On the Impact of Homework Format in an Elementary Statistics Course

Lisa Kaylor & Bradford Westgate
Alma College*

Abstract

This study compared online and written homework in an Elementary Statistics course with respect to student outcomes and opinions. Two instructors taught two sections each, one with online homework and one with written homework. There were no significant differences between homework formats for grades on exams and quizzes or on an embedded explanation question, although the mean homework grade for sections with online homework was significantly higher (uncorrected for multiple comparisons). Student survey results had a higher mean rating for written homework than online, indicating a more positive opinion towards written homework. Student feedback was synthesized to contextualize these results.

Keywords: statistics education, homework, traditional, web-based

1 Introduction

There are many factors to consider when deciding on a homework format for a mathematics course. Online homework systems allow assignments to be graded automatically, which provides students immediate feedback on the accuracy of their solutions. This format may also allow students multiple attempts at achieving a correct solution. Research has shown that these characteristics may enhance student learning (Butler et al., 2014). In comparison, written assignments require manual grading, a waiting period to receive feedback, and a single attempt at achieving the correct solution, unless resubmissions are permitted. An advantage of written homework is that students gain practice in written communication of concepts that is not possible when submitting only a final answer in an online system. Although feedback is delayed, instructor-graded assignments also provide individualized feedback on process.

The advantages and disadvantages pertaining to homework formats motivated this study, which was conducted in four sections of an Elementary Statistics course at our college. Our interest was on the impact of homework format on quiz and exam grades and specifically on students' ability to answer explanation questions, measured by grades on a final exam question. Student opinions on format were obtained through a survey tailored to address students' perceived preparedness, based on homework, for computation and explanation questions.

At our small liberal arts college, we offer an average of six sections of Elementary Statistics per year. Each section enrolls 23 students on average. This course satisfies our college's math competency requirement and may be the only math course a student takes at our institution. This motivates our desire to ensure students have a positive, successful academic experience

*The two co-authors contributed equally to this study.

while improving their numerical literacy and ability to interpret mathematics. Course content includes sampling and experimental design, descriptive statistics, linear regression, and statistical inference.

All sections of Elementary Statistics use the textbook *OpenIntro Statistics* (Diez et al., 2019). This textbook can be used with the online homework platform *MyOpenMath* (Lippman et al., 2025). This platform is free to both the student and institution with no additional implicit costs. It is of high importance to minimize costs, which led to our use of both this textbook and homework system. Previous sections of this course have used various homework formats, including textbook problems, instructor-written problems, and problems via the *MyOpenMath* system. For this study, we focused on online homework and instructor-written problems submitted on paper.

1.1 Literature Review

There is extensive research on the effect of homework on students' academic success. It has been observed that graded homework generally has a positive impact on academic success and that the positive impact increases as students move up in grade level (Bas et al., 2017; Cooper et al., 2006; Hattie, 2008). The impact in math courses has been shown to be particularly strong when aspects such as repeated practice and timely, quality feedback are provided (Butler et al., 2014). The advent of alternative homework formats has led to studies on the impact of these formats compared to the traditional pencil and paper method in a variety of disciplines (Cox and Singer, 2011; Gok, 2011; Hauk et al., 2015; Kashyap and Mathew, 2022; Kodippili and Senaratne, 2008; Quinlan, 2023).

The literature on homework format is inconclusive with respect to which format leads to more successful student outcomes. Magalhães et al. (2020) reported mixed results when considering student outcomes in a review of studies comparing online and traditional homework. Eleven of the 31 studies included were conducted within undergraduate mathematics courses. Of these studies, four studies found that online homework had better results than traditional, five studies had neutral results (no difference between formats), and two studies had mixed results. Another meta-analysis by Bishop et al. (2024) observed similarly mixed results to Magalhães et al. (2020). Of the 18 studies considered, 11 studies reported a more beneficial impact on academic success via online homework, while seven studies reported the analogous result for traditional homework.

Although these meta-analyses were conducted within the last five years, the majority of studies considered were significantly older, took place at large institutions, and primarily examined homework format's impact on grades. As educational technologies continue to advance, it remains valuable to examine their impacts in the classroom. This sentiment was expressed by Quinlan (2023) in his study on homework format in a first-semester calculus course. In this study, we investigated not only academic success as measured by overall grades, but also student preparation for explanation questions, which we do not find in the literature.

2 Methods

This study was performed in Elementary Statistics sections taught by the two co-authors during the Fall 2024 semester. Each instructor taught two sections of Elementary Statistics. One of these sections used online assignments via the MyOpenMath platform and the other used instructor-written assignments that the students submitted on paper. The assignments were common amongst the two sections using each homework format. Students were not given access by instructors to assignments in the alternative format. Instructor A's section

with written assignments was taught in the morning and their section with online assignments was taught in the afternoon, while the reverse was true for Instructor B's sections. This choice was made to mitigate the impact of class meeting time on results.

2.1 Homework Formats

There were 10 homework assignments throughout the semester. Online assignments were created using problems from the *MyOpenMath* content library. For written assignments, instructor-written problems were supplemented by problems taken or adapted from the textbook and *MyOpenMath* platform. Problem types such as numerical answer, multiple-choice, true/false, matching, short answer, and creating graphical displays were included in both formats. However, each homework format was used in a manner that is natural to it: online assignments used more problems that could be graded automatically, such as multiple-choice and true/false, while written assignments included more short answer problems requiring students to show work and write explanations.

The online homework assignments were longer than the written homework, with a mean of 12.7 problems per homework compared to a mean of 5.6 problems for the written homework; however, several problems per assignment were multi-part for both formats. *MyOpenMath* allows the instructor to choose the number of attempts students have on a problem. We allowed students unlimited attempts, no matter the problem type.

Written homework assignments were graded according to a common rubric by the instructor of the respective section. Online homework assignments were graded automatically by the *MyOpenMath* platform, which gave feedback on correctness without detailed solutions.

2.2 Response Variables

Student outcomes were assessed by eight quizzes, two preliminary exams, and a cumulative final exam. All quizzes and exams were common amongst the four sections of the course and given in the traditional pencil and paper format. Instructors graded their own section's homework, quizzes, and exams according to a common rubric. The quizzes accounted for 15% of the final course grade, while the preliminary exams and final exam accounted for 20% each. The remaining 25% of the course grade came from the homework assignments (20%) and reflection surveys following the exams (5%).

Four response variables related to student outcomes were evaluated in this study: the 75% of the overall course grade from quizzes and exams, the final exam grade, the grade on a particular explanation problem on the final exam dealing with correlation vs. causation, and the overall homework grade. The explanation question was graded on a 0-3 scale. The text of this question was the following: "An educator wants to determine how effective tutoring is in raising students' grades in her class, so she offers free tutoring for those who want it. She finds that there is a positive correlation between number of tutoring sessions attended and final exam scores. Can the instructor conclude that tutoring causes an increase final exam scores? Explain your answer."

Statistical analysis for these response variables was performed using MANCOVA and follow-up ANCOVA, with the homework format and instructor as factors and the class year (1 for first-year students, 2 for sophomores, 3 for juniors, and 4 for seniors) as a covariate. Since the students were not randomly assigned to the class sections, including the class year helped to protect against differences between students registering for each of the sections. Permutation-based MANCOVA and ANCOVA were also performed, using the 'vegan' (specifically, the adonis2 function with the Mahalanobis distance) and 'permuco' packages in R (R Core Team, 2025).

Additionally, an anonymous survey was given to students at the end of the semester to assess students' views on the efficacy of their homework format and their satisfaction with the format. Of the 95 students in the four sections, 85 completed the survey. Students were asked to respond with "Strongly agree", "Agree", "Neutral", "Disagree", or "Strongly disagree" to the following statements:

(S1) "Overall, the homeworks in this class helped me learn the material."

(S2) "Overall, the homeworks helped me prepare for the exams."

(S3) "The homeworks helped me prepare for computational questions on the exams/quizzes."

(S4) "The homeworks helped me prepare for explanation questions on the exams/quizzes."

(S5) "Overall, I was satisfied with the homework format for this class."

Each of these statements was recorded as a response variable on a Likert scale, ranging from 5 for "Strongly agree" to 1 for "Strongly disagree". The survey also included a written question: "Please share any other thoughts you have about the homework for this course." Selected responses to this question are given in Section 4.

A composite score was calculated by taking the mean of the five statements. Statistical analysis for the composite score was performed using ANOVA, with the homework format and instructor as factors. Permutation-based ANOVA was also performed. All statistical analyses were performed in R, using the 'car', 'effectsize', 'MVN', 'permuco', 'psych', and 'vegan' packages.

3 Results

Table 1 shows means and standard deviations for the response variables related to student outcomes. Via MANCOVA, the homework format was statistically significant for the response variables (Pillai's trace = 0.177, p = 0.0018). The instructor was not significant (p = 0.085), the interaction effects between the homework format and instructor were not significant (p = 0.50), and the class year was not significant (p = 0.103).

Via univariate ANCOVA, the homework format was significant at the 5% level for the homework grades (p = 0.017). We interpret this result cautiously, because Bonferroni's correction reduces the 5% significance level to 1.25%, after which it is not significant. Partial eta-squared equaled 0.06 (90% confidence interval $[0.01, 0.16]$), indicating a medium effect size. The homework format was not significant for the quiz and exam grades (p = 0.59), final exam grades (p = 0.77), and explanation question grades (p = 0.75).

The MANCOVA residuals exhibited severe non-normality (Royston's test $p < 0.0001$). For comparison, permutation-based MANCOVA and ANCOVA tests were also performed. These tests gave similar p-values and reached the same conclusions for statistical significance as the previous tests.

Table 1. *Means and standard deviations (SD) for the four response variables related to student outcomes. The explanation question was graded on a scale from* 0 – 3*, while the other three variables were graded from* 0 – 100*.*

		Written		Online	
		A	B	A	B
Sample size		23	22	24	26
Course grade (quizzes/exams)	Mean	77.8	78.2	74.5	77.3
	SD	18.2	15.8	17.4	15.2
Final exam grade	Mean	73.1	74.1	67.7	76.3
	SD	24.0	18.6	21.3	17.5
Explanation question grade	Mean	2.35	2.14	2.23	2.12
	SD	0.68	0.99	0.86	0.95
Homework grade	Mean	81.3	80.8	88.8	88.4
	SD	16.4	15.6	11.3	18.7

Table 2 shows means for the five statements on the survey. Exploratory factor analysis strongly supported unidimensionality for the statements. Cronbach's alpha equaled 0.86, indicating high internal consistency. Therefore, statistical analysis was performed on the composite score formed by taking the mean of the five statements.

Table 2. *Mean responses for the five statements on the reflections survey. Each response variable was measured on a Likert scale from 1–5. The statements were aggregated into a composite score by taking their mean.*

	Written		Online	
Statement	A	B	A	B
Sample size	18	21	21	25
S1: Overall, homework helped	4.56	4.29	4.33	4.16
S2: Homework helped for exams	4.39	4.38	4.14	3.96
S3: Homework helped for computations	4.61	4.38	4.24	3.96
S4: Homework helped for explanations	4.39	4.33	3.95	3.64
S5: Overall, satisfied with homework	4.28	4.38	4.14	3.88
Composite score	4.44	4.35	4.16	3.92

Via two-factor ANOVA, the homework format was significant at the 1% level for the composite score ($p = 0.006$), while the instructor and interactions between the homework format and instructor were not significant ($p = 0.18$ and $p = 0.56$, respectively). The mean of the students' responses was significantly higher for the written homework sections than for the online homework sections. Partial eta-squared equaled 0.09 for the homework format (90% confidence interval [0.02, 0.20]), indicating a medium effect size. Because the residuals of the ANOVA exhibited non-normality (Shapiro-Wilk $p = 0.016$), a permutation-based ANOVA was also performed, which gave similar p-values and reached the same conclusions for statistical significance.

4 Discussion

This study showed no statistically significant differences in outcomes for students assigned online and written homework, measured by grades on quizzes and exams and the final exam. This result aligns with the literature, which indicates that online homework is, at minimum, as effective as written homework with respect to student outcomes. Moreover, there was no significant difference between homework formats on the final exam explanation question, even though students with the online homework format had less preparation for writing explanations. The mean grade for online homework was significantly higher than for written homework (uncorrected for multiple comparisons), which we attribute primarily to the unlimited attempts allowed. A higher online homework average has also been noted in previous studies (Cox and Singer, 2011; Lunsford and Pendergrass, 2016).

Survey results indicated that students were more positive towards written homework. A composite score formed by taking the mean of the five statements on the survey showed significantly higher means in the written homework sections than in the online homework sections.

In the written question on the survey, students from the online homework sections appreciated the ability to submit multiple attempts, which aligns with the preferences found in a study by Cox and Singer (2011). These are example student responses:

> "I liked that we had multiple attempts on the questions in the homework. It took a lot of stress off of doing the homework and helped me concentrate and learn better."
>
> "I really loved how you gave us so many attempts at each question! Although I always tried my best on the homework, there were some parts I just couldn't get right away."

A disadvantage of online homework highlighted in survey responses was the tolerance intervals for answers on numerical response problems. Some students expressed frustration that they performed the correct calculations but were not given credit by the platform for the correct answer. Example responses include:

> "I had trouble getting a good grade when the rounding was sometimes off."
>
> "The online homework was very frustrating at times, I thought some of the questions were extremely long and if you didn't have the calculations perfect you often wouldn't get it right."

A specific example of this issue involves calculations with the Normal distribution. If the student uses a different method for obtaining values from the Normal distribution than is coded in the online system, such as a table in the textbook, the student may not be able to obtain an accepted answer. We addressed this issue in an ad-hoc way, by manually giving credit to answers off due to rounding. A study by Williams (2012) cites rounding tolerance as an issue impacting homework scores. Similar frustrations were expressed by students when changes to an online homework system resulted in difficulty in getting better grades in a study (Jonsdottir et al., 2017).

The higher mean composite score in the written homework sections may also be due to the greater similarity between the format of written homework and exams. The following response from the written homework sections highlights this similarity:

> "This type of homework is my personal favorite for math as the questions were very much like the exam questions so I felt prepared."

In contrast, two responses from the online homework sections highlight the difference between formats:

> "I think an improvement for the homework's would be to have similar questions that we go over in class because some questions asked questions in different formats from the class work and quizzes, which confused me at times."
>
> "I think I would have liked more practice with writing conclusion sentences but overall I enjoyed the homework."

An article by Lundsford and Pendergrass (2016) provided suggestions for working with the inherent differences noted here between formats.

The lack of random assignment to homework format is a limitation of our study. Efforts to control for course meeting time and class year of students were made, but other potentially influencing factors on student outcomes were not controlled for. One example, which was found to be influential by Palocsay and Stevens (2008), is mathematics background. Since our students come to this course with a wide range of experience in mathematics and statistics, including this variable may prove insightful for outcomes.

In conclusion, this study revealed no significant difference in student outcomes on quizzes and exams, final exam, or an embedded explanation question between homework formats. Students were more positive towards written homework than online homework. We attribute this in part to technical difficulties with the online homework system, but also, to similarities in question format between written homework and quizzes and exams. In the review by Magalhães et al. (2020), authors in nine of the 11 studies conducted in undergraduate mathematics courses recommended online homework. We recommend that instructors using online homework keep problem types as similar as possible between homework and exams, and consider how to increase students' perceived preparedness for exam problems.

5 Disclosures

5.1 Conflicts of Interest

The authors have no relevant financial or non-financial conflicts of interest.

5.2 Generative Artificial Intelligence (GenAI)

Generative AI searches were used during this study for finding reference articles, for recommendations about statistical methods and R code examples, and for questions about grammar and academic style. The authors evaluated and refined all results from AI searches, and take full responsibility for the content of the publication. AI tools used were GPT-4 Turbo (for reference articles), xAI Grok 3 (for statistical methods and grammar), and Microsoft Copilot (for statistical methods and grammar).

5.3 Human Participants

This study received approval from Alma College's Institutional Review Board (IRB), approval number: R_1kebJM2jLmjt8nJ. The IRB determined that informed consent was not required, as the study used accepted pedagogical practices and de-identified data.

References

Bas, G., Senturk, C., & Cigerci, F. (2017). Homework and academic achievement: A meta-analytic review of research. *Issues in Educational Research, 27*(1), 31–50. https://www.iier.org.au/iier27/bas.pdf

Bishop, R., Best, T., & Horsley, M. (2024). Comparison of academic achievement between online vs. traditional homework: A meta-analysis. *e-Journal of Business Education and Scholarship of Teaching, 18*(1), 8–20. https://eric.ed.gov/?id=EJ1428382

Butler, A., Marsh, E., Slavinsky, J., & Baraniuk, R. (2014). Integrating cognitive science and technology improves learning in a STEM classroom. *Educational Psychology Review, 26*, 331–340. https://doi.org/10.1007/s10648-014-9256-4

Cooper, H., Robinson, J., & Patall, E. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research, 76*(1), 1–62. https://doi.org/10.3102/00346543076001001

Cox, T., & Singer, S. (2011). Taking the work out of homework. *Mathematics Teacher*, 104(7), 514–519. https://doi.org/10.5951/MT.104.7.0514

Diez, D., Cetinkaya-Rundel, M., & Barr, C. (2019). *OpenIntro Statistics* (3rd ed.). OpenIntro. https://www.openintro.org/book/os/

Gok, T. (2011). Comparison of student performance using web-and paper-based homework in large enrollment introductory physics courses. *International Journal of the Physical Sciences 6*(15), 3740–3746. https://academicjournals.org/article/article1380721731_Gok.pdf

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge. https://doi.org/10.4324/9780203887332

Hauk, S., Powers, R., & Segalla, A. (2015). A comparison of web-based and paper-and-pencil homework on student performance in college algebra. *PRIMUS, 25*(1), 61–79. https://doi.org/10.1080/10511970.2014.906006

Jonsdottir, A., Bjornsdottir, A., & Stefansson, G. (2017). Difference in learning among students doing pen-and-paper homework compared to web-based homework in an introductory statistics course. *Journal of Statistics Education, 25*(1), 12–20. https://doi.org/10.1080/10691898.2017.1291289

Kashyap, U., & Mathew, S. (2022). Comparison of students' performance using web-based and paper-based homework in an undergraduate statistics course. *Journal of Education and Practice, 13*(27), 26–31. https://doi.org/10.7176/JEP/13-27-09

Kodippili, A., & Senaratne, D. (2008). Is computer-generated interactive mathematics homework more effective than traditional instructor-graded homework? *British Journal of Educational Technology, 39*(5), 928–932. https://doi.org/10.1111/j.1467-8535.2007.00794.x

Lippman, D., et al. (2025). MyOpenMath: An online course management and assessment system. https://www.myopenmath.com

Lunsford, M., & Pendergrass, M. (2016). Making online homework work. *PRIMUS, 26*(6), 531–544. https://doi.org/10.1080/10511970.2015.1110219

Magalhães, P., Ferreira, D., Cunha, J., & Rosário, P. (2020). Online vs. traditional homework: A systematic review on the benefits to students' performance. *Computers & Education, 152*, 103869. https://doi.org/10.1016/j.compedu.2020.103869

Palocsay, S., & Stevens, S. (2008). A study of the effectiveness of web-based homework in teaching undergraduate business statistics. *Decision Sciences Journal of Innovative Education, 6*(2), 213–232. https://doi.org/10.1111/j.1540-4609.2008.00167.x

Quinlan, J. (2023). Efficacy and attitudes towards online homework systems. *Ohio Journal of School Mathematics 95*, 26–31. https://ohiomathjournal.org/index.php/OJSM/article/view/9725

R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Williams, A. (2012). Online homework vs. traditional homework: Statistics anxiety and self-efficacy in an educational statistics course. *Technology Innovations in Statistics Education, 6*(1). https://doi.org/10.5070/T561011683

Lisa Kaylor is an Assistant Professor of Mathematics at Alma College in Alma, Michigan. She earned her Ph.D. from Wesleyan University in 2019 with work focused in number theory. Her research interests have expanded to include teaching innovations and how students learn mathematics.

Bradford Westgate is an Associate Professor of Mathematics at Alma College in Alma, Michigan. He earned a Ph.D. in Operations Research from Cornell University, focusing on applied Bayesian methods. His teaching interests are in probability, statistics, and data science.