for significant difference using Pearson's chi-square test. This test was deemed suitable since both independent (instructional strategy) and dependent variables (engagement) were categorical variables with two levels, the distributions were found to be non-normal from Shapiro-Wilk test and cell count of the 2x2 contingency table was more than 5.

5. Results

5.1 Post-test Results

Post-test scores of both groups exhibit good representation in high scores range (Fig. 2). The mean of total post-test score of the experimental group was 7.91 (SD =3.07) while the mean of the control group was 7.89 (2.92). We found no statistically significant difference between either of the group means in the total post-test score or in question-level means. Both groups at Viewing and Responding engagement levels did equally well in the post-test with $p > 0.10$. However, it was observed that experimental group was able to complete the post-test in half the time (10mins.) taken by control group (20 mins.). We compared the rate of correct solution between the groups ($U = 3.84 \times 10^3$ $p = 0.000$) and found a statistically significant difference in favor of the experimental group. Thus, even though the absolute post-test scores did not exhibit a significant difference, the significantly faster rate of correct solution of the experimental group leads to rejection of the null hypothesis (H0_1). The alternate hypothesis (H1_1) that prediction activity interleaved with visualization leads to better learning outcome than simply viewing the visualization with parallel instructor commentary in a programming class, is accepted.

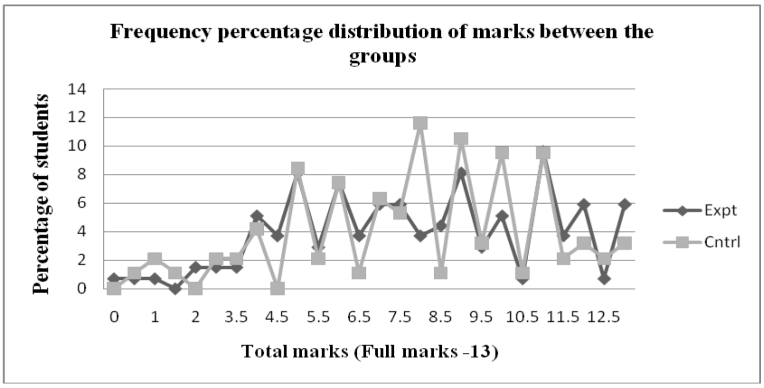

Fig. 2: Frequency percentage distribution of post-test marks across the groups

5.2 In-lecture Observation Results

The 240 observation codes in each group were categorized into engagement (active and passive) and non-engagement. The percentage of observations out of the total that were in each category is reported in Table 1. The chi-square (χ^2) test on the engagement and non-engagement categories of each group revealed a significant difference between the two groups (Table 1).

Table 1: In- lecture student behavioral engagement observation results

Observed Engagement	Observation Frequency (percent)		Chi-square results
	Responding	Viewing	
Non-engagement	26 (10.83%)	47 (19.58%)	χ^2 (1) = 7.13, $p< 0.05$
Engagement (total)	214 (89.17%)	193 (80.41%)	
Active engagement (AET)	56 (23.33%)	23 (9.58%)	χ^2 (1) = 4.42, $p< 0.05$
Passive engagement (PET)	158 (65.83%)	170 (70.83%)	

Both groups showed high behavioral engagement in the classroom with total engagement of responding group (89.13%) being higher than Viewing (80.41%) which was also found to be statistically significant from Pearson's chi-square test (Table 1). We analyzed the total engagement data further into active and passive engagement based on BOSS terminology and found the responding group (23.30%) to be more actively engaged than viewing (9.58%). The chi-square (χ^2) test yielded a significant difference between the groups on active engagement. Thus the null hypothesis, H0_2 is rejected. The prediction activity with visualization led to significantly more active behavioral engagement in classroom than viewing alone.

5.3 Student Perception Survey Results

The responses of both groups to the short 2-question survey on a 5-point Likert scale were analyzed. The survey questions asked were: Q.1) 'Did you learn from the strategy used?' and Q.2) 'Do you recommend the strategy used for rest of the course?' While analyzing the survey responses, the 'Strongly Agree' and 'Agree' responses were clubbed together into a single category of 'Agree'. Likewise 'Strongly Disagree' and 'Disagree' were clubbed into a single category of 'Disagree'. Analysis of the responses showed both groups highly recommended the instructional strategy used with visualization in the lecture. 91.9% of the responding group favored the use of visualization with instructor's parallel commentary and prediction activity and agreed that this strategy helped them learn. For the viewing group, 87.4% favored the use of visualization with instructor's parallel commentary whereas 84.2% agreed that this strategy helped them learn. We did Mann-Whitney U test with these survey responses on a 5-point Likert scale and did not find a statistically significant difference in responses of the two groups on either question (Table 2). So, for students exposed to active learning at tertiary level and being trained in procedural thinking (learner characteristics), highly positive response for visualization alone with instructor commentary as also visualization with commentary and prediction worksheet was obtained. Thus the null hypothesis H0_3 is accepted.

Table 2: Mann-Whitney U test results for student perception survey

Question	Group	Agreed	Disagreed	U	p
Q1. Instructional strategy with visualization helped me learn	Responding (N= 136)	91.9%	2.8%	7255.5	0.245
	Viewing (N= 95)	84.2%	2.1%		
Q2. I would recommend this instructional strategy with visualization for the course	Responding (N= 136)	91.9%	2.2%	7838.5	0.958
	Viewing (N= 95)	87.4%	2.1%		

6. Discussion and Conclusion

6.1 Answering RQ1

The first research question – 'Does prediction activity with visualization (Responding) lead to higher levels of learning outcome than simply viewing the visualization (Viewing) with parallel instructor commentary for programming topics?', was answered by the post-test scores. We found no statistically significant difference in post-test scores either for the total score or at question level between the responding and viewing groups. But the time taken by experimental group to solve the same post-test paper was half (10mins.) that of control group (20 mins). Thus the different instructional strategies with visualization implemented at two different engagement levels, appears to have had an effect on the learning outcome in terms of the relative rate of correct solution between the two groups. The strategies however did not lead to any significant difference in terms of absolute post-test scores. Some of the possible reasons could be:

i. The significantly increased level of behavioral engagement of the experimental group in the class may have led to the significantly higher rate of correct solution exhibited by the experimental group relative to the control group.
ii. Post-test questions were possibly not sufficiently challenging to measure the difference in procedural understanding between the two groups. Byrne et.al, 1999 found significant differences in procedural understanding but only on the challenging questions.

iii. The learner characteristic of both these groups may have played a role. Students of both groups were highly trained in procedural thinking and had prior training in predicting with visualization for learning programming topics. This prior exposure to prediction activity with visualization may have conditioned the viewing group to apply their prediction skills, even when they were not explicitly asked to. The fact that viewing group took double the time to solve the same paper as responding group is possibly because they took the extra time to apply their prediction skills to answer the post-test questions. However, further experiments are required to explore effect of prior training in prediction activity with visualization on learning outcome.

While the specific instructional strategy implemented at two different engagement levels with visualization did have an effect on the learning outcome in terms of rate of correct solution but did not show an effect in terms of post-test scores. Both strategies were found to lead to high learning outcome, as seen from the score distribution across groups as given in Fig. 2. Similar results were reported by Byrne et.al, (1999) in which groups that viewed animation with prediction as well as the group that viewed animation without prediction performed well. (Byrne's study compared these groups with a No viewing-No prediction group and a No Viewing – prediction group while our study did not have such a group).

6.2 Answering RQ2

The second research question, 'Do prediction activity with visualization lead to higher behavioral engagement than just watching visualization for programming topics?' was answered by our in-lecture observations. From Table 1 we see that there is a significant difference in behavioral engagement between the responding and viewing groups. Further, there is a difference in the quality of engagement in the two groups. The Responding group is found to exhibit higher Active Engagement than the Viewing group, whose behavioral engagement is largely passive. Thus the instructional strategy used to implement two different engagement levels with visualization had a significant effect on the behavioral engagement of students in the two groups. The relation between classroom behavioral engagement and learning outcome needs further study.

6.3 Answering RQ3

The third research question, **'**What are student perceptions about learning from visualization with the strategy used?', was answered by the student perception survey. From Table 2, we find both groups highly favored the instructional strategy used with them for learning with visualization and there was no significant difference in response between the groups. A probable reason can be that viewing group gave a positive response to visualization only strategy, since they automatically did prediction on their own due to prior training of prediction with visualization, thus converting the strategy used with them to visualization with prediction.

6.4 Overall Conclusion

We find that there is a significant difference in the relative rate of correct solution on questions of procedural understanding on a programming topic like pointers in a 1st year programming course. However, no significant difference in the procedural understanding was found in terms of the absolute post-test scores. A plausible explanation can be that our student sample was highly trained in procedural understanding. Hence, the control group was able to comprehend and then apply their procedural knowledge in solving the post-test questions. However, it took them more time compared to the experimental group who had already comprehended the procedure in class and only had to do the application in the post-test. We also found significant difference in the classroom behavioral engagement between the groups but non-significant difference in student perception of the respective instructional strategy used with visualization. Our conclusion is that there may be other conditions such as learner characteristics, topic complexity, and challenge level of questions, along with engagement level with visualization that play a role in the learning outcome. Similar conclusions have been stated in other research work (Byrne et.al., 1999; Urquiza-Fuentes et.al., 2013). Our study points to the need for further research which tries to isolate the effects of these conditions, so that instructors can choose the optimal strategies for teaching-learning with visualizations, based on the set of conditions in their context.

