May 2017

Effective Online Practice Strategies Using the ASSISTments Framework

Adam Nathan Croteau
Worcester Polytechnic Institute

Jacob Mackenzie Aki
Worcester Polytechnic Institute

Ryan Patrick Walsh
Worcester Polytechnic Institute

Effective Online Practice Strategies Using the ASSISTments Framework

An Interactive Qualifying Project
Submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfilment of the requirements for the
Degree of Bachelor of Science

by
Jake Aki
Adam Croteau
Ryan Walsh

Neil Heffernan, Co-Advisor
Cristina Heffernan, Co-Advisor

Abstract

Online practice and teaching tools have existed for awhile now. Companies like Pearson have cornered the market in higher education with prohibitively priced products. ASSISTments seeks to provide an equivalent or superior service through an open source platform. This project developed content on this platform to serve the needs of AP physics classes and to survey the efficacy of different instruction methods.

Acknowledgements

This project would not have been possible without the hard work and dedication of many individuals. We would like to take the time to thank everyone who made this research possible.

ASSISTments is online education ecosystem designed by Professor Neil Heffernan and Christina Heffernan and hosted by Worcester Polytechnic Institute. We used the ASSISTments framework to design all of our content and to implement all of our random control trials. Without the power and flexibility of this tool, this project would not have been possible.

Rick Doherty worked with the team to help develop our AP physics content. His insight into high school physics classrooms from his experience as a teacher coupled with his enthusiasm provided invaluable support for the team as we built out most of our content.

Our advisors Neil and Christina Heffernan have been with us for the entire project giving us guidance and support. Their knowledge of ASSISTments is about as exhaustive as possible, and their feedback shaped the progress of this project.

Thank you so much.

Executive Summary

Our Goals

The goal of this project was to develop effective content for the ASSISTments online tutoring platform. Specifically, the focus was on material for preparing high school students in AP Physics classes. The team also wanted to help Rick Doherty develop his content in parallel with ours as there was considerable overlap.

Methodology Design

This project is roughly broken down into three categories: there was content creation, random control trial (RCT) design, and the content dispersal. The content was designed within the ASSISTments platform and was modelled after some of the content designed by Rick Doherty. RCT design was simple and undertaken independently by team members so that three RCTs could be attempted. Content dispersal was set up in B term, managed during C term, and then data was gathered and analyzed in D term.

Findings

Unfortunately, the team’s RCTs generated minimal participation. However, attempts to ameliorate the lack of raw empirical data was compensated by citing some research. Furthermore, analysis was put into why the participation was low.

Recommendations

The team saw the greatest success at garnering interesting while at the aweSTEM teachers conference, so the main recommendation was to promote this content with other teachers even more. Another important thought would be to run the trails during a more applicable timeframe. e.g. kinematic tends to be taught in the fall semester not the spring.

Contributions

Jake Aki

Jake Aki is an electrical and computer engineering major concluding his junior year at WPI. He worked heavily on fleshing out the problem sets for the AP physics content but his major contribution to the project was the design and implementation of the website that hosts the team’s content. He also wrote and edited portions of the paper such as the introduction, background, and extraneous pieces.

Adam Croteau

Adam Croteau is a chemical engineering major concluding his junior year at WPI. He worked on surveying students to gather feedback allowing the team to make more informed decisions about the design of the problem sets. He also helped gather data in the later stages of the project and helped to write the paper with emphasis on the findings, conclusion, and recommendations.

Ryan Walsh

Ryan Walsh is a computer science major concluding his junior year at WPI. He worked closely with Rick to develop problem sets for the AP physics content, but also to help Rick on his projects developing physics content for the classes he teaches at Shrewsbury High School. He also helped to write portions of the paper with emphasis on the methodology and implementation of our problem sets.

1 Introduction

“To improve education through scientific research while not compromising student learning time” is the mission statement for ASSISTments (Heffernan & Heffernan, 2014). ASSISTments is a platform, which allows teachers to use, build, and edit content that they can then use in their classrooms to provide their students with immediate feedback. As a platform, ASSISTments brings a lot to the table. Teachers who use it are able to get a large amount of consistent data on the efforts of their students. Students can immediately know what they are doing right versus what they are doing wrong, and perhaps most importantly, ASSISTments provides everyone with an ecosystem in which to work.

The essence of the idea behind this is that, quoting the creators, “The ASSISTments platform assists while it assess.” (Heffernan & Heffernan, 2014). There’s nothing in the design of the platform that inhibits a teacher, research, or IQP student from running a research trial while simultaneously providing useful content. In fact, that’s the whole point. Because ASSISTments allows the content creators to monitor the use of individual problem sets and generate a large amount of data, studies can be run in parallel with the implementation of a particular problem set so that you can constantly roll out content and iterate based on feedback.

When the IQP team was brought on board with the ASSISTments project, they were initially brought up to speed by working on content to expand the ASSISTments certified portion of the website. However, the team quickly began to branch off with the help of Rick Doherty, a Ph.D. student who was instrumental to a lot of our work on this project. Doherty’s work was primarily in developing physics content specifically geared towards high schools trying to expand their support for AP physics curriculum. Thus, the bulk of the team's initial work was into developing and editing a lot of Doherty's content while simultaneously expanding the reach of his content to cover more topics.

Another important component of the project work was into developing some random control trials (RCTs) to test different ideas about the efficacy of different teaching strategies. In order to do this, the team developed a website to serve as an access point for all of the content and then seeded the RCTs throughout the content. This left a final phase of finding people to consume the content and provide feedback.

There were three primary groups of clients that were targeted to find consumers. First, all three team members work as freshmen RA's, so they took advantage of floor wide email aliases to gain participants working on freshmen level college physics. Second, there was the pool of general acquaintances, coworkers, and fraternity brothers. Lastly, the team went to the aweSTEM conference held in D term to advertise to high school teachers with the goal of getting them to incorporate ASSISTment content into their curriculum.

2 Background

2.1 Introduction

There exists several forms of online tutoring softwares or web-based homeworks. At WPI perhaps the most notorious is WebWork. Another contender for popularity amongst campus faculty is Pearson's Mastering line of software (Mastering Physics, Mastering Chemistry, etc). There also exists a measure of academic literature on the topic of online tutoring. VanLehn built what he called the Andes physics tutoring system, and released a five year evaluation of it (VanLehn, Lynch, Schulze, Shapiro, Shelby, Taylor, Treacy, Weinstein, & Wintersgill, 2005). An older study carried out by researchers at Carnegie Mellon reviews a 10 year project based on research by J. R. Anderson (Anderson, Corbett, Koedinger, & Pelletier, 1995).

2.2 Competitors

WebWork is well known on campus. A great deal of introductory freshmen classes use WebWork to provide some extra practice work, and some courses even incorporate it into the class rubric to make exams and quizzes less punishing. It's easily recognizable from its minimal HTML-esque appearance, and it isn't particularly robust in responding to different questions.

Pearson provides a very pretty product. It's professionally designed, often comes packaged with an electronic copy of the textbook, and provides more robust answer processing and feedback than WebWork. However, a common complaint from students is the hefty price tag. A semester of use can cost around $75, which can be untenable especially if multiple classes use it.

Furthermore, MasteringPhysics is almost infamous on campus for creating an extremely frustrating homework environment. The typical structure for the homework is a five attempts for partial credit with hints provided after you can no longer get credit for your answers. Because professors tend to incorporate MasteringPhysics as a mandatory, graded portion of a class, this leads to a lot of frustration when students see points getting taken off in front of their eyes from typos, rounding errors, etc.

2.3 Academic Literature

The Andes physics tutoring system is a robust system with a lot of tools that allow it to be extremely flexible and logical in application. However, it is an extremely focused tool. It is specifically designed to work with physics problems. Furthermore, it serves to primarily accommodate instructors by providing some automated labor as opposed to serving the students with feedback. VanLehn's work is primarily into what he describes as "web-based homework (WBH) grading services" (VanLehn et al, 2005).

However, the Andes system does seek to offer some rudimentary tutoring for students. Fundamentally, Andes is structured in a way that is logical for a student working through a physics problem on paper. There's the problem statement, a section for writing down variables, a section for describing a free-body diagram, and a final quarter for writing down equations (VanLehn et al, 2005). This is substantially more detail than either WebWork, MasteringPhysics, or ASSISTments is capable of offering. However, it's easy to contrive situations where this additional functionality doesn't make sense when trying to stretch this to other subjects. Equations and variables aren't needed in biology, for example, and free-body diagrams are really only applicable in physics, and even then, they are less relevant to certain topics such as electromagnetism.

The hints in the Andes system are also quite simple. They tend to follow a three step system. VanLehn calls them “a pointing hint, a teaching hint, and a bottom-out hint.” (VanLehn et al, 2005). This is similar to the strategy that Doherty used in his physics content, and the strategy that the team used in the content that was built for the IQP. The ASSISTment hints simply expanded on the idea with two “pointing hints”, two “teaching hints” and then the necessary “bottom-out” hint.

The bottom-out hint is an interesting concept in itself, because it is necessitated in an automated tutoring system. There are two main options for allowing progress throughout a problem set. The system can either provide the student with some kind of option to give up, or there must be someway to get the answer to the question and sacrificing credit. The later option is obviously more desirable as the student then gets the opportunity to figure out their mistake. However, pains must be taken in order to avoid any danger of academic dishonesty.

Regardless, VanLehn found substantial success with his Andes tutoring software, and the results from his five year study show a promising correlation between the Andes software and improved test scores amongst students at the U.S. Naval Academy.

VanLehn’s work was inspired by the Advanced Computer Tutoring (ACT) theory developed at Carnegie Mellon. This research was conducted by Anderson and was primarily applied to LISP, geometry, and algebra (Anderson et al, 1995). The main ideas of the ACT theory had to do with three assumptions: procedural-declarative distinction, knowledge compilation, and strengthening.

Procedural-declarative distinction is the idea that there exists a distinction between procedural knowledge and declarative knowledge. Procedural knowledge being the ability to take a theorem or law and apply it to a problem in order to solve the problem. Whereas, declarative knowledge is limited to simply being able to describe or state the theorem/law.

Knowledge compilation is a description of how declarative knowledge can be converted into procedural knowledge through practice in problem solving activities. This process of interpreting problems generates the rules by which the procedural knowledge is built up.

Strengthening is the idea that both procedural and declarative knowledge improve and strengthen with practice. So, if you forget how to do something, with practice you can recover that skill. And, if initially you're not that good at something, you can eventually become skilled with practice (Anderson et al, 1995).

Ultimately, Anderson's research had to do with using these fundamental tenets of learning to promote what is described as "cognitive skill" (Anderson et al, 1995). He identified certain forms of feedback and help that promote cognitive skill and different models and schemes for tutoring. Because the scope of his research was primarily mathematical, it can be difficult to stretch some of the concepts to other disciplines, but with relation to the computational portions of physics, it is easy to see where comparison can be made.

3 Methodology

3.1 Building A User Base

In order to obtain an adequate sample size for random control testing, the group tried several different ideas to attract users. To get more use, the RCT's would need to be part of a package of material that was in demand. Furthermore, the material would need to be comprehensive to retain interest and to provide enough material to embed our RCTs in. The best way to create this kind of breadth was to cover an entire course or multiple courses. At first, it was considered only catering the content to the students at WPI taking PH 1110 Mechanics and PH 1120 Electricity and Magnetism. The decision was made to move away from a WPI-centric group as it seemed that the potential pool of users would be smaller, and that the group did not make much head way in their attempts to get in contact with professors at the school. The attempts to get in touch with professors proved to be far more difficult than anticipated at the beginning of the project. The initial impression was that several of the professors in the physics department were highly interested in trying new practice material for their classrooms. That did not end up being the case, the professors became busy with the school year and did not have the time or the ability to change their current course structure for this school year. Along with inability to change their courses, professors at WPI seemed to want a fully developed system early on in the year. They essentially wanted a completed physics course of content at the start of the year when the project began, and the group could not deliver on that much content until near the end of B term.

There was still the possibility of reaching out to students on campus and that would also prove to be quite difficult. Students here at WPI did not seem to have the time or ambition to do physics practice that was outside of what was normally assigned by the professors.

Therefore, the decision was made to move to a much larger demographic, which was high school students enrolled in AP physics across the country. Typical AP physics review is available through expensive review books, so a free alternative seems like it has a place in this market. This changed the problem of user engagement. Instead of explicitly trying to contact professors and fellow students, the team was also attempting to get in touch with people outside the group's direct community.

3.2 Creation of the Website

Since the team had decided to try to engage users from outside the university, an attempt was made to make a more presentable product for outreach. It was determined that a website would be the best way to go. A website enables students to use the content via embedded ASSISTments direct links. Therefore, the group created a website using the URL http://jmakiusa.wixsite.com/-Assistmentspracticep, which was then abbreviated to tiny.cc/physicsassist for convenience. Once the website was created, the idea of a sales pitch was tossed around some, because even though the group wasn't selling a product they still needed users, and it was clear that just having a website would not be sufficient. Therefore, they added an "About Us" section to the website. The sales pitch was centered around how the group was all students at WPI and had undertaken this project as their IQP. That being said, emphasis was on the group pushing their product to be used by students to help prepare them for the exams and to help the group develop data and debug our system. In addition to the “About Us” page, they also organized the website such that it would be user friendly for students who were looking to study a specific subject area. Therefore, the website was broken into 5 major categories and subcategories that are pertinent to individual equations giving the user the ability to focus closely on what they are having difficulty with.

3.3 Personal Connections

The group attempted reaching out to their personal connections that were available to them. This included attempting to get in contact with previous professors of AP physics that both Adam and Ryan were in touch with from high school. The professor Adam reached out to informed him that Fairhaven high school was no longer offering AP physics, and Ryan’s professor told him he would take a look at the material and potentially implement it into his classroom the following school year. Following that, the group had no more personal connections at the highschool level and appealed to students of WPI again via email and in person to get more potential users. This proved to be of minimal value and would not be of much help in getting statistically significant data. It did prove to be useful in developing a better anecdotal response to the material as many of the students who did go through the problems sent back critiques for the group to work on.

3.4 AweSTEM

The group also attended the AweSTEM event on campus where they table sat and provided information about the content to interested teachers. The team collected the teachers’ emails and sent out the link to the website. In the talks with teachers at the event, it was clear that many of them enjoyed the content that had been created and were interested in using it in their classrooms in the future. The group received a good deal of positive feedback from AP physics teachers on the product but also from some teachers for younger age students. Teachers who taught science in the later years of middle school, where some of the material would be applicable to their students, thought it could be quite beneficial. In particular, much interest was expressed about how indepth the hint system was and how it could address many problem areas for potential students. This does raise the question of how beneficial this content could be to users in courses less rigorous than AP physics. Overall, the conference left us feeling that there is definitely potential for the content that was created and that it could certainly benefit those taking the appropriate courses.

3.5 RCT Development

Toward the beginning of the project, the goal was to create content that was comparable to that which was offered by MasteringPhysics by Pearson. The overall idea was to try and create a system that would be implemented in WPI's basic physics classes PH 1110 Mechanics and PH 1120 Electricity and Magnetism to replace MasteringPhysics by Pearson. This focus would eventually be changed to creating content for AP physics, which overlaps heavily with the content in these two classes but has a much larger pool of students. With the initial intention the group wanted to look into how the platform of ASSISTments could be adapted to more adequately fit the user's needs than the product being delivered by Pearson.

There were several main questions that were being assessed in the open response surveys. Most had to do with the likes and dislikes of the user about the interface that was provided by Mastering Physics by Pearson, and the educational value of the work compared to that assigned from a textbook. Since the survey was mostly open response the surveys were looked through for frequency of similar problems. In total 39 open response surveys were filled out. The full frequency of answers with the questions is in Appendix A. A copy of the survey is in Appendix B.

One thing that the group wanted to check immediately was how beneficial students felt the material provided was. This was done by a comparison of the educational value of assignments provided from a textbook and the work being provided via online content. The setup for that particular question of the survey gave a multiple choice question with 3 options. The options ranged from being less than, equivalent, and greater than what was offered from book work. The results of the survey showed that students felt that the material offered by Pearson was not as educational as book work but was similar in practice work. Some of the reviews that felt the work was highly valuable said that the work was great for letting the student know where they went wrong immediately. Other reviews that felt the material was not too helpful described the questions from Mastering Physics to be unlike those that would be seen on the exams. The ASSISTments content that the group created was able to incorporate questions from the AP physics exam as post test items, which would make the practice material feel more like what the students would see on the exam.

Another topic area the group wanted to evaluate was the overall satisfaction of students who used MasteringPhysics by Pearson. This survey question was also carried out via multiple choice and had 5 options ranging from highly dissatisfied to highly satisfied. The results of the survey showed that the average student was between being indifferent and dissatisfied. This would suggest that there may not be an overwhelming need for change rather a chance for improvement. The survey also included open response questions about likes or dislikes. These open response sections were made to give the group ideas on where to improve their content and what the users might like to see.

Some of the characteristics that students enjoyed from the interface included the instant feedback on how they were doing, the ease to submit homeworks, and the ease of use (i.e. the user interface). All three of these characteristics apply to the material that is supported by the ASSISTments platform. Since the major upsides were hit by the ASSISTments platform, the group had no reason to try and build on these features. Therefore, the team looked into some of the criticism. Out of 39 responses, 19 complained of poor answer boxes and system failures to read answers correctly. This problem proved to also be a sticking point for the ASSISTments system as the system can only really read the correct answer in the forms that are supplied by the programmer. This means that slight variations such as rounding or miscalculations were marked as incorrect. One of the largest issues with this is that there can be multiple correct ways to describe direction in classical mechanics, but ASSISTments can only recognize the notation used by the designer. Thus, a student can submit a completely correct answer but have it marked as incorrect.

The group tried to compensate for this problem with Jake’s RCT which has the identification PSA3HGT. The premise of this RCT was to break up the answering section of each problem to give the student a better chance of getting one piece of information at a time. The problems that were assessed by this RCT were on the kinematic equation for displacement: $x = x_0 + v_0 t + \frac{1}{2}at^2$. The problems were broken up into 3 separate parts. The first asked for the magnitude only. The second part asked for the units of displacement (e.g. meters). And, the last part asked for the direction of displacement. By breaking up the problem into multiple parts, it was the group's hope that it would avoid some of the problems that were exhibited by the other physics problems created in ASSISTments. Since the control problems typically required that the student answer the magnitude, units, and direction on the same submission. This created problems should the student answer correctly but put those three in a different order than what was anticipated by the programer, didn't leave spaces, or left a different number of spaces than intended. In the end, it simplified what the person was responsible for putting in the answer box.

Another large problem based on the response frequency (12 out of 39 responses) was poor hints and explanations provided when a student is stuck on a problem. The hints that were provided by Pearson sometimes left the students with no more information than they had before they requested the hint. Students felt that the explanations of the problems were not adequate to get them through the mental block they were having with the problem. The hints they were too focused on the conceptual piece of the problem and did not give ample explanation of the analytical pieces of the problem. In order to assess this problem, it was decided to create an RCT around this topic. Ryan wrote this RCT which has the identification PSA3Q3X.

This RCT runs off of a control trial that exists with short hints that are typically provided with the physics problems, and a treatment that has much more in depth and explanations. The explanations for the more in depth hints was provided by Rick

Doherty who provided a critical view of all of the problems created for this IQP. In particular, he helped to develop the much more in depth hints using his knowledge of physics and time spent teaching in the classroom. He provided guidance when it came to helping the users to more fully understand the concepts of each problem and giving them the tools to finish each problem. During the AWEstem event that the group attended to show the content to teachers in the science community, this particular RCT was shown to them and the teachers were very impressed by what they saw in the hints.

The third RCT revolved around student comprehension of a topic before doing the assignments. Adam developed this RCT. The idea was meant to help students preemptively when they were having difficulty with topics. Online software can be quite difficult to use due to the specificity of answer boxes and lack of flexibility. If a student is having trouble with the problems to begin with, the work can be quite frustrating and hard to do.

A good idea might be to give the students a brief overview of how to solve the problems before they enter the problem set. Several ideas were evaluated before coming up with the final solution. Specifically, the idea of going for a video or a written explanation. The video was highly considered, but the group did not feel comfortable or qualified recording their own to present the material to users. Then, the idea of whether to use a video from the web was a viable option, but due to the possibility of that causing problems with copyright infringement, the group dropped that idea. Therefore, the decision was made to go to a text based explanation for students. This RCT is quite simple. Either the student is given the control, which has no explanation, or the student is given the treatment, which includes an explanation before the problem set. This RCT was done on the kinetic equation for work: w=fd.

The explanation gives a short description of what all of the units used in the formula are. It also describes the direction of the applied forces. Hopefully, this information was going to give the user enough of an understanding of the material to help them with the problem set that followed.

Those were the three RCTs that the group ended up designing. However, in this section there is some additional data from the survey that was used in helping to develop the problems that are worth mentioning. Some of the problems that were stated by students in the surveys are solved via the ASSISTments platform, and the group did not have to fix. The largest of which had to do with students who expressed frustration around losing credit for incorrect answers and using hints. Granted the ASSISTments platform does deduct points for hint usage and that makes sense for a student just pushing through the hints and getting the question correct would teach them little. The nice part of ASSISTments that students found about the problem set system is the three in correct method allows students to get questions wrong and to learn from their mistakes without losing credit on the assignment. Another advantage that ASSISTments showed in the surveys was that most questions that are asked by Mastering Physics are available for free on the internet. Which lead to a large contingent of students just searching for the correct solution on the internet, because they did not want to lose credit on those problems.

4 Findings

4.1 Introduction

Throughout all of the attempts to obtain data, the group was unable to collect enough data for it to be statistically significant. Therefore, the analysis done in this section will be highly speculative. In order to get more conclusive results, more data would be needed and in addition, it would help to enhance the tests by providing a questionnaire along with the data to get more explanatory information into why the data comes out the way it does. Regardless, some conclusions can be drawn from the data that was collected.

4.2 Ryan’s RCT

Starting with the data that was collected from the RCT designed by Ryan with identification number PSA3Q3X. The full data is available in Appendix C. There were 25 students who started the problem set and 7 who finished. The low rate of completion was seen through all 3 RCTs. However, some were lower than others. In this RCT, the rate that finished the control was 38.9% (7 of 18 trials) compared to the treatment which had a completion rate of 14.3%(1 of 7 trials) which was quite a bit lower. When looking at the problems, it is relatively clear to see why this was the result though.

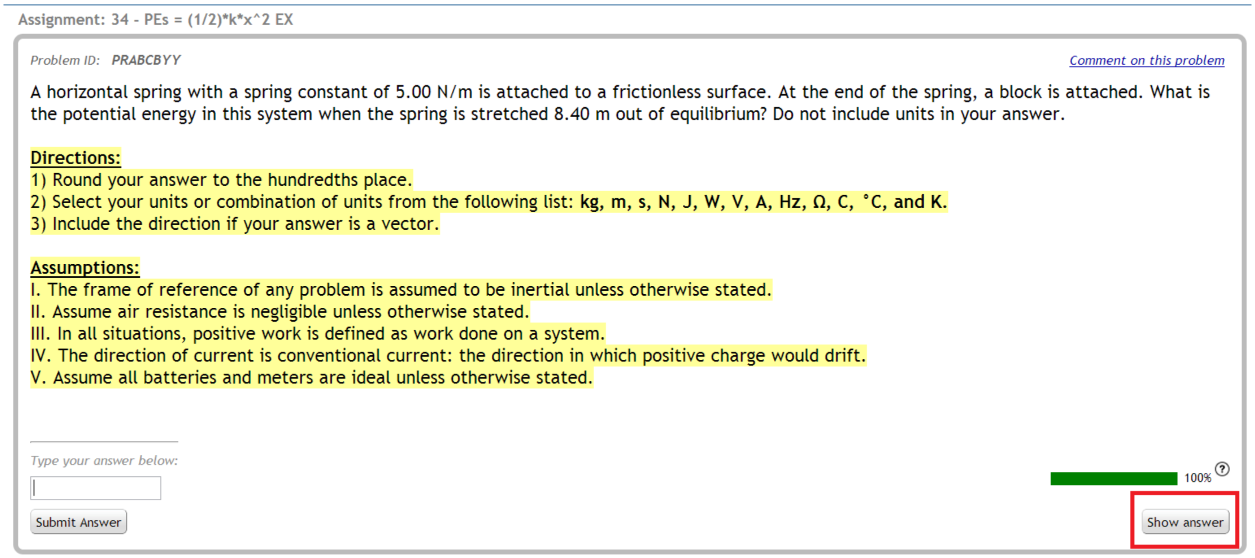

Fig. 4.1 Screenshot of the control question without the hint expanded

Fig. 4.1 shows the control question. Show hints is replaced with show answer (as seen in Fig. 4.2). It is considerably less frustrating and less time-consuming to just click through the show answer button a few times until the user becomes familiarized with the content and then, use rote memorization to solve three in a row and complete the problem set.

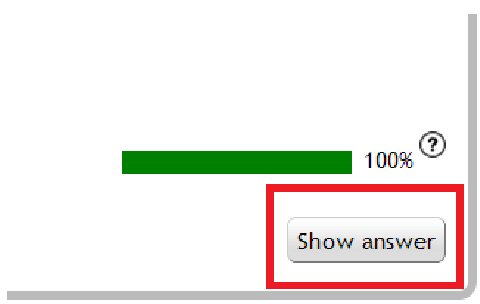

Fig. 4.2 Screenshot closeup showing that the hint button, is replaced by a show answer button

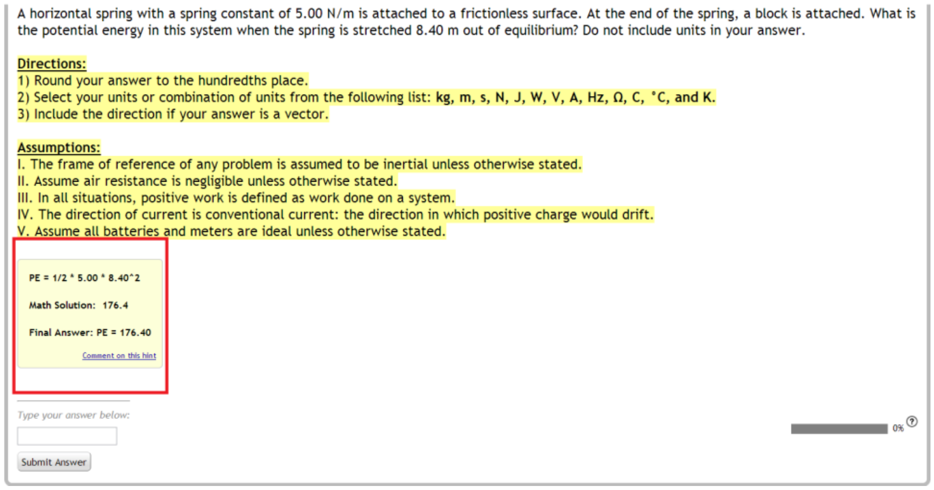

Fig. 4.3 Screenshot of the expanded hint for the control experiment

As you can see in Fig. 4.3, the solution is handed to student along with necessary significant figures and units. It's much easier to identify the pattern that ASSISTments is expecting using the control.

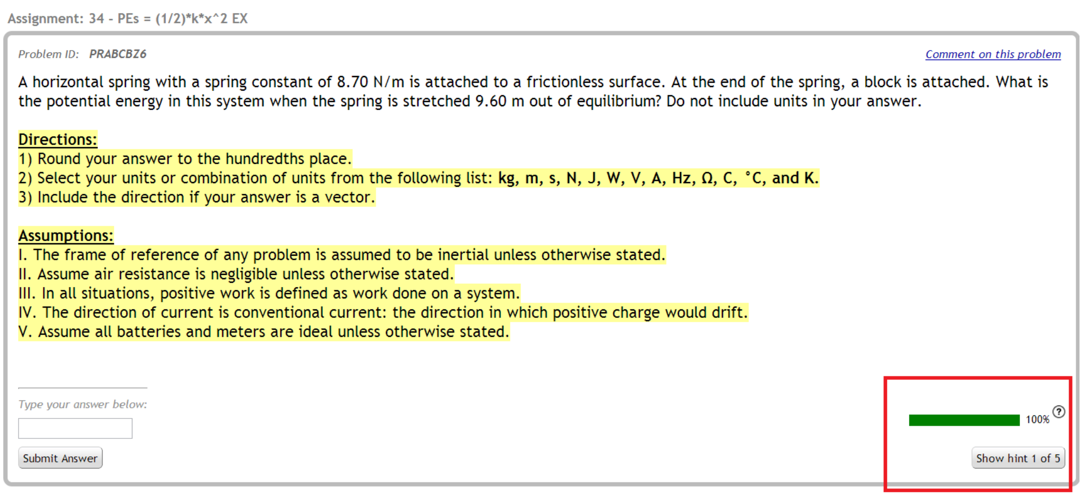

Fig. 4.4 Screenshot of the treatment version of the question

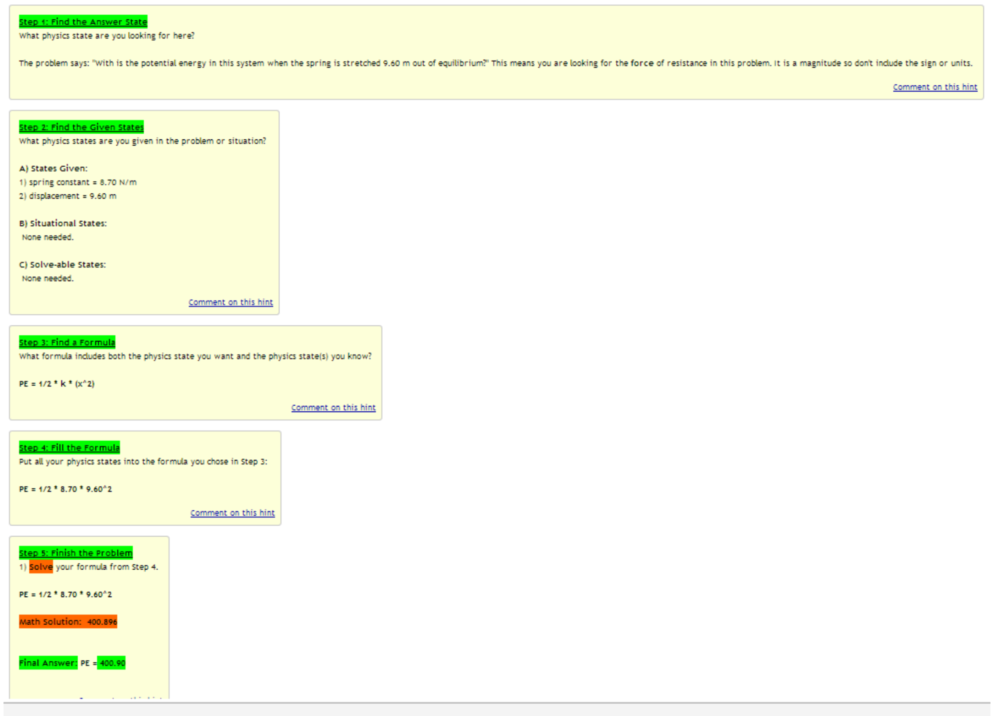

Fig. 4.5 Screenshot showing what the expanded hint of the treatment looks like.

As you can see in Fig. 4.4 & 4.5, the more in-depth solution is definitely more educational but takes longer to click through. Furthermore, the team can contribute that anecdotally talking to participants, it is more impactful to click through as you see yourself slowing "losing credit" for each hint, even though ASSISTments technically uses only a pass/fail system for evaluating correctness of an answer.

The purpose of this RCT was to give more in depth hints to assist students. It would appear at least in completion rate that this had an opposite effect on the students. This could potentially be due to the simplicity of the problems. They are algebraic and follow the same general form for all of the problems in this problem set. Therefore, in the control when it gives the student the form filled out it teaches them the method to doing this problem. However, it lacks the conceptual component. Therefore, since the problems go directly to the solution in the control students can quickly learn the form and finish the problem set.

4.3 Adam's RCT

The next RCT is Adam's with identification number PSA3QYY. The problem set focuses on giving students a brief overview of the kind of problem they are undertaking before getting into the problem set. The intro treatment is shown in Fig. 4.6, and the complete data set is attached in Appendix D.

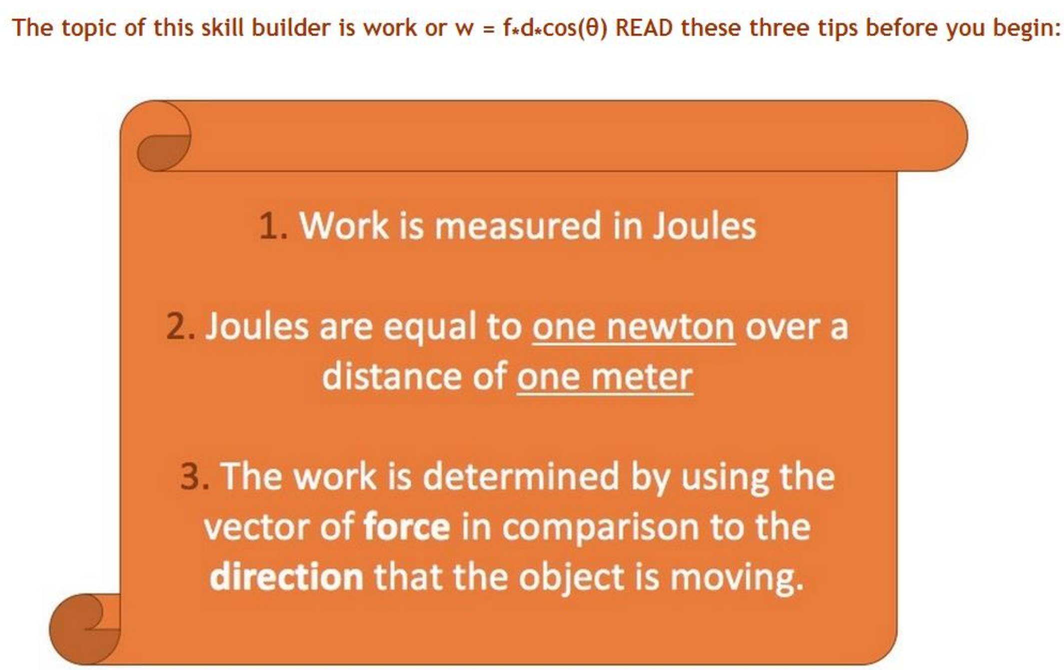

Fig. 4.6 Screenshot of the treatment introductory question

There were 12 overall trials for this RCT: 9 for the control and 3 for the treatment. The completion rates were 11.1%(1 out of 9) for the control and 66.7% (2 out of 3) for the treatment. Since the trial number was so low the reasoning for the higher completion rate is highly speculative, but it could be due to the students having a slightly better understanding of the material. With just 12 students, that could be on a more individual basis that was pre-existent to this RCT. The post test data for this RCT was quite low at just 3 trials the results are shown in Fig. 4.7.

	A	B	C	D	E	F
1	User ID	Number All Problem:	Number Complete P	Number Correct Prob	Assigment Complete	Condition
2	419968	2				control
3	415921	2				control
4	402302	2				control
5	415920	2				control
6	416107	2				control
7	427536	2				control
8	427533	2	2	2	1	control
9	427663	2	2	1	0	control
10	423691	2				control
11	422918	2	2	0	1	treatment
12	427642	2	2	2	1	treatment
13	427664	2				treatment

Fig. 4.7 Graph of problem level data for PSA3QYY

By the results in the table, it would appear that there was actual a 4th finisher who was in the control, but the system did not register them as having finished the problem set. In reference to the post test problems completed, 3 of 4 were done correctly for the control, and 2 of 4 were done correctly for the treatment. The findings are hard to draw conclusions from but would suggest that the intro had little effect on comprehension. That said, it is important to remember the feedback received from teachers in AweSTEM, which was very positive and suggests that there is room to grow.

4.4 Jake's RCT

The last RCT that was written by Jake had the least participation. This is likely due to the complexity of the material that this RCT covered. The identification number is PSA3HGT, there were a total of 4 trials and none of them were completed. One reason why there may not have been any students who finished was the design of the problem. The RCT was testing the effect of having a split up problem to help with answer box problems. However, three separate sections may have been too time consuming for students. However, there were no finishers for the control trials either, which could suggest that this particular topic could just be difficult for students in general. The results of this trial can be seen Appendix E.

Outside of the reasons discussed above there should be the additional discussion of bias. The problems were sent out heavily to WPI students whom have for the most part either taken AP physics or they have taken the Physics: Mechanics and Physics: Electricity and Magnetism which makes them more well qualified to understand the material than the average AP student. Their grasp on the material may be above average, but their commitment to completing the material is far below that of the average AP student as they do not stand to gain from finishing the problem sets.

The other users that the group advertised out to were teachers who like students at WPI are quite qualified to finish the material but lack a desire or need to finish it. This should create low completion numbers and high percentages correct. This was partially the case for these trials. The sample size proved to be too small for any definitive statements to be made, if a larger sampling of people completed the trials it would lead to much better correlations. Moving forward more outreach and sample size would be needed to draw and conclusions from these RCTs

5 Conclusions

5.1 Advertising

The largest issue that faced this IQP was the lack of content users. The team identified several primary impediments in developing a stronger user base. It is difficult to line up the timing of the IQP with the high school physics schedule. Our team was getting up to speed and learning in A term and then developing content in B term. From an advertising perspective, content should've been ready before A term so that A term could be used to propagate the material and generate users so that our material would be arriving at the time that students needed it. Another problem was our target audience. AP physics made sense because of the support we had from Rick Doherty and the scale of the audience. There are many more AP physics students than there are introductory physics students at WPI. However, this means that there exists only one window of opportunity to start spreading our content, which is before the start of the fall semester. Otherwise, they start moving into content that hasn't been covered, such as electromagnetism.

5.2 RCT Conclusions

Despite that, the team found that they were able to generate some user feedback. Insight from teachers at the AweSTEM conference suggests that Adam's RCT was a success. Providing some degree of preparation and advice before the start of a problem set seems to be a very popular idea and is worth building on. Our implementation was simple, but given the positive response, it would be worth developing even better introductory content. Perhaps even an entire IQP could be spent on recording video tutorials, example problems, and similar content.

Ryan's RCT confirmed that the more comprehensive hint strategy was worth doing. During meetings with Rick, concerns had been expressed that it wasn't worth the effort of having such a complicated hint structure, but feedback suggests that this kind of feedback is one of the best selling points of an online tutoring platform. Furthermore, this RCT generated the most user participation due to its simplicity, and it demonstrates that the simpler a problem is (even in the hint strategy) the more participation you get from users. This suggests that a potentially valuable avenue for exploration is trying to find a way to design problems to be simpler while teaching the same content.

The RCT with the most negative reception was definitely Jake's. Unfortunately, it seems that the complexity of the answer (i.e. needing both magnitude, direction, and units) stressed the limits of the ASSISTments software. Both entering through multiple submissions and as a single submission irritated users. In the multiple submission strategy, a single mistake meant that the entire problem was marked wrong, But, the single submission strategy made it very difficult to identify where the error was in the user's submission. It's difficult to identify an effective strategy to move forward from here with, but the general consensus from the team is that for more complex problems a multiple choice as opposed to free-response might be a less frustrating option.

6 Recommendations

6.1 Outreach Problems

Looking at the attempts made to distribute the content, it is clear that the real way to get the content to get significant usage would be to get teachers to use the content as part of their lesson plan. The attempts that were made to obtain individual users were relatively futile as many students are not in need of this specific content. When not taking the appropriate courses, students tended to not complete the trials as they had little motivation to do additional school work. When students were in the correct courses for the material to be applicable, they also did not tend to want to test the material. This was mostly due to the structure of coursework. In a majority of cases, students who were in the correct classes had adequate practice between both the written work assigned out of textbooks and other competitive software such as MasteringPhysics by Pearson. This would lead back to the conclusion that the best way to make headway would be advertising it at a teacher level. If the support of a teacher is gained, then there will be a grouping of required users. Gaining the support of individual teachers at any level is also a relatively difficult task from what can be assessed from the distribution work that we did. If the group was to do more advertising, there are several avenues that they feel would be best to promote this content in the future.

6.2 Future Advertisement

One of the best ways that the group could see for this product to be advertised would be for the product to be brought to more teacher conferences. Specifically, it could be brought to conferences on how to teach physics. This would give a great opportunity for teachers to see the content and to view its accessibility. It could be said that one of the greater selling points for the product has to do with accessibility to the student in terms of getting them connected with the content that they need to study. Each of the topics for all of AP physics 1 are laid out on the website, and that would create a great resource for students who are struggling in a specific problem set area.

It should be highlighted why ASSISTments was created. The goal is to provide free educational software to students could use it. With that in mind, it is relatively easy to determine which schools participate in the AP physics 1 exam. Once a list of schools is acquired, it may be good to look into the districts and individual schools that have low budgets and come from lower income areas. With this cross reference of schools, looking into the individual school's websites the email directory should be able to point to who their science or physics teachers are. Sending those teachers emails next year would probably lead to a much better population of people to use this content and would likely be the best shot at getting into teachers' plans for their classes.

Another element of the advertising for this set of content to get more trials on the RCTs would have to be timing. It is quite unusual for a teacher to be willing to change the work they are currently assigning in a classroom due to what they receive in an email. It is much more common for them to evaluate what is being given to them over a period of time to ensure quality for their students. