International Journal of Environmental & Science Education
Vol. 5, No. 2, April 2010, 189-200

A comparison of V-Frog© to physical frog dissection

James P. Lalley, Phillip S. Piotrowski, Barbara Battaglia,
Keith Brophy, Kevin Chugh

Received 24 June 2009; Accepted 13 November 2009

The purpose of the present study was to examine and compare the effectiveness of virtual frog dissection using V-Frog© and physical frog dissection on learning, retention, and affect. Subjects were secondary students enrolled in year-long life science classes in a suburban high school (N=102). Virtual dissections were done with V-Frog©, a virtual reality software application that allows users to work with a virtual specimen that can be cut and explored in ways that are therefore unique for each individual user. The study employed a pretest, posttest, delayed posttest design using the pretest as a covariate in the analysis of the posttest and delayed posttest. Scores on a posttest administered immediately following treatment indicated that the virtual group learned more than the physical group ($p \leq .001$). Delayed posttest scores indicated there were no effects for treatment found. In the area of affect, survey results were fairly even between the two groups. Students did not show superior retention using V-Frog©. However, it should be noted that with no additional instructional cost, students could repeat the virtual dissection to improve retention. The results of the study indicated that the V-Frog© provides a viable alternative to physical dissection that produces effective learning outcomes and may be appealing to teachers and students for a number of practical and/or ethical reasons.

Keywords: adolescent, biology, dissection, frog dissection, virtual reality

Introduction

Physical dissection is a long-accepted teaching practice in high school curricula (Physicians Committee for Responsible Medicine, 2009). Marszalek and Lockard (1999), Offner (1993), and McCollum (1988) all agree that its value lies in being hands-on and exploratory, which promotes student inquiry.

However, there is resistance to physical dissection by some based on a number of issues and in the United States a number of states have enacted legislation stating that schools requiring physical dissection must offer an alternative to students voicing objections (Duncan, 2008). According to People for the Ethical Treatment of Animals (PETA), there are very real moral and ethical concerns over killing an animal for the sake of learning (PETA, 2004). In addition, there may be limited opportunity for student learning due to specimen decay which undermines the effectiveness of physical dissection (American Anti-Vivisection Society, 1996). Madrazo (2002) similarly notes the inability to repeat a procedure due to cost of additional specimens and logistical concerns. According to Orlans (1988) another challenge to the practice of physical dissection is the need for students to share a specimen due to the high cost. This limits hands-on learning opportunities. According to Balcombe (1997), there are also health and safety issues

ISSN 1306-3065

http://www.ijese.com/

related to the use of dissection instruments and chemicals used in the preservation of specimens. Balcombe (1997) points out the poor integration of multiple media (specimen and lab report/text). Furthermore, Balcombe (2000) states that the compromise of ecosystems when too many frogs are harvested is a major concern as well.

Viable alternatives to physical dissection are available, enabled by computer technology. These “digital dissections” allow students to use interactive software that features animated objects, charts and diagrams, and multimedia such as video clips. It should be noted that there are also challenges inherent to alternative dissections in the form of computer simulations. Resources for virtual dissection include computers and software; while most schools currently have computers available for student use, there are costs in terms upgrades, software, and support. Technological malfunctions during instruction can also pose a challenge.

While research has demonstrated that dissection alternatives can produce comparable learning outcomes to physical dissection, they do not provide opportunities for actual interaction with a specimen in terms of cutting, probing, and exploration. One product that was designed to bridge this digital divide is V-Frog© by Tactus Technologies, a virtual reality software application that allows users to work with a virtual specimen that can be cut and explored in ways that are therefore unique for each individual user. Unlike multimedia packages which feature either precut specimens, or require the user to use a mouse to connect predetermined hot points to reveal a predetermined cut, using V-Frog© the cutting for each individual user is a live, real-time interaction that is unique for each user. According to *Scientific Computing* (2008):

> V-Frog© virtual-reality based dissection software allows not mere observation, but actually simulates nearly unlimited manipulation of specimen tissue. As a result, every dissection is different, reflecting each student’s individual work. Students can ‘pick up’ a scalpel, cut open V-Frog’s skin, and explore the internal organs with true real-time interaction and 3-D navigation (p.11).

The introduction of this new type of dissection simulation may represent the beginning of a paradigm shift in this instructional domain. The purpose of the experiment described below was to examine outcomes in learning, retention, and affect when comparing V-Frog© to physical dissection.

Review of Literature

Research on virtual dissection has produced inconsistent results. In some cases there were no differences in learning found between physical and virtual dissection. Montgomery (2008) found no differences in learning about frog anatomy between adolescent biology students learning with *Cyber Ed Dissection Series* and those learning with physical dissection. This is consistent with the research of Kinzie, Strauss, and Foss (1993), who found similar learning outcomes for students who completed physical dissection and students learning with an interactive videodisc.

Other research has found differences in learning outcomes. Cross and Cross (2004) compared advanced adolescent biology students’ performance completing a physical dissection protocol. Prior to completing the protocol, they completed either a computerized frog dissection using the multimedia application *Biolab Frog Dissection®* or a physical dissection. They found that students completing the physical dissection performed better on the protocol. Similarly, Marszalek & Lockard (1999) found that adolescent science students completing a physical dissection produced superior learning gains from pretest to posttest when compared to *Digital*

Frog®, a multimedia dissection application. When they measured retention over time, however, they found that these differences dissipated.

Conversely, Predavec (2001) found that first year undergraduates learned more during rat dissection using *E-Rat* than those dissecting a physical specimen. Similar outcomes were found by Velie and Hall (1999), who found better learning outcomes for adolescent biology students learning with online instruction using *Frog Dissection Lab* (http://www.ofsd.k12.wi.us/science/frogdiss.htm) than those completing physical dissection. This was true whether the learning measures were based on materials used in the online dissection or the physical dissection.

Based on the varied outcomes of dissection research conducted to date, further research is warranted on the effectiveness of virtual dissections when compared to traditional physical dissections. This is particularly true in light of the continuing technological innovation in terms of the instructional materials and delivery systems available in classrooms. One such innovation is the real-time, 3-D virtual reality software *V-Frog*©. Therefore, the research question for the current study is: will virtual dissection produce comparable learning and affect outcomes when compared to traditional dissection? The first hypothesis being tested is that comparable learning outcomes will result for students completing a virtual dissection using *V-Frog*© when compared to those completing a physical dissection. The second hypothesis being tested is that comparable retention outcomes will result for students completing a virtual dissection using *V-Frog*© when compared to those completing a physical dissection. The third hypothesis is that students will prefer virtual dissection.

Method and Procedure

Subjects

Subjects were secondary students enrolled in year-long life science classes in a suburban high (N=102). All classes followed the same curriculum throughout the school year. The self-reported demographic data indicated that the make up of the sample for gender was 48 males and 54 females. For year in school there were 100 sophomores and two seniors, with an average age of 15 and one-half years. There was one senior in each treatment group. The racial make up of the sample was three African-American, two Asian, two Hispanic, two Native American, 91 White and two Other (*Other* was a choice on the demographic form provided to avoid students having to make a forced decision that did not truly represent how they perceived themselves). The initial subject pool consisted of 106 students who received treatment. However, four subjects were eliminated from the data analysis due to being absent for a data collection session or having completed a prior frog dissection.

Treatment

There were two treatment groups: virtual dissection and physical dissection. Students were randomly assigned to treatment at the class level (i.e., existing, intact classes were assigned to one of two treatments). Treatment for both groups consisted of one learning session that concluded with a 15 item multiple-choice test on the content of the lesson and a seven item survey to measure affect regarding dissection. Students in both treatments completed the same pencil-and-paper lab to guide them through the lesson. Students completing a virtual dissection completed their lesson and lab report independently. Students completing a physical dissection worked in pairs completing the dissection but were closely monitored to assure they completed their lab reports independently. Working in pairs during physical dissection was a common practice for the students involved and is common practice during many dissections lessons (Orlans, 1988). This practice was adhered to during the study to avoid creating an artificial learning environment for physical dissection. Class periods were 42 minutes in length. Students were allowed 31 minutes to complete the lesson and 11 to complete the test and survey. Students had little difficulty completing the lesson, test, and survey in one class period. Students completed treatment and assessment during their regular life sciences classes.

