A pilot study of students’ learning outcomes using didactic and Socratic instructional methods: An assessment based on Bloom’s taxonomy

Oluwatoyin Adenike Akinde

Department of Graduate Teacher Education, College of Education and Science, Concordia University, St. Paul, Minnesota, USA.

Received 15 October, 2015; Accepted 26 October, 2015

This work is a pilot study on the learning outcomes of students, who were taught a research course for seven weeks, using didactic and Socratic instruction methods. The course was taught in two sessions concurrently. The students were divided into two groups (A and B) and both groups were taught either with Socratic instruction method or didactic instruction method. At the end of the 7 weeks, the students were tested. The test, which was valid and reliable, was categorized into 5 domains of Bloom’s taxonomy: analysis, comprehension, evaluation, knowledge and synthesis. Based on the results, there was no evidence to show if there is a difference in the learning outcomes of groups A and B. There was an outlier in the synthesis domain. Interpretation of students’ projects and final papers shows a difference in the degree of creativity. This is however tangential to the main research question.

Key words: Didactic method, Socratic method, Bloom’s taxonomy, pilot method.

INTRODUCTION

This study was an attempt to compare students’ learning outcomes, using didactic and Socratic instructional methods. The student participants in this study were registered in one of two 7-week research courses that were taught concurrently. Group A consisted of students taught with didactic method, and group B consisted of students taught with Socratic method. In the process of fine-tuning and standardizing the instructional methods for the course, a pilot method was necessary to know which methods the students responded better to. The syllabus and learning outcomes for the course were established, but the means of knowing the learning outcomes were not established. In addition to the syllabus and learning outcomes, the assessment questionnaire (the main instrument used in this study) was formed and categorized into the different learning levels of Bloom’s taxonomy. Hence, this paper includes a review of prior empirical studies on didactic teaching method, Socratic teaching method and Bloom’s taxonomy. The primary goal of the study is to show which of the methods students responded best to. The findings can be used to inform course facilitators of the best
practices for teaching.

Research questions and hypotheses

RQ1: Is there a difference between the learning outcomes means of Group A (didactic group) and Group B (Socratic group)?

RQ2: How are the learning outcomes of Group A (didactic group) and Group B (Socratic group) compared in the five domains of Bloom’s taxonomy?

H₀: There is no difference between the total means of the learning outcomes of Group A and Group B.

H₁: There is a difference between the total means of the learning outcomes of Group A and Group B.

LITERATURE REVIEW

Didactic method

Traditional teacher-centered (didactic) direct instruction is a method of teaching in which students are passive receptors of knowledge. However, in recent times, educators have been moving toward more student-centered understanding-based (constructivist) teaching that focuses on exploration and experimentation (Smerdon et al., 1999).

Leacock (1969), says in didactic teaching method, students have minimal initiation and do little exploration, as it involves little probing for personal or intellectual meaning. Simply put, didactic teaching is facilitated with strong adherence to established curricula. In all respect, didactic teaching falls in line with traditional hermeneutics, which focuses on interpreting text (Outhwaite, 1985).

Nadler et al. (2003), considered didactic instruction as a principled-based teaching. Moreover, these authors reported that in a study of training negotiators, that is, principle-based training, the students’ confidence decreased. However, other investigations reveal that providing learners with a principle is not as effective as other types of learning, such as analogical learning (Gick and Holyoak, 1983; Loewenstein et al., 1999; Ross and Kilbane, 1997).

One of the drawbacks of didactic method is that unless there is a close connection between a principle and relevant examples, students cannot understand abstract principles (Ross and Kilbane, 1997). Therefore, this method should be complemented with other teaching methods, in order to increase students’ confidence.

Socratic method

Questioning is one of the key elements of Socratic method (Overholser, 1993). According to Overholser’s argument, Socratic method encompasses other things, such as inductive reasoning and active learning. He also argued that the method is a complex interplay of questions, content and process. Proponents of Socratic method also refer to it as Socratism, which emphasizes high-level cognitive reasoning, and content that induces independent problem-solving skills (Kearney and Beazley, 1991; Seeskin, 1987). The method requires collaborative interaction between the learner and the learned. This method can be used to facilitate self-guided discovery, helping students realize the answers they already have (Overholser, 1993). Butler (1997), postulated that focusing on the Socratic questioning method is a key to constructivist education. In an earlier advancement, Freire (1973) described Socratism in line with constructivism, saying it is a way of guiding learners to develop their own critical consciousness.

Tjosvold et al. (1977), described inquiry teaching as a method of teaching that supports students in examining, investigating, and exploring questions and situations to help their understanding. It enables students to discover their own insights. Socratic method is meant to help students develop problem-solving skills (Overholser, 1993) rather than simply accumulating factual information on subject matter. For Tjosvold et al. (1977) the expected outcome is for students to experience scientific methods of discovering and creating knowledge. Supporters of inquiry teaching believe that this method of teaching requires students to develop complex learning skills.

Bloom’s taxonomy

Bloom’s taxonomy is a multi-tiered model of classifying thinking according to six cognitive levels of complexity. Forehand (2010), depicted these levels as a stairway that leads teachers to encourage their students to “climb to a higher level of thinking.” She added that the lower three levels are knowledge, comprehension and application; and the higher three levels are analysis, synthesis and evaluation.

At the inception of Bloom’s taxonomy, it included knowledge, comprehension, application, analysis, synthesis and evaluation (Bloom, 1956). In the 1990s, Sosniak (1994) reexamined and re-conceptualized it as remembering, understanding, applying, analyzing, evaluating and creating. Figure 1 is a representation of the differences between the original version of Bloom’s taxonomy by Benjamin Bloom (1959), and the revised
version by Sosniak (1994) and Anderson (1999). Their new illustrations note a change between the versions from nouns to verbs. Sosniak (1994), in particular, described this change as an effort to describe the different levels of the taxonomy. He further extrapolated that the top two levels are essentially exchanged in the old to new versions.

Krathwohl (2000), argued that both versions are inherently similar. In his attempt to demystify this argument, he stated that it was a matter of verb vs. noun; suggesting that the revised gives greater weight to teachers’ usage. He observed that synthesis and evaluation were exchanged, therefore, create was coined. The revision represents a hierarchy in the cognitive processes, which differ in their levels of complexity. In this vein, he cited, for instance that remember was less complex than understand and also less complex than apply. The levels of complexity are however not the focal point in this study; it is a tangential premise at best.

Thus, the aim of this study is to help the faculty of the student participants in the study to gauge the values of the two instructional methods (didactic and Socratic methods) and the degree to which they support students’ learning as examined by Bloom’s taxonomy. In terms of value added, this study prods faculty members to intentionally use instructional methods while teaching the research course, based on insights from this study. As typified by Forehand (2010), researchers and educators use either the original or the new version. Therefore, this study will adopt the original version of the learning taxonomy as presented by Anderson (1999) and Huitt (2004).

**METHODOLOGY**

Pilot-study was the design employed for this study. It permits preliminary testing of a methodology and provides the researcher with ideas, approaches, and clues that are unforeseen (Fraenkel and Wallen, 1993). For instance, the ancillary section of this study provides a few tangentially related evidences gathered through the course of this study. Though there were not anticipated, they appeared to be relevant findings. Hence, the use of pilot-study is invaluable in this instance.

**Context of study**

This study was premised on two types of instructional methods in an undergraduate-level research course, which is a science and general education requirement at the institution where this study took place. The purpose was to evaluate students’ learning outcome under the instructional model of didactic method and Socratic Method. Students’ learning was evaluated based on administering a 20 questions test, which was comprised of 15 multiple-choice and five binary questions. The test questions were grouped in the following domains: Analysis, comprehension, evaluation, knowledge and Synthesis. The activities associated with the instructional methods differ in terms of how the learning objectives were met. Appendix A, being the lesson structure, provides an illustration.