References

Ben-Ari, M., Bednarik, R., Ben-Bassat Levy, R., Ebel, G., Moreno, A., Myller, N., & Sutinen, E. (2011). A decade of research and development on program animation: The Jeliot experience. *Journal of Visual Languages & Computing*, *22*(5), 375-384

Ben-Bassat Levy, R., Ben-Ari, M., & Uronen, P. A. (2003). The Jeliot 2000 program animation system. *Computers & Education*, *40*(1), 1-15.

Blake, C. & Scanlon, E. (2007). Reconsidering simulations in science education at a distance: Features of effective use. *Journal of Computer Assisted Learning*, *23(6)*, 491–502.

Byrne, M. D., Catrambone, R., & Stasko, J. T. (1999). Evaluating animations as student aids in learning computer algorithms. *Computers & education*, *33*(4), 253-278.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, *74*(1), 59-109.

Grissom, S., McNally, M. F., & Naps, T. (2003, June). Algorithm visualization in CS education: comparing levels of student engagement. In *Proceedings of the 2003 ACM symposium on Software visualization*, 87-94.

Hansen, S. R., Narayanan, N. H., & Schrimpsher, D. (2000). Helping learners visualize and comprehend algorithms. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, *2*(1), 10.

Hundhausen, C. D., Douglas, S. A., & Stasko, J. T. (2002). A meta-study of algorithm visualization effectiveness. *Journal of Visual Languages & Computing*, *13*(3), 259-290.

Hundhausen, C., & Douglas, S. (2000). Using visualizations to learn algorithms: should students construct their own, or view an expert's? *Proceedings of IEEE International Symposium on Visual Languages, 2000*, 21-28.

Ilomäki, L., Jaakkola, T., Lakkala, M., Nirhamo, L., Nurmi, S., Paavola, S. & Lehtinen, E. (2003). *Principles, models and examples for designing learning objects (LOs). Pedagogical guidelines in CELEBRATE*. A Working Paper for the European Commission, CELEBRATE Project, IST-2001-35188, May 2003.

Jarc, D. J., Feldman, M. B., & Heller, R. S. (2000). Assessing the benefits of interactive prediction using Web-based algorithm animation courseware. In *ACM SIGCSE Bulletin , 32(1),* 377-381.

Laakso, M. J., Myller, N., & Korhonen, A. (2009). Comparing learning performance of students using algorithm visualizations collaboratively on different engagement levels. *Journal of Educational Technology & Society*, *12*(2), 267-282.

Lindgren, R. & Schwartz, D. (2009). Spatial learning and computer simulations in science. *International Journal of Science Education, 31* (3), 419-438

Linn, M. C. & Eylon, B. S. (2006). Science education: Integrating views of learning and instruction. In P. A. Alexander & P.H. Winne (eds), *Handbook of Educational Psychology,* 2nd ed., 511-544. Mahwah, NJ: Erlbaum.

Myller, N., Bednarik, R., Sutinen, E., & Ben-Ari, M. (2009). Extending the engagement taxonomy: Software visualization and collaborative learning. *ACM Transactions on Computing Education (TOCE)*, *9*(1), 7.

Naps, T. L., Rößling, G., Almstrum, V., Dann, W., Fleischer, R., Hundhausen, C. & Velázquez-Iturbide, J. Á. (2002). Exploring the role of visualization and engagement in computer science education. In *ACM SIGCSE Bulletin, 35(2),* 131-152.

Parsons, J., & Taylor, L. (2011). *Student Engagement: What Do We Know and what Should We Do?* University of Alberta.

Riess, W., & Mischo, C. (2010). Promoting systems thinking through biology lessons. *International Journal of Science Education*, *32*(6), 705-725.

Rutten, N., van Joolingen, W. R., & van der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Computers and Education, 58,* 136–153.

Shapiro, E. S. (2003). Behavioral Observation of Students in Schools (BOSS).*Computer Software*. *San Antonio, TX: Psychological Corporation*.

Stasko, J., Badre, A., & Lewis, C. (1993, May). Do algorithm animations assist learning?: an empirical study and analysis. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* p. 61-66.

Student project, University of Pittsburg (2012). Retrieved February 10, 2013, from http://www2.sis.pitt.edu/~is2470pb/Fall02/Final/rm/interactive.html

Urquiza-Fuentes, J., & Velázquez-Iturbide, J. Á. (2009). A survey of successful evaluations of program visualization and algorithm animation systems. *ACM Transactions on Computing Education*, *9*(2), 9.

Urquiza-Fuentes, J., & Velázquez-Iturbide, J. Á. (2013). Towards the Effective Use of Educational Program Animations: the Roles of Student's Engagement and Topic Complexity. *Computers & Education, 67,* 178-192.