

Question-to-Learn Science in Higher Education: A Quantitative Study

Lars Bengtsson

Department of Physics, Faculty of Science, University of Gothenburg, Gothenburg, Sweden

Email address:

lars.bengtsson@physics.gu.se

To cite this article:

Lars Bengtsson. Question-to-Learn Science in Higher Education: A Quantitative Study. *Higher Education Research*.

Vol. 5, No. 3, 2020, pp. 83-93. doi: 10.11648/j.her.20200503.13

Received: April 22, 2020; **Accepted:** May 27, 2020; **Published:** June 4, 2020

Abstract: This work reports the results of a case study where traditional activities in an engineering/science classroom, such as demonstrations and self-paced activities, were compared with ‘writing across the curriculum’ (WAC) activities. One group did writing-to-learn assignments and one group did our seminal construct ‘question-to-learn’ where they designed exam problems for their peers. A model is presented which describes the parameters that influence the exam score outcome. Some of these variables were carefully controlled during the project (labs, textbook, lecturer) and some other parameters were measured (lecture attendance, time-on-task and previous knowledge) in order to minimize data corruption due to confounding variables. The main parameter of interest, the ‘predictor’ of the exam score was the *extra-curricular activity*. A pre-test and a post-test were also conducted in order to establish the students relative gain. We also tested the hypothesis of using the quality of the students’ WAC outputs as a predictor of academic achievements. Data is analyzed both with parametric and non-parametric methods and results show that there was no significant difference between the groups on exam scores and that the relationship between WAC quality and exam scores is not significant. The main reason for the non-significant results is concluded to be due to low participation rates and too low “dosage”.

Keywords: Writing-Across-the-Curriculum, Writing-to-Learn, Question-to-Learn, Higher Order Thinking Skills

1. Introduction

In the ancient Greece, Platon argued that writing was a threat to intellectual training since it would diminish the need for memorization and foster only a pretense of wisdom [1]. Today, writing is generally recognized as an important element of any curriculum in higher education and one of the most efficient vehicles for learning. It benefits content learning and higher order thinking skills at the same time [2-5], it promotes critical thinking [6-14], metacognition [15-17], helps students develop their general ability to express themselves [18] and converts their thoughts to words [19]. The importance of a verbal language to support the learning process was emphasized already by Vygotsky [20] since the development of higher cognitive functions (such as analysis and synthesis) benefits immensely from a proficiency in verbal expression, *written* verbal expressions in particular. Exactly why writing has such a beneficial influence on learning is still debated, but it has been suggested that it slows down the brain [21, 22] and thereby gives the brain

more time to process the information and also helps students realize that they don’t understand the material as well as they might think. Writing facilitates meaningful learning because it forces students to analyze (break down) and synthesize (re-connect) fundamental scientific concepts [23, 24]. Others emphasize the metacognitive processes that are stimulated by writing; the students are forced to think about a) how they perceive the information, what they understand (and what they don’t understand) [25], and b) how they connect new information to what they already know [19, 26, 27].

There are several different ways to implement writing in the curriculum (and reasons for doing it). The general term is “writing across the curriculum” (WAC) and refers to any writing activity within the curriculum. The most common WAC activity in STEM classes (science, technology, engineering, mathematics) is “writing in the discipline” (WID) (sometimes referred to as ‘writing to communicate’ or ‘transactional writing’). WID includes lab reports and bachelor theses. It is not unusual that the bachelor thesis is the first (and only) serious WAC activity in an undergraduate

STEM program and this deficiency has been emphasized in the literature [3, 28, 29]. WID writing is characterized by very formal writing complying to strict templates [30] and there is little room for informal spontaneity.

Writing to learn (WTL), on the other hand, is the opposite of formal writing [18]. Students are encouraged to “just write”. Never mind the formalities, just try to word your thoughts. The idea is that the simple act of writing, in itself, benefits learning [31]; “writers generate knowledge at the point of utterance” [32, 33, p. 211]. A typical WTL assignment is less than one page of writing and characterized by being “low-stake”, i.e. not graded [14, 21, 33, 34]. It has been debated whether these assignments should be carried out using a pen or a keyboard and in 2014, Mueller and Oppenheimer [35] showed that the use of pen or keyboard triggers different cognitive processes and when it comes to learning, the use of pen is preferable.

It is also generally recognized that WID and WTL activities promote different cognitive levels. WID is considered more beneficial for the higher order thinking levels (analysis and synthesis) [22] while WTL is mainly a tool for the lower order thinking levels (remember and understand). In order to close the gap between WID and WTL, “writing to engage” (WTE) has been suggested [36] which is an amalgamation of WID and WTL, but no explicit research on WTE in higher education has been found in the literature. Figure 1 illustrates the relationship between the different WAC activities and Bloom’s taxonomy of higher-order cognitive thinking [37].

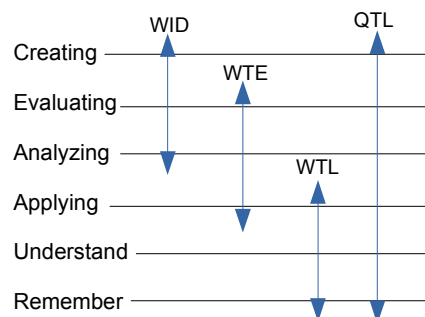


Figure 1. Different WAC activities versus thinking level orders.

It could be argued that the benign influence on learning from writing could be explained by the extra time spent on the course material. In 2007, Drabick, Weisenberg, Paul and Bubier [18] compared thinking versus writing; the students who *wrote* about the subject for some time each week outperformed those who *thought* about the subject during the same time. It is the *act of writing* that promotes learning. This conclusion about writing versus thinking was confirmed in another study in 2011 [38].

Even if the consensus of the community is in favour of WAC activities and extolls it as an important vehicle for learning, some works have reported discordant results. Armstrong, Wallace and Chang [39] reported that writing assignments had only a minimal impact on the overall content learning and Bangert-Drowns, Hurley and Wilkinson

[16] even suggested that writing activities could have an adverse effect on learning since it “takes time away from other important educational activities and decreases content coverage” (ibid, p. 33). Libarkin and Ording [40] concluded that even if students’ ability to discuss scientific concepts and draw reasonable conclusions based on data improved by using writing assignments, their ability to state a hypothesis or draw connections between human actions and environmental impacts did not improve. Sabrio, Sabrio and Tintera [41] could not detect any difference in performance between a writing and a non-writing class in mathematics and Bargate [21] purported that writing only improved students’ grade on essay-type questions.

It has been recognized by the community that learning by writing does not occur automatically but only under certain conditions [16] and some evidence of the benefits may even be anecdotal [33]. For example, tasks that promote deep processing are more likely to enhance learning [42] and the extent to which instructors explain the learning benefits of writing also affects the learning outcome [34]. Klein [33] stressed the social context’s influence on learning by writing.

Hence, writing-to-learn is a complex process influenced by a diversity of external (context) and internal (cognitive) factors and successful implementations require careful planning.

1.1. Definition of Concepts

WAC activities are generally considered to improve *critical thinking* and *scientific literacy*. To obviate any misunderstandings, we define these concepts here:

By ‘critical thinking’, we refer to the ability to conceptualize, apply, analyze, synthesize and evaluate data acquired from observations, experiments, reasoning or communication with others, and to use this as a guide in decision makings and actions [43]. It is “an intentional, self-regulated process that provides a mechanism for solving problems and making decisions based on reasoning and logic” [10, p. 141].

‘Scientific literacy’ implicates a level of knowledge and understanding of scientific concepts, conjunctions and processes to such an extent that it benefits decision making in socio-scientific issues [44]. This implies an ability to describe, explain and predict phenomena from a natural science point of view and to identify scientific issues that are likely to have significant influence on society and/or individuals. A scientific literate is able to assess evidence, pose arguments and draw conclusions based on these evidences and previous knowledge [45].

1.2. Brief Review of WAC-WTL

WAC activities have been used for a long time in higher education but did not really draw scholars’ attention until the 1970th [46, 47]. Until then, WID had been the dominating WAC activity, but first Britton [46, 47] and then Emig [22] acknowledged writing as “a unique mode of learning” [22, p. 122] since it is self-paced and scaffolds both analysis and

synthesis of new information. Griffin [48] purported that it is the *process* of writing, rather than the *product* of writing, that promotes learning. This position has been advocated repeatedly by others [14, 42, 49]. Moskowitz and Kellog [50] stated that WTL activities “treat writing as a means, rather than an end” (ibid, p. 919). Applebee [1] advocated writing because it fosters reasoning skills; “good writing and careful thinking go hand in hand” (ibid, p. 577).

WTL can be, and has been, implemented in several different ways. The most common formats seem to be informal personal journals (“micro themes” or “five-minute essays”) [33], writing informally about an experiment [31], blogs and chat forums [19], inquiry-based writing as response to external material [44, 50], written arguments for/against research articles [39] and case studies [43]. In order to emphasize the informal character of WTL, Sabrio, Sabrio and Tintera [41] prompted students to “write a letter to a relative or a friend describing what you do to prepare for the upcoming exam” and “write an introduction to chapter x for an algebra student who has not yet studied it” (ibid, p. 422). Almer, Jones and Moeckel [51] proved that as little as “one-minute papers” improved students’ performance on essay-like questions in an accounting exam. As a matter of fact, in 2019, Twitter (140 character text messages), was ranked as the 4th (out of 200) best technological tool for learning by the Center for Learning and Performance Technologies [52].

1.3. Introducing QTL

The seminal approach to WAC in this work is the introduction of the QTL concept (Question-to-Learn). Rather than writing essay-like papers, we hypothesized that students’ learning would benefit more from being prompted to design exam problems for their peers. Students solve each other’s problems in pairs and afterwards they discuss problems and solutions (providing an inherent peer-review process to the QTL concept). It has been reported repeatedly that the most important prerequisite for successful WAC is that the writing assignments must focus on the students’ cognitive processes [15, 16, 27, 42] and they must challenge students to define and explain concepts and expand on ideas [28]; we believed that the QTL writing task would meet both these prerequisites even better than WTL and WID. Prompting students to design exam problems is also consonant with other purported aspects of successful WAC; students must be allured to restructure knowledge [3] and meta-reflect on their own understanding [16]. Forced to design a challenging exam problem in combination with peer-review, we believe that both cognitive and meta-cognitive processes are triggered and we hypothesise that this will benefit learning in general and improve exam performance on question items requiring higher-order thinking skills in particular. We hypothesised that the QTL approach will elicit even deeper thinking processes than plain writing and that this will have a significant impact on both content learning and higher order thinking question items. This is illustrated in *Figure 1*. The main objective of this study was to benchmark our QTL approach to WAC in relation to traditional WTL assignments.

2. Method

The target group of this study was the “Introduction to Natural Science” program cohort at University of Gothenburg. This program is for students with a high-school diploma but not eligible for third cycle studies in the STEM disciplines (or medicine). It is a one-year program consisting of introductory courses in mathematics, physics, chemistry and biology and 179 students were accepted for the academic year 2019/2020 and approximately 130 students took the Electricity course during the second semester of fall 2019. The cohort was randomly divided into three groups; a QTL group, a WTL group and a reference group.

The project was first presented to the students in a plenum session where the presumptions were explicated: 1) participation is voluntary but attendance implicates a tacit consent to use students’ data for research purposes, 2) these extra-curricular activities do not have any (direct) influence on their grade on the course (but will hopefully have an indirect, positive influence), 3) we ask the students not to migrate between groups and/or participate in multiple activities since that would corrupt the data, and, 4) all data will be treated with professional secrecy (teacher-student privilege).

In order to establish a “baseline” of specific knowledge (in Electricity) and higher order thinking skills, all students took a pre-test consisting of a selection of test items from “Teaching Physics” by Edward Redish [53]. Neither the results nor the solutions were communicated to the students at this time. On the last lecture preceding the exam, the students took the same test again (in order to establish any progress) and this time the solutions were presented orally in plenum. On the pre-/post-test, for each question item students were also asked to indicate their degree of confidence in their answer (‘How sure are you that your answer is correct?’) on a 5-level Likert scale, ranging from ‘Absolutely sure’ to ‘Just guessing’.

The pre-test and post-test scores were used to determine each student’s *relative* improvement and the exam score was used to determine the students’ *absolute* degree of learning. Finally, one month after the exam, a *retention* test was scheduled in order to measure any differences in the retention of knowledge between the groups.

The three groups met for one hour weekly for a total of five occasions. This may seem like a short time to register any improvements in higher order thinking skills, but Quitadamo and Kurtz [10], showed that higher order thinking skills can be measurably changed in weeks. The “one-minute paper” study by Almer, Jones and Moeckel [51] also supports this approach.

2.1. QTL Instructions

The QTL group was tasked with the following assignment:

- a) Based on what was covered in the last week’s lectures, design a “good” exam problem, including an elaborate solution (on a separate paper).
- b) Exchange your problem sheet with one of your peers and solve the problem your peer designed.
- c) Exchange sheets again and correct your peer’s solution.

- d) Finally, discuss your problems and solutions until you agree on the solutions. If necessary, correct the solution to your problem.
- e) The students were also encouraged to keep this exercise in mind for the upcoming lectures and try to figure out a good problem already during the lectures (and during the in-between-lecture-times).

The instructor collected the problem and solution sheets and they were subsequently “graded”. The problems were assessed by considering legibility, originality, relevance, level of complexity, elucidating figures and “degree of effort” (do you only need to apply a formula or is higher order thinking required to solve it?). All problems were graded using a Likert scale from 1 to 5, where “5” represents a very creative problem illustrating central concepts and demonstrates a profound understanding on the designer’s part. The problem should also be considered to significantly contribute to the peer’s learning. At the other end, grade “1” represents a substandard problem characterized by minimum efforts, deficient (or faulty) assumptions and negligibly contributing to the peer’s learning. (Problems copied directly from the textbook were also graded with “1”).

2.2. WTL Instructions

The WTL assignments were designed according to ideas proposed by Kovac and Sherwood [7, 8] and Angelo and Cross [54]; short, informal and not intended for communication. The WTL group was instructed as follows: “Summarize what was presented on the lectures during the last week. Focus on central concepts, conjunctions and context and write no more than one page. In particular, try to explicate the parts you did not understand.” This exercise was carried out using pen and paper in order to optimize the learning [35].

The sheets were collected for grading (1-5), where “5” represents a student who demonstrate an excellent ability to express his/her understanding verbally (and/or lack of understanding), the student has a high level of self-awareness (of his/her deficiencies), an excellent language and has correctly identified the gist of the course material. A “1” represents substandard works with severe deficiencies in language, verbal expression and an inability to catch the central concepts and conjunctions.

NB! In both the QTL and WTL groups, the grading was surreptitious to the students. According to MacKinnin-Slaney [55], students are more likely to focus on content and clarity of expression, rather than formal aspects of writing (spelling and grammar) if papers are not graded and Bargate [21] reported that an ungraded paper group improved more than a graded group. All QTL/WTL works were always graded by the same teacher.

2.3. The Reference (REF) Group

The reference group had 45 minutes of “traditional” activities in a STEM classroom; the teacher demonstrated problems on the whiteboard half the time and the rest of the time was dedicated to questions or self-paced problem-solving. (This is the “placebo” group; activities are only intended to even out differences in teacher time between groups.)

2.4. The NULL Group

There will inevitably be a group of students that ignore any extra-curricular activities. We will refer to this group as the “NULL” group and their contribution to this study is also important. This group’s performance was registered and compared with the other three groups. The reason this group is interesting is because it has been suggested that WAC activities may not be the best way to spend students’ time [16]. Even if this group is not randomly selected, their performance may be an indication of the credence of that suggestion.

2.5. Variables

The main objective of the study was to determine if the extra-curricular activities (X_0) have any impact on the students’ learning (Y_i). The most fundamental tenet in scientific methodology states that a correlation between X and Y does not imply a causal relationship (causality); significant differences in exam scores between the groups could be explained by confounding factors. Figure 2 illustrates the confounding factors that were considered in this work. All students used the same textbook, did the same laboratory exercises and were lectured by the same teacher; these factors were constant during the experiment. Lecture attendance, Time-on-Task (how much did you study out-of-class?) and previous knowledge had to be measured in order to establish any differences between the groups. Lecture attendance and Time-on-Task were measured by a questionnaire at the end of the course and previous knowledge was measured by the pre-test.

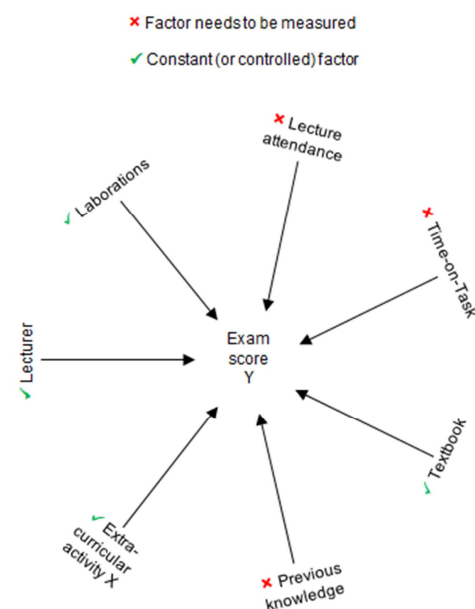


Figure 2. Learning model.

The extra-curricular group activity is the primary independent variable (‘predictor’) in this study; this is a categorical (nominal) variable (NULL/REF/WTL/QTL). This is the only independent variable that was manipulated in this experiment. Multiple dependent variables (‘outcomes’) were

considered; the students *relative gain* (post-test minus pre-test), the students *absolute gain* (exam score) and the students *retention score* (one month after exam). These three cases were also subdivided into two groups; the total score and the score on the HOCS questions only.

Table 1 summarizes the variables considered and their relationship.

Table 1. Dependent and independent variables (X_i are predictors, Y_i are outcomes).

Variable	Label	Depends on
Extra-curricular activity	X_0	-
Pre-knowledge	X_1	-
Pre-knowledge, HOCS	X_2	-
Time-on-Task	X_3	-
Lecture attendance	X_4	-
Relative gain (post-pre)	Y_0	X_0, X_3, X_4
Relative gain (post-pre), HOCS	Y_1	X_0, X_3, X_4
Absolute gain (exam score)	Y_2	X_0, X_1, X_2, X_3, X_4
Absolute gain (exam score), HOCS	Y_3	X_0, X_2, X_3, X_4
Retention	Y_4	$X_0, X_1, X_2, X_3, X_4, Y_2$
Retention, HOCS	Y_5	X_0, X_2, X_3, X_4, Y_3

2.6. Hypotheses

Referring to Table 1, six null hypotheses were formulated:

H_{00} : There is no difference in the total relative learning between the groups.

H_{01} : There is no difference in the HOCS relative learning between the groups.

H_{02} : There is no difference in the total absolute learning between the groups.

H_{03} : There is no difference in the HOCS absolute learning between the groups.

H_{04} : There is no difference in the total retention of knowledge between the groups.

H_{05} : There is no difference in the HOCS retention of knowledge between the groups.

2.7. Proposed Analysis

From Table 1, it is obvious that the dependent variables depend on multiple variables and hence ‘multiple-way’ ANOVA analysis will have to be applied in order to get an indication of any significant differences. To this end, all variables will be defined in SPSS¹ and a ‘univariate, general linear model’ will be applied. If any of the null hypotheses are rejected, we will perform a multiple comparison analysis to establish which groups are significantly different; this will be performed using Fisher’s (improved) LSD test (least significant difference), including either Holm [56] and/or Dunn [57] compensated P -values or a simple Tukey test [58, p. 555]. If data does not follow a normal distribution, or if too few samples are collected, non-parametric methods will be applied.

We will also analyze individuals’ improvements by analyzing the correlation between the *quality* of QTL/WTL works and post-test/exam/retention scores.

3. Results

3.1. Screening

The students’ participation in these extra-curricular activities was voluntary and attendance rates were less than expected. In order to represent a ‘valid sample’ for the QTL, WTL or REF group, the following criteria were established:

1. The student must have attended at least two extra-curricular classes.
2. The student must have attended both the pre- and post-test.

In order to represent a valid sample for the NULL group the student must have attended both the pre- and post-test sessions but not attended any of the extra-curricular activities. Also, students must have provided information about their lecture attendance and their Time-on-Task (outside the classroom).

Hence, most students represented invalid samples because they either only attended one extra-curricular session or they did not take both the pre- and post-test. One student attended multiple extra-curricular sessions and was therefore excluded. Two students did not provide information about their lecture attendance and time-on-task and were also excluded.

After the screening stage, a total of 43 students remained. Table 2 summarizes their scores on the dependent variables $Y_0 - Y_3$. The retention test results ($Y_4 - Y_5$) had to be excluded from the experiment since only five of the 43 students took the retention test.

Table 2. Average and standard deviation of outcome variables.

Outcomes, Y_i	QTL ($n=6$)		WTL ($n=11$)		REF ($n=4$)		QTL ($n=22$)	
	Av.	SD	Av.	SD	Av.	SD	Av.	SD
Y_0	6.2	1.3	5.5	3.4	7.1	1.8	4.2	3.6
Y_1	2.2	1.6	2.0	2.1	2.1	.25	1.8	1.2
Y_2	10.0	3.4	11.0	3.4	11.1	2.2	10.6	4.2
Y_3	5.0	1.4	5.4	1.0	4.7	1.3	5.3	1.9

Figures 3-6 illustrate the box plots of the four outcome variables $Y_0 - Y_3$; the boxes represent the middle 50% and the whiskers represent the top and bottom 25%.

3.2. Testing for Normality

First, data was investigated for normality. Figure 7 illustrates the histograms of the relative gain results for the four groups. (The histograms of the other output variables were very similar.) Due to the low number of samples in some of the groups, normality is hard to infer. Figure 8 illustrates a P-P plot of the pre-test gain for the *merged* sample of all groups. The meandering of data around the straight line indicates a minor *skewness* problem but a 1-sample Kolmogorov-Smirnov test of the merged data produced a $p=.083$ which would suggest that data is normally distributed but since it is just about not significant ($\alpha=.05$), both parametric and non-parametric methods will be applied. Also, the low sample size indicates that non-parametric methods should be applied [58, pp. 283-284]. However, both parametric and nonparametric methods will be applied for comparison.

¹ SPSS is a statistics analysis tool provided by IBM and the most common software tool for hypothesis testing.

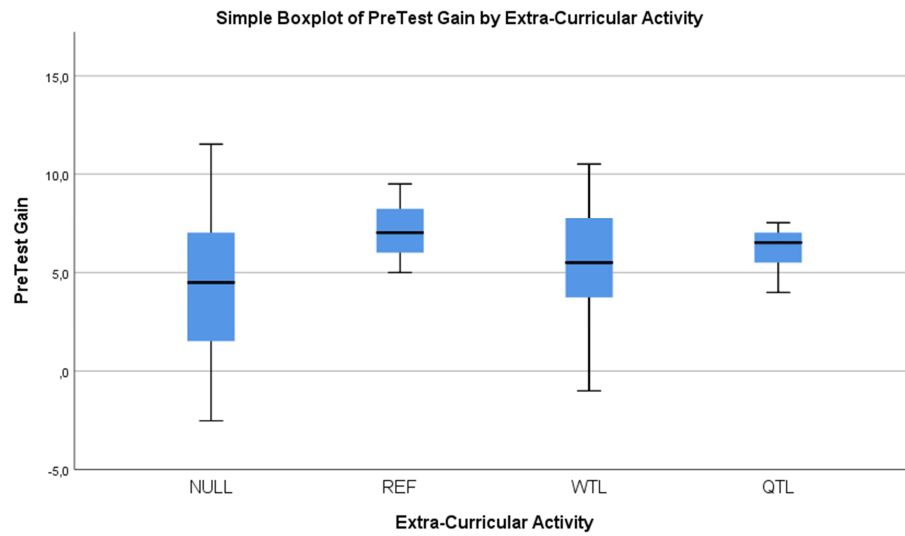


Figure 3. Box plot of relative gain.

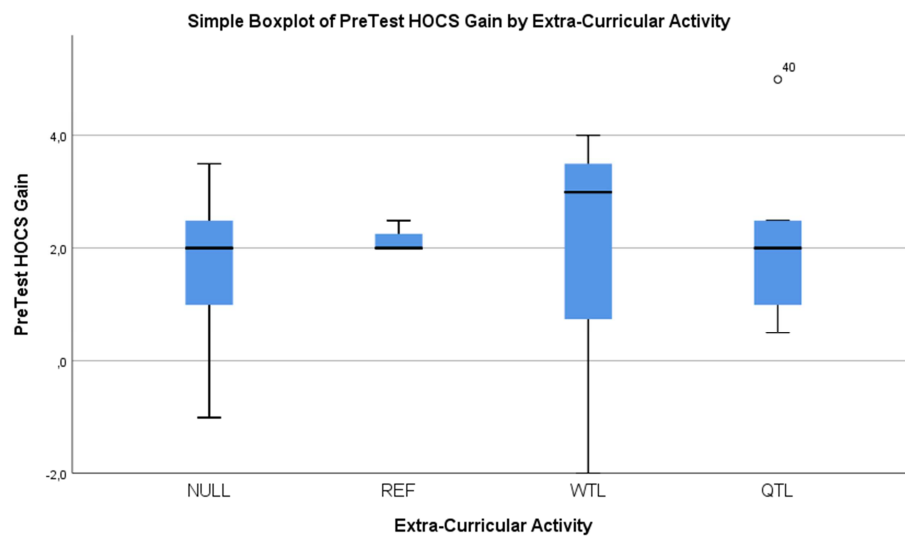


Figure 4. Box plot of relative gain, HOCS only.

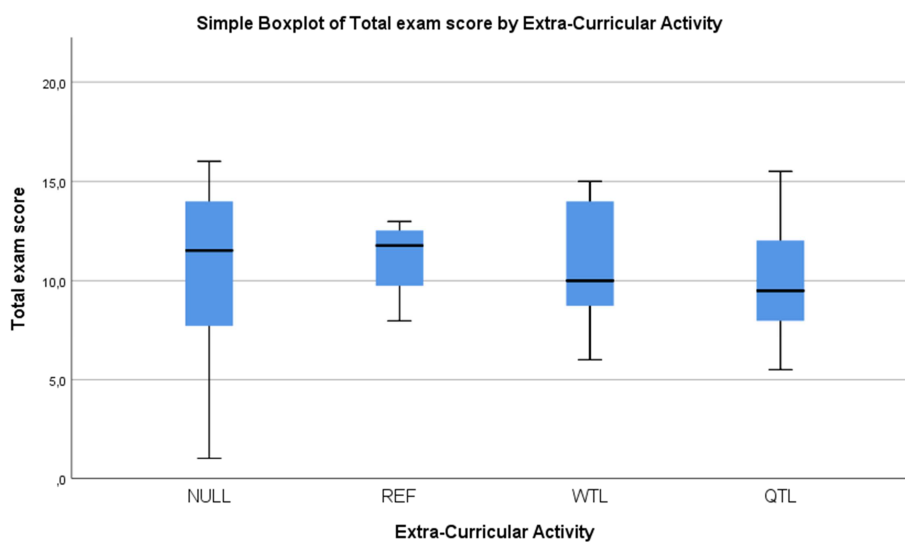


Figure 5. Box plot of total exam score.

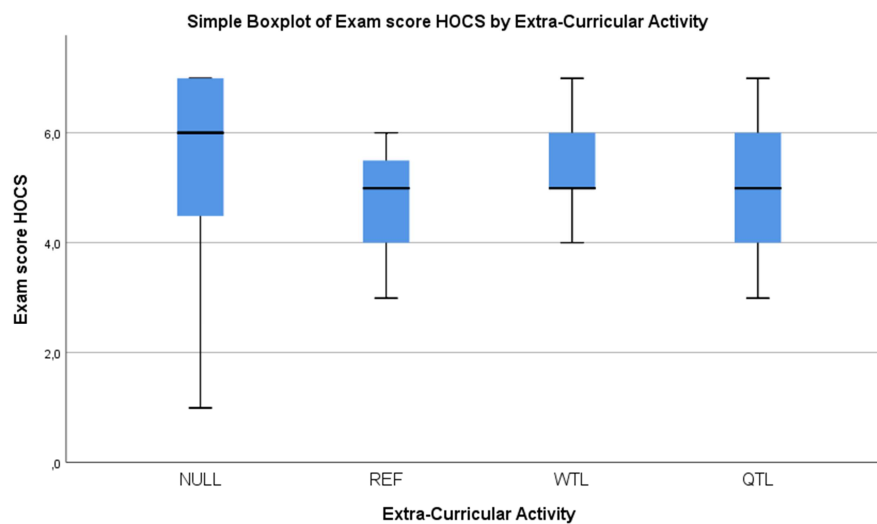
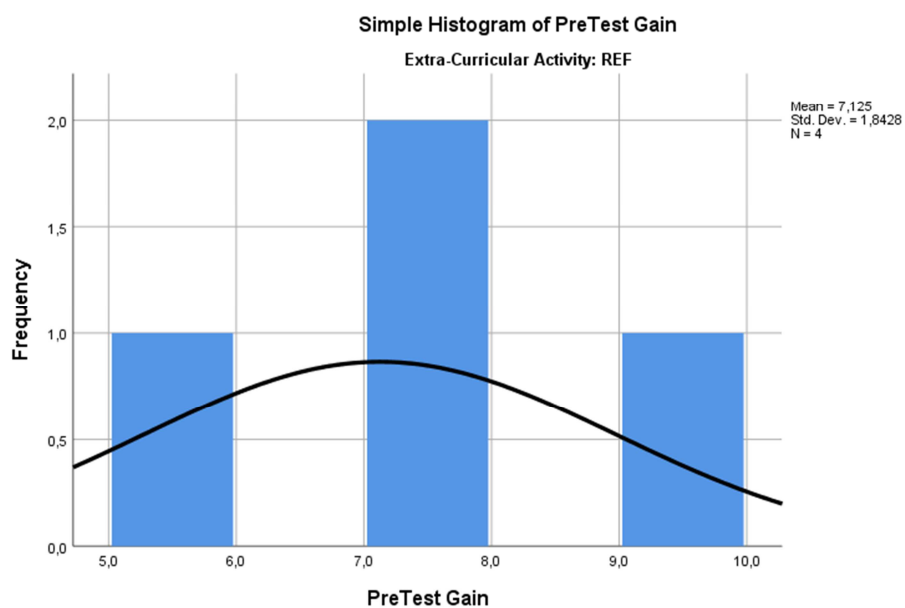
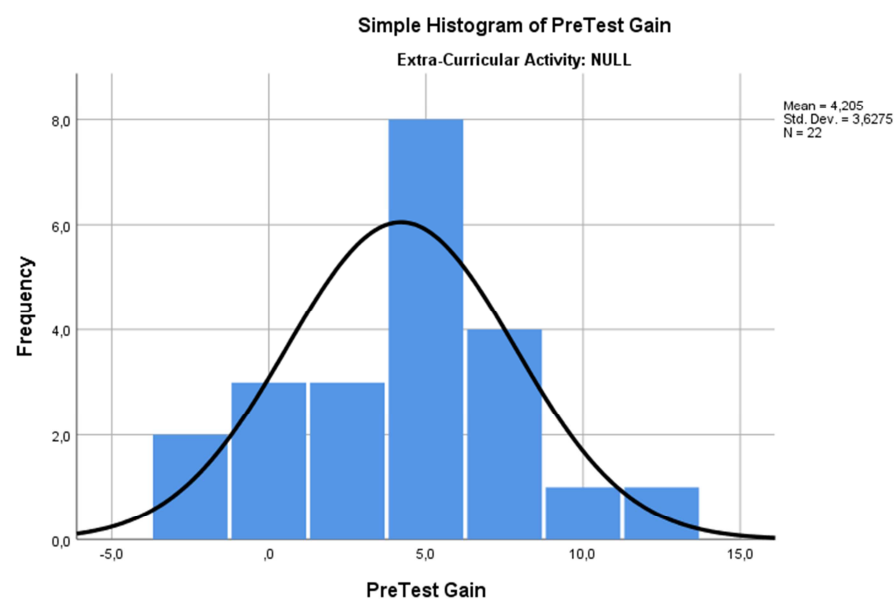


Figure 6. Box plot of exam score, HOCS only.



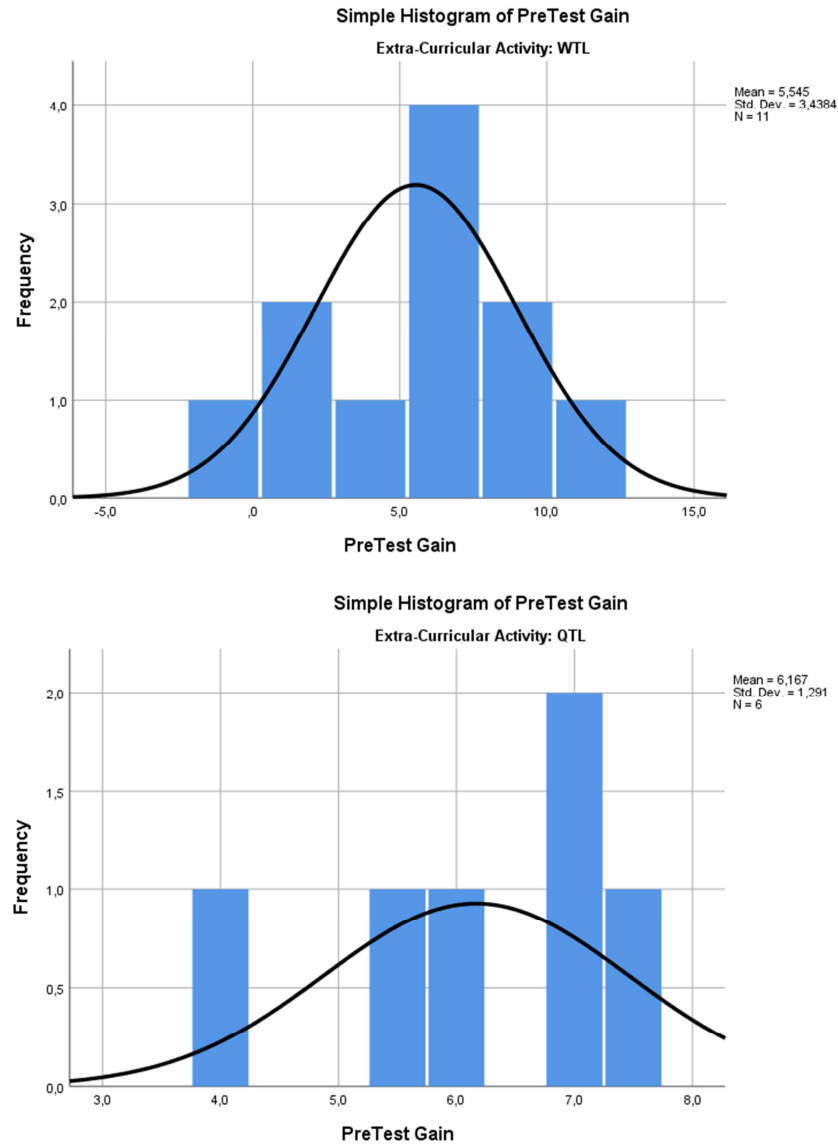


Figure 7. Testing PreTest Gain data for normality.

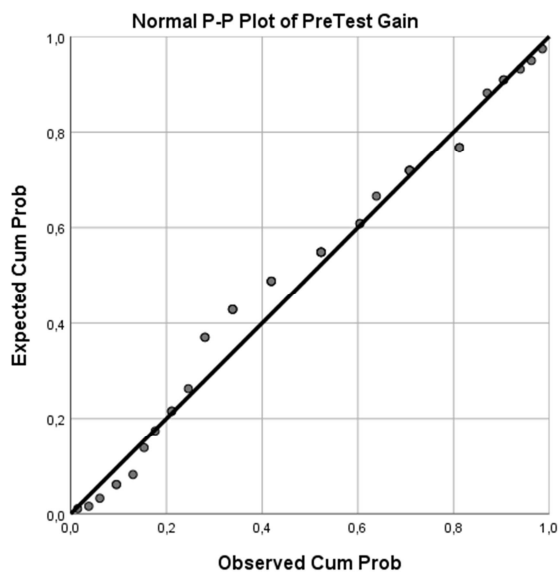


Figure 8. P-P plot of PreTest Gain (all samples).

4. Analysis

4.1. Parametric Methods

First, a 1-way ANOVA analysis was performed on output variables $Y_0 - Y_3$ with respect to independent variables $X_0 - X_3$. The result is illustrated in Table 3; no scores are significantly different. Table 4 reports the 2-way ANOVA between the extra-curricular activity (X_0) and the other independent variables ($X_1 - X_4$) for the different Y_i s. Since no significances were found, no post-hoc multi-comparison analyses were applied.

Table 3. 1-Way ANOVA analysis: Y_i versus extra-curricular activity (X_0).

Dependent variable	F	Sig.
Y_0	.114	.952
Y_1	.825	.488
Y_2	.124	.945
Y_3	.221	.881

Table 4. 2-Way ANOVA analysis: Y_i versus extra-curricular activity and X_1 - X_4 .

Dependent variable	$X_0 \times X_1$		$X_0 \times X_2$		$X_0 \times X_3$		$X_0 \times X_4$	
	F	Sig.	F	Sig.	F	Sig.	F	Sig.
Y_0	.991	.490	.679	.705	.591	.735	.733	.605
Y_1	1.109	.415	.972	.480	1.065	.406	.672	.648
Y_2	.611	.780	.371	.925	.242	.959	.522	.757
Y_3	.661	.741	.460	.871	.294	.935	.538	.746

4.2. Non-parametric Methods

A Kruskal-Wallis 1-Way ANOVA test was applied to all the dependent variables (Y_0 - Y_3) with the extra-curricular

activity as the independent variable (X_0). The result is illustrated in Table 5. No significant differences were reported across categories of extra-curricular activities.

Table 5. Kruskal-Wallis, 1-Way ANOVA.

Null Hypothesis		Sig.	Decision
H_{00}	There is no difference in the total relative learning across categories of extra-curricular activities	.231	Retain the null hypothesis
H_{01}	There is no difference in the HOCS relative learning across categories of extra-curricular activities	.763	Retain the null hypothesis
H_{02}	There is no difference in the total absolute learning across categories of extra-curricular activities	.931	Retain the null hypothesis
H_{03}	There is no difference in the HOCS absolute learning across categories of extra-curricular activities	.580	Retain the null hypothesis

4.3. Performance Predictors

We also investigated the possibility of using the quality of the QTL and WTL students' writing assignments as a predictor for $Y_0 - Y_3$; the average score on their QTL/WTL assignments was plotted against their performance scores and a linear graph was fitted to the data. Results show that there was no significant correlation between QTL/WTL scores and their academic performance.

5. Conclusions

From the graphs in Figures 3-8, it can be concluded that the extra-curricular activities implemented in this experiment did not have any significant impact on the students' learning; not in terms of relative or absolute knowledge, regardless of whether the total learning is considered or if only HOCS questions are considered. None of the six initial null hypotheses can be rejected. Also, this work cannot conclude that the quality of their extra-curricular writing assignments is a valid predictor of their exam performance.

6. Discussion

The most interesting results of this study is the NULL groups' performance. Even if this group is not a random selection, it is generally accepted that the students that don't participate in any extra-curricular activities are the low achievers. The results of this work do not support that hypothesis; no significant differences between participants and non-participants were observed.

The fact that none of the hypotheses could be rejected could suggest a support for Bangert-Drown, Hurley & Wilkinsons' [16] contention that WAC activities may be a waste of students' time. There could, however, be other explanations for the absence of significant results. First of all,

the 'dosage' was too small; each group was only scheduled for one extra-curricular event each week, a total of five occasions, and few students attended more than one occasion. Another reason for the meagre outcome could be that the low dosage effects reported by Almer, Jones & Moeckel [51] do not apply to STEM teaching (their work was in accounting). Quitadamo & Kurtz [10] proved that effects could be detected in only a few weeks (in a biology class) but they were specifically targeting critical thinking.

However, the main source of uncertainty in this work is of course the low number of samples in the four groups which makes any analysis and conclusions precarious.

7. Suggestions

Since the main issue in this work was the low participation rate, a new study should focus on stimulating the attendance rates. Extra-curricular classes should be mandatory or some other incentives for participation must be implemented. Also, the logistics and scheduling should be considered. Some of our extra-curricular classes were scheduled early in the morning or late in the afternoon and in classrooms geographically far away from their normal domains.

In order to improve the quality of the QTL assignments, we also suggest that students should not design the exam problems in the classroom, but rather design them at home and bring them ready-made to the QTL classroom. That would save classroom time and probably increase the quality of the problems (and their learning).

Author Note

This work was conducted at the department of Physics, University of Gothenburg, during the fall of 2019 with permission of the department head. The author would like to acknowledge Jonas Enger, Carlo Roberto and Andreas Johansson for their help collecting the data.

The author has no other conflicts of interest to report.

References

- [1] A. N. Applebee, "Writing and reasoning," *Review of Educational Research*, vol. 54, no. 4, pp. 577-596, 1984.
- [2] J. Brown and E. C. Stephens, "Writing as transformation: A new paradigm for content writing," *Clearing House*, vol. 69, pp. 14-16, 1995.
- [3] J. Reynolds, C. Thaiss, W. Katkin and R. Thompson Jr, "Writing-to-learn in undergraduate science education: A community-based, conceptually driven approach," *Life Science Education*, vol. 11, pp. 17-25, 2012.
- [4] P. Rowell, "Learning in school science: The promise and practices of writing," *Studies in Science Education*, vol. 30, pp. 19-56, 1997.
- [5] L. Yore, B. Hand and V. Prain, "Writing-to-learn science: Breakthroughs, barriers and promises," in *The International Conference of the Association for Educating Teachers in Science*, Austin, TX, 1999.
- [6] D. Bensley and C. Haynes, "The acquisition of general purpose strategic knowledge for argumentation," *Teaching of Psychology*, vol. 22, pp. 41-45, 1995.
- [7] J. Kovac and D. Sherwood, "Writing in chemistry: An effective learning tool," *Journal of Chemical Education*, vol. 76, no. 10, pp. 1399-1403, 1999.
- [8] J. Kovac and D. Sherwood, *Writing across the chemistry curriculum*, 1 ed., Upper Saddle River, NJ: Prentice Hall, 2001.
- [9] M. T. Oliver-Hoyo, "Designing a written assignment to promote the use of critical thinking skills in an introductory chemistry course," *Journal of Chemical Education*, pp. 899-903, 2003.
- [10] I. Quitadamo and M. Kurtz, "Learning to Improve: Using writing to increase critical thinking performance in general education biology," *Life Science Education*, vol. 6, pp. 140-154, 2007.
- [11] N. VanOrden, "Critical-thinking writing assignments in general chemistry," *Journal of Chemical Education*, pp. 506-507, 1987.
- [12] C. Wade, "Using writing to develop and assess critical thinking," *Teaching of Psychology*, vol. 22, pp. 24-28, 1995.
- [13] J. Waller, "Philosophies of psychology: A discovery process for undergraduates," *Teaching of Psychology*, vol. 21, pp. 33-35, 1995.
- [14] L. Welch, S. Gradin and K. Sandell, "Enhancing engineering education with writing-to-learn and cooperative learning: Experiences from a software engineering course," in *Proceedings of the 2002 American Society for Engineering Education Conference & Exposition*, Montreal, 2002.
- [15] M. Balgopal, A. Wallace and S. Dahlberg, "Writing to learn ecology: A study of three populations of college students," *Environmental Education Research*, vol. 18, no. 1, pp. 67-90, 2012.
- [16] R. Bangert-Drowns, M. Hurley and B. Wilkinson, "The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis," *Review of Educational Research*, vol. 74, no. 1, pp. 29-58, 2004.
- [17] J. Rich, J. Miller and L. DeTora, "From concept to application: Student narratives of problem-solving as a basis for writing assignments in science classes," *Across the Disciplines*, vol. 8, no. 1, pp. 1-14, 2011.
- [18] D. Drabick, R. Weisberg, L. Paul and J. Bubier, "Keeping it short and sweet: Brief, ungraded writing assignments facilitate learning," *Teaching of Psychology*, vol. 34, no. 3, pp. 172-176, 2007.
- [19] A. Cooper, "Today's technologies enhance writing in mathematics," *The Clearing House*, vol. 85, pp. 80-85, 2012.
- [20] L. S. Vygotsky, *Studies in communication. Thought and Language*, E. Hanfmann and G. Vakar, Eds., Cambridge: MIT Press, 1962.
- [21] K. Bargate, "Effects of a writing-to-learn tutorial programme in a higher education setting: An interactive qualitative study," *Africa Education Review*, vol. 12, no. 3, pp. 493-507, 2015.
- [22] J. Emig, "Writing as a mode of learning," *College Composition and Communication*, vol. 28, no. 2, pp. 122-128, 1977.
- [23] S. M. Glynn and K. Muth, "Reading and writing to learn science: Achieving scientific literacy," *Journal of Research in Science Teaching*, vol. 31, no. 9, pp. 1057-1073, 1999.
- [24] W. Holliday, L. Yore and D. Alverman, "The reading-science learning-writing connection: Breakthroughs, barriers and promises," *Journal of Research in Science Teaching*, pp. 933-946, 1994.
- [25] S. Zemelman, H. Daniels and A. Hyde, *Best practice: Today's standards for teaching and learning in America's schools*, 3 ed., Portsmouth, NH: Heinemann, 2005.
- [26] L. Mason, "Sharing cognition to construct scientific knowledge in school context: The role of oral and written discourse," *Instructional Science*, vol. 26, pp. 359-389, 1998.
- [27] D. Sanders, "Writing activities can improve learning in computer science courses," *Computer Science Education*, vol. 2, no. 2, pp. 171-181, 2006.
- [28] H. Christensen and A. Rasmussen, "Writing to learn in the natural science: Does source material matter," *Research and Teaching*, vol. 47, no. 1, pp. 57-65, 2017.
- [29] L. Rivard, "Are language-based activities in science effective for all students, including low achievers," *Science Education*, vol. 88, pp. 420-442, 2004.
- [30] H.-Y. Lo, G.-Z. Liu and T. Wang, "Learning how to write effectively for academic journals: A case study investigating the design and development of a genre-based writing tutorial system," *Computers & Education*, vol. 78, pp. 250-267, 2014.
- [31] P. Klein, "Constructing scientific explanations through writing," *Instructional Science*, pp. 191-231, 2004.
- [32] J. Britton, *Shaping at the point of utterance*, G. Pradl, Ed., Monclair, NJ: Boynton/Cook Publishers, Inc, 1982, pp. 139-145.
- [33] P. Klein, "Reopening inquiry into cognitive processes in writing-to-learn," *Educational Psychology Review*, vol. 11, pp. 203-270, 1999.