The course under study was offered in two different sessions concurrently, in a Midwest college campus. Group A consisted of 11 students and Group B consisted of 14 students. The study ran for 7 weeks. Each week, students were required to spend three hours in the classroom, with 21 total hours of in-class instruction. The students, being the participants in the study, were considered homogenous on the basis of having met the prerequisite for the course, which were statistics and composition. Furthermore, the students were at the same level on their programs course sequence That is, how many courses they have taken in their programs. The course under study took place in the college setting, and the course was a required liberal arts general education course in partial fulfillment of a baccalaureate degree.
Students, as participants, were predominantly male, over 18 years of age (therefore able to consent) were either in the criminal justice, or information and systems security or electronics programs. No other demographic data was collected.

Students who registered for the course were approached by the course instructor (who was also the researcher) to discuss the research agenda and solicited students’ participation. This researcher informed the students that the course would be facilitated based on the didactic method or Socratic method. Further explanations of both methods were provided, and that two sessions of the course were offered concurrently, hence having Group A and Group B. Moreover, students were informed that their being in the Socratic or didactic group was random and that no specific requirement was considered for the grouping. After students were provided this information, the researcher/instructor asked for students’ consent. Students were given a choice to opt-out of the study, but none of them did. They were guaranteed anonymity, in that their academic record, names, and pertinent data was not going to be included in the study. That their participation meant that they are registered in the course, are required to attend class, participate in the learning activities designed for the course, and take a final exam; all of which were the norms, if they were to pass the course. That said, they were not required to do anything extra. Finally, as a part of the information to students, and consent process, they were informed that their final exam results would be evaluated to determine the effectiveness of the instructional method they were part of. It should be noted that after being presented with the information, all the students consented in writing, by each signing a consent form.

Group A class took place on Tuesdays using a traditional face-to-face model. The course was didactically taught in ways that involved the use of traditional hermeneutics. In this regard, the teaching was heavily based on delivering content and providing heuristics. The goal of the instructor was to impart knowledge. The students were required to write a weekly paper on assigned chapters. They were required to take a 20-question examination, write a final paper on a student-selected topic, and present the paper as a project.

Group B class took place on Wednesdays, in a face-to-face classroom. Their learning was facilitated using the Socratic method. The instructor took more of a facilitator’s role, encouraged challenges, induced learning through questioning, allowed some degree of constructivism, and used debate as a means for teaching. The students participated in small- and large-group discussions. As with Group A, Group B students took a 20-question examination, wrote a final paper on a student-selected topic, and presented the paper as a project.

The 20-question paper-and-pencil examination was the way in which student’s learning was assessed, based on the five learning domains of Bloom’s taxonomy which measured the differences in learning outcomes between Group A (didactic pedagogy) and Group B (Socratic pedagogy). Also, both groups were required to develop a research proposal.

**Sampling technique**

Being a pilot study, the goal was to test two methods of teaching in order to know which was most effective for students. Hence, two versions of the course were offered concurrently. Convenience sampling was employed for the entire class in both sessions.

**Statistical test**

The parametric two-sample t test is robust and not particularly sensitive to distribution, and therefore was used due to the relatively small sample size instead of the Mann–Whitney test. In this study, Minitab was used for the calculation of the results. Furthermore, a confidence interval of 95% was used because an alpha of 0.05 is assumed for most research and the researcher saw no reason to deviate from this.

**Learning taxonomies**

The students were evaluated using five of six Bloom’s taxonomy domains. The domains measured were analysis, comprehension, evaluation, knowledge, and synthesis. Application, the sixth domain, was not included in the examination. The examination included a total of 20 questions. They were categorized as follows: Nine questions measured analysis, seven measured comprehension, two measured evaluation, one measured knowledge, and one measured synthesis.

**Instrument**

The instrument used in this study was a questionnaire consisting of 20 multiple choice questions. The 20 questions were from a 50-question test designed for the course. For proprietary reasons, the actual questions are not included in this paper. The questions were written using verbs adopted from sample questions of Bloom’s taxonomy by Huit (2004). According to Huit, for questions pertaining to analysis, each included at least one of the following verbs: analyze, appraise, categorize, compare, contrast, criticize, differentiate, distinguish, examine, and experiment. Questions in the comprehension domain included the following verbs: classify, describe, discuss, explain, identify, indicate, recognize, report, select, and translate. In the evaluation domain, the questions included the following verbs: evaluate, assess, and compare. The question in the knowledge domain included the verb define. The question in the synthesis domain included the following verbs: construct and organize. The instrument used was considered as proprietary material and not available to the public domain. That is, the instrument was designed by a third-party consortium, in collaboration with this researcher. Nonetheless, it was tested for validity and reliability at the inception of the course. However, instructional method was the main question that required answers; hence the significance of this study.

**RESULTS**

The aim of this study is to ascertain if there are differences in learners’ outcomes in the 20-question examination, using different teaching methods. In the results, the degrees of freedom are different because of unequal variances. This may also explain why there are differences between the degrees of freedom for all the domains (Table 1). The total mean of Groups A and B is not likely to be different. In the analysis, comprehension, and knowledge domains, were not likely to be different.
On the other hand, in the evaluation domain, both groups were more likely to be different, as group A had 9.091 and group B, 6.43. In synthesis domain, there was a similarity, as Group A had 3.636 and Group B, 5.000 (Table 2).

As an alpha of 0.05 is smaller than p-value of 0.410, the null hypothesis cannot be rejected. This means the data do not show that there are differences between the groups as shown in the total scores. In doing other tests in future, to have higher degrees of freedom, the sample size can be increased. In this case, the degrees of freedom were somewhat reduced due to the difference in sample size. In Table 3, the p-value of 0.945 is larger than 0.05. Therefore, the null hypothesis cannot be rejected. This means the data do not show a difference between the groups, as seen in the evaluation scores. However, if an alpha of 0.10 were used, it would likely be significantly different.

In Table 4, the p-value of 0.076 is a bit larger than an alpha of 0.05. Therefore, the null hypothesis cannot be rejected. This means the data do not support a difference between the groups, as seen in the knowledge scores. There is an outlier in Group A whereas Group B had some variability. The impact of the outlier is minimal in particular because the null hypothesis
Table 5. Evaluation domain results (Two-sample t-test and CI: Evaluation 10% A vs. Evaluation 10% B).

<table>
<thead>
<tr>
<th>Variable (%)</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev</th>
<th>SE Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation 10 A</td>
<td>11</td>
<td>9.09</td>
<td>3.02</td>
<td>0.91</td>
</tr>
<tr>
<td>Evaluation 10 B</td>
<td>14</td>
<td>6.43</td>
<td>4.13</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Difference = μ (Evaluation 10% A) - μ (Evaluation 10% B); Estimate for difference: 2.66; 95% CI for difference: (-0.30, 5.63); T-test of difference = 0 (vs not =): T-Value = 1.86 P-Value = 0.076 DF = 22.

Table 6. Knowledge domain results (Two-sample t-test and CI: Knowledge 5% A vs. Knowledge 5% B).

<table>
<thead>
<tr>
<th>Variable (%)</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev</th>
<th>SE Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge 5 A</td>
<td>11</td>
<td>4.55</td>
<td>1.51</td>
<td>0.45</td>
</tr>
<tr>
<td>Knowledge 5 B</td>
<td>14</td>
<td>3.21</td>
<td>2.49</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Difference = μ (Knowledge 5% A) - μ (Knowledge 5% B); Estimate for difference: 1.331; 95% CI for difference: (-0.343, 3.005); T-Test of difference = 0 (vs not =): T-Value = 1.85 P-Value = 0.113 DF = 21.

Figure 2. Mode and frequency graph.

cannot be rejected.

For the synthesis domain, based on the boxplot, one could hypothesize that the mean of Group B is higher than that of Group A. However, this cannot be called a statistical difference because Minitab was not able to calculate it due to the fact that Group B values were identical; that is, had no variability (Figure 8).

Results of the research questions

Research question 1: Is there a difference in the learning outcomes means of Group A (didactic group) and Group B (Socratic group)? There is no evidence to support if there is a difference in the total scores based on the p-value of the two-sample t test (Figure 2).
Research Question 2: What is the comparison in the learning outcomes of Group A (didactic group) and Group B (Socratic group) in the five domains of Bloom's taxonomy?

For the analysis, comprehension, evaluation, and knowledge domains, there is no evidence to support if there is a difference based on the p-value of the two-sample t test. The Boxplots (Figures 3 to 7) provide a visual
For the synthesis domain, it appears that Group A was higher than Group B based on what was observed from the Boxplot (Figure 7), but it cannot be called statistical difference.
DISCUSSION

The student participants in this study took the same course over a 7-week period using a traditional face-to-face model. However, Group A (11 students) was exposed to the didactic teaching model, whereas Group
B (14 students) used the Socratic teaching model. At the end of the course, both groups did projects individually that included a self-selected topic, research paper, and presentation. Group B also participated in a debate. And, both groups were assessed using the same 20-questions examination. Each participant could score up to 100% in the examination. The comparison of the results of the 20-questions examination between the two groups included assessing the overall results of the examination and using of Bloom’s taxonomy domain. Statistically, the examination results for Group A and Group B do not appear to be significantly different.

CONCLUSION

First, based on the findings in this study, instructors may use either Socratic method or didactic method in teaching the course. From this study, it could be inferred that students’ test results are not likely to be different. Secondly, an observation from this study shows that the Socratic group employed creativity based on the artifacts and diversity of approach used in presenting their projects. Group A’s projects, from the students who were exposed to didactic instruction, appeared to be generally linear. In part, they all used PowerPoint presentations, did not introduce any artifacts, and did not raise new questions (greater illustration of this is provided in the ancillary findings section that follows). Bringing both the test results and project outcomes to bear, it is best to employ both Socratic and didactic methods in teaching the course. This may help to balance students’ acquisition of the content and prepare them for further inquiry, especially in using different research methods, such as qualitative to address complex issues. Simply put, didactic teaching methods can enhance knowledge acquisition—that is, provide students with the heuristics of inquiry. However, employing the Socratic teaching method will help students beyond possessing heuristics to produce creative outcomes.

Limitations of study

Pilot-study is considered a small-scale study trial of a proposed procedure, which is used to test models, hypotheses, and detect problems that could be remedied for future practice or study (Fraenkel and Wallen, 1993). Therefore, findings from a pilot-study may only apply to the context of study – which makes generalizability limited. Secondly, the use of convenience sampling, based on using the course that was accessible, and sample size being 11 and 14 students pose as limitations. Finally, the result of this study is limited; in part, it could only be used to inform future instructors of this course about what outcome to expect if the same testing questionnaire is used and if the course study is facilitated using didactic or Socratic method. Simply put, students may perform differently if other instructional method other than didactic or Socratic method is used.

Future research

Based on the limitations of this study, future researcher could consider a different research method other than a pilot-study. Therefore, using a random sampling technique, with a larger sample size would be better.

Ancillary findings

For the students’ self-selected projects, they were required to turn in a paper and present their findings using different visuals including PowerPoint presentations, videos and collages. Group A’s (didactic method) papers reflected key terms which indicated that the students understood the topic vastly. The themes from Group A’s papers suggest that the students had knowledge of the topic and comprehended the subject. The visuals, which mainly included PowerPoint presentations, reflected the students’ understanding of the course contents. All 11 students used PowerPoint presentations and no other visuals. Outcomes were direct, simple and concise. Group B’s (Socratic Method) papers appeared to demonstrate knowledge and comprehension of the topic. Nine of the 14 students’ papers had some degree of dialectic exchange of ideas and corroboration, and raised challenging points as well as unique line of inquiries, which reflected synthesis, analysis and evaluation. Visuals included PowerPoint presentations, videos, collages and documentaries. Outcomes were creative.

ACKNOWLEDGEMENTS

The author would like to acknowledge the Dean of the college where this study took place, for granting her the permission to carry out this study, for the purpose of sharing best practices and to improve students’ learning. She also wants to thank the participants (students) in the study for their support.

Conflicts of interest

The author has not declared any conflicts of interest.
REFERENCES


# APPENDIX A

## Lesson Structure

<table>
<thead>
<tr>
<th>A. Sessions</th>
<th>B. Objective</th>
<th>C. Group A Didactic Activities</th>
<th>D. Group B Socratic Activities</th>
<th>E. Notes and Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introductions</td>
<td></td>
<td>C.1.1 Students and Instructor</td>
<td>D.1.1 Students and Instructor</td>
<td>C&amp;D.1.1 Getting acquainted. Shared experience, background, and career interest.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C.2.1 Reviewed syllabus, course objectives, assignments, and assessment.</td>
<td>D.2.1 Reviewed syllabus, course objectives, assignments, and assessment.</td>
<td>C&amp;D.2.1 Explained the course objectives, how it would be met, assignments, required reading, and students’ evaluation at the end.</td>
</tr>
<tr>
<td>2 Syllabus</td>
<td></td>
<td>C.3.1 Scavenger hunt</td>
<td>D.3.1 Scavenger hunt</td>
<td>C&amp;D.3.1 Students went around the class to identify what activities that are research based that students have been involved in prior to the course.</td>
</tr>
<tr>
<td></td>
<td>Audit of students’ prior research experience.</td>
<td>C.4.1 Students report</td>
<td>D.4.1 Students report</td>
<td>C&amp;D.4.1 Students self-reported what they would like to learn in the class.</td>
</tr>
<tr>
<td></td>
<td>4 Students’ interest</td>
<td>C.4.2 Cluster grouping by interest</td>
<td>D.4.2 Cluster grouping by interest</td>
<td>C&amp;D.4.2 Students where grouped into cluster of: qualitative, quantitative, or mixed method. Group met throughout the course to discuss presentation projects, brainstorm and peer-reviewed.</td>
</tr>
</tbody>
</table>

**Week 2:**
- Introduction to research

| C.1. | 90 minutes - Lecture: Instructor gave PowerPoint based on the course textbook (traditional hermeneutics approach, Outhwaite, 1985). |
| D.1. | 30 minutes - Lecture: Instructor gave PowerPoint, providing an overview of the required reading. |
| C.1 & D.1. All students were present, some appeared engaged based on readiness to ask clarifying question. |
| C.1. | some students seemed tired after about one hour of lecture, based on some students appearing sleepy, doodling on their books, and getting up to stretch. |
| D.1. | appeared engaged, maintained eye-contact with instructor, took notes, and nodded periodically. |

**Week 3:**
- Research writing
- APA and MLA Citations

| C.2. | 25 minutes - Question and Answer: Students asked questions to clarify any ambiguity from the lecture or reading materials (Smerdon, Burkham, & Lee, 1999). |
| D.1.1. | 20 minutes - Instructor showed videos/ted-talk on research methodologies. |
| C.2. | Students took notes as their questions were answered by the instructor. |

**Week 2-6**

<p>| C.3. | 30 minutes - Students were instructed to meet in their cluster group to work on final project. |
| D.2. | 90 minutes - Question and Answer: Instructor posed questions to the class. Each group (qualitative, quantitative, and mixed-methods) were given a scenario. Students were excused to find the answer (Freire, 1973), and reported back to the class with their findings. |
| D.2. | Two groups did a show-tell of their findings and presented it to the class, while one group did a presentation of their findings. At the end, each group had new questions as a result of the group work (Seeskin, 1987). Instructor refrained from answering those questions, rather students were directed to explore their questions (Butler, 1997) using the text, conducting a research, and reporting back the following week. |</p>
<table>
<thead>
<tr>
<th>Week 7</th>
<th>Testing (20 questions test)</th>
<th>Presentations</th>
<th>D4. Debate</th>
<th>Testing</th>
<th>Presentations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Paper-pencil test.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Presentation styles were at students' liberty.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Debate was setup in a way whereby the three cluster groups were given the same question. They each defended the methodology from the perspective of their group emphasis, being: qualitative, quantitative, and mixed-methods.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>