Curriculum

Because of the vast amount of information to be learned during a dissection, the amount of content used for the study had to be limited to give students sufficient time for learning and assessment in one class period. The researchers worked with the participating teachers in determining a reasonable amount of content to be learned given the one period time constraint. Because they involve the standard procedures of cutting and examination, internal anatomy and the digestive system were chosen as the curriculum for the study.

Teachers and Instruction

Both of the teachers participating in the study had a minimum of 10 years experience teaching life sciences (previously referred to as biology). Both groups completed parallel lab reports with guiding questions. Students in the virtual group obtained information from the interactive software, while students in the physical group had the same information (e.g., labeled drawings, definitions, organ functions, and locations) available to them in their lab report. This guaranteed the same information was provided to all students and that the only difference between treatments was the media used to present them. Further, teachers provided only procedural and safety instruction, they did not teach about content.

Independent and Dependent Variables

The study employed a pretest, posttest, delayed posttest design. The independent variable for the study was treatment type (virtual or physical). The dependent variables were the posttest to measure learning completed immediately after treatment, the survey measuring affect completed immediately after the posttest, and the delayed posttest to measure retention completed one week after instruction. The survey consisted of seven items that required students to indicate their degree of agreement with provided statements on a five-point Likert scale with one indicating low agreement and five indicating high. The survey items were as follows:

1. How much do feel you learned from this lesson?
2. This was a great way to learn about anatomy.
3. This is the way I like to learn.
4. I wish there more lessons like this in my classes.
5. Lessons like this make learning more fun.
6. Lessons like this are easy to understand.
7. Will you tell your friends about this lesson?

Pretest/Posttest/Delayed Posttest

Because students were enrolled in life science classes, it was expected that they would have some prior knowledge of anatomy taught in the treatment lessons. In an effort to control the impact of prior knowledge, all students previously completing a frog dissection were not included in the data analyses. A fifteen-item, multiple-choice pretest was administered the day before instruction. Pretest scores were used as a covariate in the data analyses of the posttest and delayed posttest. It should be noted that, because of the specific nature and limited content of the curriculum (based on limits imposed by one instructional period and the need to avoid contamination of data), the pretest, posttest, and delayed posttest were the same instruments with items re-ordered for each administration. Sample questions from the test are listed below. Based on Bloom's taxonomy for the cognitive domain (Bloom, Englehart, Furst, Hill & Krathwohl, 1956), the former is at the knowledge/comprehension level and the latter is at the analysis level.

Of the four components of digestion listed below, which of the following functions using sticky mucus?

a. stomach
b. esophagus
c. duodenum
d. tongue

During digestion, in what order does food pass through these organs?

a. stomach, small intestine, large intestine, cloaca
b. small intestine, stomach, large intestine, cloaca
c. stomach, cloaca, small intestine, large intestine
d. stomach, small intestine, cloaca, large intestine

Data Analyses

All data analyses were performed with the *Statistical Package for the Social Sciences* (SPSS) version 16.0. The data were analyzed for all complete cases (demographics, pretest, posttest, affect, and delayed posttest). As noted, students indicating that they had previously dissected a frog were excluded from the analyses (they did, however, complete treatment because assignment to treatment was done at the class level). This resulted in 60 students in the virtual group and 42 in the physical. The learning and retention data were analyzed using a Repeated Measures Multiple Analysis of Covariance (MANCOVA) using the factor treatment as the between subjects variable and pretest as covariate, with the posttest and delayed posttest as the repeated measures. Preference data were analyzed with a Multiple Analysis of Variance MANOVA using the factor treatment as the independent variable and the seven survey items as dependent variables. Average preference scores were also compared using an Analysis of Variance (ANOVA) using the factor treatment as the independent variable and average score on the seven survey items as a dependent variable. In addition to these analyses, Cohen's d (1988) was calculated for each of the significant differences found.

Results

Means and standard deviations for all analyses are shown in Table 1. The between-subjects MANCOVA (Table 2) indicated that there were differences between the groups [$F(1,99)=11.37$, $p \leq .001$]. Posttest scores from the 15 item test administered immediately following treatment indicated that the virtual group had higher scores than the physical group and the effect for treatment was significant [F(1,99)=18.74, $p \leq .001$]. The effect for the delayed posttest was not significant. The effect size, reported as Cohen's d (1992), was .92 for the immediate posttest.

Table 1. Means and standard deviations by dissection type

	Virtual		**Physical**	
	Mean	**SD**	**Mean**	**SD**
Posttest	12.60***	1.94	10.71	2.13
Delayed Posttest	10.87	2.18	10.00	2.05
Survey Item 1	3.60	.69	3.45	.94
Survey Item 2	3.93	.841	3.95	.91
Survey Item 3	3.23	1.23	3.74*	1.17
Survey Item 4	3.48	1.10	3.81	1.26
Survey Item 5	3.55	1.20	3.95**	1.19
Survey Item 6	3.98***	1.03	3.29	1.04
Survey Item 7	3.47	1.24	4.12**	1.02
Survey Average	3.61	.81	3.76	.73

Significance Levels for Effect of Treatment*$p \leq .05$, **$p \leq .01$, ***$p \leq .001$

Table 2. MANCOVA results for posttest and delayed posttest

	Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Posttest	Intercept	736.811	1	736.811	207.518	.000
Between	Pretest	57.463	1	57.463	16.184	.000
Subjects	Treatment	66.534	1	66.534	18.739	.000
	Error	351.509	99	3.551		
Delayed Posttest	Intercept	473.328	1	473.328	128.817	.000
Between	Pretest	89.166	1	89.166	24.267	.000
Subjects	Treatment	8.254	1	8.254	2.246	.137
	Error	363.768	99	3.674		

The MANOVA (Table 3) analyzing individual survey items indicated effects for treatment [F(1,99)=5.83, $p \leq .001$]. Specifically, effects were found for survey item 3 (This is the way I like to learn.) [F(1,99)=4.35, $p \leq .05$], item 5 (Lessons like this make learning more fun.) [F(1,99)=2.80, $p \leq .01$], item 6 (Lessons like this are easy to understand) [F(1,99)=11.18, $p \leq .001$], and item seven (Will you tell your friends about this lesson?) [F(1,99)=7.89, $p \leq .01$]. Items three (effect size .43), five (effect size .33), and seven (effect size .57) favored physical dissection while item 6 (effect size .66) favored virtual. The ANOVA (Table 4) comparing average survey scores was not significant.

Table 3. MANOVA results for survey items

	Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Survey Item 1	Intercept	1228.774	1	1228.774	1896.116	.000
	Treatment	.538	1	.538	.831	.364
	Error	64.805	100	.648		
Survey Item 2	Intercept	1536.323	1	1536.323	2031.149	.000
	Treatment	.009	1	.009	.012	.914
	Error	75.638	100	.756		
Survey Item 3	Intercept	1200.726	1	1200.726	828.931	.000
	Treatment	6.295	1	6.295	4.346	.040
	Error	144.852	100	1.449		
Survey Item 4	Intercept	1314.001	1	1314.001	970.032	.000
	Treatment	2.629	1	2.629	1.941	.167
	Error	135.460	100	1.355		
Survey Item 5	Intercept	1390.588	1	1390.588	974.110	.000
	Treatment	4.000	1	4.000	2.802	.097
	Error	142.755	100	1.428		
Survey Item 6	Intercept	1305.435	1	1305.435	1213.740	.000
	Treatment	12.024	1	12.024	11.179	.001
	Error	107.555	100	1.076		
Survey Item 7	Intercept	1421.652	1	1421.652	1066.201	.000
	Treatment	10.515	1	10.515	7.886	.006
	Error	133.338	100	1.333		

Table 4. ANOVA results for survey item average

	Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Survey Average	Intercept	1340.362	1	1340.362	2220.173	.000
	Treatment	.566	1	.566	.938	.335
	Error	60.372	100	.604		

The first hypothesis being tested, that comparable learning outcomes would result for students completing a virtual dissection using *V-Frog©* when compared to those completing a physical dissection, was rejected: students completing the virtual dissection learned more than those completing a physical dissection. The second hypothesis, that comparable retention outcomes would result for students completing a virtual dissection using *V-Frog©* when compared to those completing a physical dissection, was confirmed: there were no retention differences found between the two groups. The third hypothesis, that students will prefer virtual dissection, was rejected: scores for affect were similar and where differences were found they did not consistently favor one treatment over the other.

Discussion

Students completing virtual dissections had higher learning (posttest) scores, indicating that they learned more than those completing physical dissections. But there were no differences found for retention (delayed posttest) scores. As was noted, Cohen's d (1988) was calculated for each of the significant differences found. Cohen defines effects size as "the *degree* to which the phenomenon is present in the population" (1988, p. 9, italics in original). In the present study the *phenomena* refer to any differences found in learning, retention, and/or affect. The effect size of .92 for the immediate posttest would be considered to be a large effect (Cohen, 1992, p. 157). This indicates that, in addition to the low likelihood of making an error based on the significance level of the MANCOVA ($p \leq .0001$), the size of the significant difference was substantial.

Finding an effect for learning but not for retention is most likely the result of students forgetting some of what they learned during the time between testing sessions. This type of forgetting is common to most learning events unless there are significant opportunities for additional practice so that learning evolves into *overlearning* and skills and information are learned to the point of *automaticity* (e.g., Schneider & Shiffrin, 1977). Without such opportunities, if students were to be tested again, we would expect their scores to continue to decrease over time. Given that physical dissection is typically conducted in one day because of issues such as specimen decay and student safety, the opportunity for additional learning opportunities and their resultant effects on retention are minimal. However, because virtual dissection circumvents a number of these issues that hinder providing additional learning opportunities, students could repeat a virtual dissection to improve retention. This, of course, is contingent upon teachers recognizing and planning for forgetting being inherent to the learning process, and that additional instructional time is available for virtual dissection.

Survey results were fairly even between the two groups, based on the modest effect sizes for the four items where differences were found. Two items (three and five) had small effect sizes and favored physical dissection, while item six had a medium effect size and favored virtual dissection, and item seven had a medium effect size and favored physical. While it seems reasonable to expect that students would be enthusiastic about a virtual dissection, it may be the case that the uniqueness of physical dissection for most students may make predictions regarding affect erroneous.

In terms of how this study contributes to the body of research on this topic, its results conflict with those of Montgomery (2008) and Kinzie, Strauss, and Foss (1993), who found no learning differences between physical and virtual dissection. They also conflict with those of Cross and Cross (2004) and Marszalek and Lockard (1999), who found that students learned more with physical dissection than virtual. Its results are, however, consistent with those of Predavec (2001) and Velie and Hall (1999), who found that students learned more from virtual dissections than from physical. The results of this study are also consistent with the retention data collected by Marszalek and Lockard (1999), who found that differences in learning dissipated when measured over time. Based on the results of the current study and existing research, the answer to the research question posed: "will virtual dissection produce comparable learning and affect outcomes when compared to traditional dissection?", is largely answered in the affirmative.

Conclusion

Given that this body of research has demonstrated that students completing virtual dissections showed successful learning outcomes on several occasions, the implication for teaching is that virtual dissection is a viable alternative to physical dissection. This is likely to be particularly true if teachers can identify software, or specific components of software, that closely aligns with the experience of physical dissection. According to Duncan (2008), however, this issue extends beyond instructional choice: the need to offer choice has been mandated in many educational settings. In such circumstances, virtual dissection may provide learning opportunities to students who would not engage in, and learn from, physical dissection for either moral or ethical concerns, and/or health concerns related to chemicals and hazardous laboratory instruments. Returning to the notion of overlearning, if there is sufficient instructional time, virtual dissection and physical dissection could likely produce better learning outcomes than either would individually; given that students would be given the opportunity to learn, and possibly overlearn, on multiple occasions.

Implications for Further Research

In terms of limitations of the current study, it was conducted in a mostly white, suburban, middle-class setting. Generalizing the results to other adolescent populations should be done with careful consideration. Generalization to younger or older learners merits the same caveat. Further, the curriculum for this study was limited to internal anatomy and digestion to satisfy logistical constraints. The inclusion of a more comprehensive curriculum, particularly if delivered over multiple learning sessions, might have produced different results.

While not an actual limitation of the current study, future research that compared multiple virtual dissection programs to each other as well as physical dissection would add to the existing body of research. Also, research combining various programs with physical dissection, perhaps with a focus on sequencing, would also add to our understanding of this topic.

References

American Anti-Vivisection Society. (1996). Dissecting dissection. *The AV Magazine*, 105(3), 2-33.

Balcombe, J. (1997). Student/teacher conflict regarding animal dissection. *American Biology Teacher*, 59(1), 22-25.

Balcombe, J. (2000). *The use of animals in higher education: problems, alternatives, & recommendations*. Washington, DC : Humane Society Press.

Bloom, B., Englehart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York, Toronto: Longmans, Green.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

Cross, T. R., & Cross, V. E.(2004). Scalpel or mouse. *American Biology Teacher*, 66 (6), 408-411.

Duncan, A. (2008). To dissect or not: student choice-in-dissection laws ensure the freedom to choose. *Journal of Law & Education*, 37(2), 283-289.

Kinzie, M. B., Strauss, R., & Foss, M. J. (1993). The Effects of an Interactive Dissection Simulation on the Performance and Achievement of High School Biology Students. *Journal of Research in Science Teaching*, *30*(8), 989-1000.

Madrazo, G.M. (2002) The debate over dissection: dissecting a classroom dilemma. *Science Educator*, *11*(1), 41-45.

Marszalek, C., & Lockard, J. (1999, February). *Which way to jump: conventional frog dissection, cd-tutorial, or microworld?*. Paper presented at the National Convention of the Association for Educational Communications and Technology, Houston, TX.

McCollum, T. L. (1987). The effect of animal dissections on student acquisition of knowledge of and attitudes toward the animals dissected. *Dissertation Abstracts International*, 48(10), 2592.

Montgomery, L. (2008). A comparison of the effectiveness of virtual and traditional dissection on learning frog anatomy in high school. *Dissertation Abstracts International*, 68, 11.

Offner, S.P.R.I.Y. (1993). The importance of dissection in biology teaching. *The American Biology Teacher*, *55*(3), 147-149.

Orlans, F.B. (1988). Debating Dissection. *The Science Teacher, 55*(8), 36-40.

Physicians Committee for Responsible Medicine (2009). *Elementary and secondary schools: conscientious objection in the classroom*. Retrieved April 30, 2009 from http://www.pcrm.org/resch/anexp/elem_sec_alternatives.html

PETA. (2004). How animals are collected and killed for dissection and the alternatives you can choose, *The PETA Guide to Animals and the Dissection Industry. People for the Ethical Treatment of Animals, Norfolk, VA.*

Predavec, M. (2001). Evaluation of E-Rat, a computer-based rat dissection, in terms of student learning outcomes. *Journal of Biological Education*, *35*(2), 75.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, 84, 1-66.

Scientific Computing (2008). V-Frog©Simulates Dissection, *Scientific Computing*, 25(2), 11-11.

Velie, S., & Hall, T. (1999) *Virtual frog dissection-reality check?* Retrieved from Web site: http://www.ofsd.k12.wi.us/science/study.htm

V-Frog©: Virtual Reality Frog Dissection Software [Computer software]. Buffalo, NY: Tactus Technologies.

Authors

James P. Lalley is an Associate Professor of Education at D'Youville College in Buffalo, NY and is the Director of Curriculum Development for Tactus Technologies. He has conducted multiple research studies in the area of adolescence science education and has also published on the topics of mastery learning and differentiated instruction. He received his doctorate in Educational Psychology from the University at Buffalo, NY in 1997. Email: lalleyj@dyc.edu

Phillip Piotrowski is an Assistant Professor of Education at D'Youville College in Buffalo, NY. He received his doctorate in Curriculum Planning from the University at Buffalo, NY in 1986. He has taught at the elementary and post-secondary levels for over 35 years. His previous research has been in the areas of the academic impact of all-day kindergarten, gender representation among elementary teachers, and adolescence science education. **Correspondence:**