- [34] J. Nevid, A. Pastva and N. McClelland, "Writing-to-learn assignments in introductory psychology: Is there a learning benefit?," *Teaching of Psychology*, vol. 39, no. 4, pp. 272-275, 2012.
- [35] P. A. Mueller and D. M. Oppenheimer, "The pen is mightier than the keyboard: Advantages of longhand over laptop note taking," *Psychological Science*, vol. 25, pp. 1159-1168, 2014.
- [36] J. Bean, *Engaging ideas: The professor's guide to integrating writing, critical thinking and active learning in the classroom*, 2 ed., San Francisco: Jossey-Bass, 2011.
- [37] L. Andersson, D. Krathwhol, K. Cruikshank, R. Mayer, P. Pintrich and M. Wittrock, *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, London: Pearson, 2001.
- [38] P. Papadopoulos, S. Demetriadis, I. Stamelos and I. Tsoukalas, "The value of writing-to-learn when using question prompts to support web-based learning in ill-structured domains," *Educational Technology Research and Development*, vol. 59, pp. 71-90, 2011.
- [39] N. Armstrong, C. Wallace and S.-M. Chang, "Learning from writing in college biology," *Research in Science Education*, pp. 483-499, 2008.
- [40] J. Libarkin and G. Ording, "The utility of writing assignments in undergraduate bioscience," *Life Science Education*, vol. 11, pp. 39-46, 2012.
- [41] D. Sabrio, S. Sabrio and G. Tintera, "Writing to learn and learning to write mathematics: an experiment," *Problems, resources and issues in Mathematics Undergraduate Studies*, vol. 3, no. 4, pp. 419-429, 2007.
- [42] K. Arnold, S. Umanath, K. Thio, W. Reilly, M. McDaniel and J. Marsh, "Understanding the cognitive processes involved in writing to learn," *Journal of Experimental Psychology: Applied*, vol. 23, no. 2, pp. 115-127, 2017.
- [43] D. Chamely, J. Haky and J. Galin, "From Bhopal to cold fusion: A case-study approach to writing assignments in honors general chemistry," *Journal of Chemical Education*, vol. 89, pp. 502-508, 2012.
- [44] M. Balgopal and A. Wallace, "Decisions and dilemmas: Using writing to learn activities to increase ecological literacy," *The Journal of Environmental Research*, vol. 40, no. 3, pp. 13-26, 2009.
- [45] National Science Education Standards, "Scientific Literacy," 10 October 2009. [Online]. Available: <http://www.literacynet.org/science/scientificliteracy.html>.
- [46] J. Britton, *Language and learning*, London: Allen Lane, 1970.
- [47] J. Britton, *Writing to learn and learning to write*, Urbana, IL: National Council of Teachers of English, 1972.
- [48] C. Griffin, "Using writing to teach many disciplines," *Improving College and University Teaching*, vol. 31, no. 3, pp. 121-128, 1983.
- [49] M. Balgopal and A. Wallace, "Writing-to-learn, writing-to-communicate & scientific literacy," *The American Biology Teacher*, vol. 75, no. 3, pp. 170-175, 2013.
- [50] C. Moskovitz and D. Kellogg, "Inquiry-based writing in the laboratory course," *Science Education*, pp. 919-920, 2011.
- [51] E. Almer, K. Jones and C. Moeckel, "The impact of one-minute papers on learning in an introductory accounting course," *Issues in Accounting Education*, vol. 13, no. 3, pp. 485-497, 1998.
- [52] J. Hart, "Top 200 Tools for Learning 2019," Centre for Learning & Performance Technologies, [Online]. Available: <https://www.toptools4learning.com/>. [Accessed 18 May 2020].
- [53] E. F. Redish, *Teaching Physics*, Danvers, MA: John Wiley & Sons, 2003.
- [54] T. A. Angelo and K. Cross, *Classroom assessment techniques: A handbook for college teachers*, 2 ed., San Francisco: Jossey-Bass, 1993.
- [55] F. MacKinnin-Slaney, "Discovery writing in introductory college student personnel graduate course," *Journal of College Student Development*, vol. 32, pp. 92-94, 1991.
- [56] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, vol. 6, pp. 65-70, 1979.
- [57] O. J. Dunn, "Multiple comparison among means," *Journal of the American Statistical Association*, vol. 56, pp. 52-64, 1961.
- [58] A. Field, *Discovering Statistics Using IBM SPSS Statistics*, London: SAGE Publication Ltd, 2018.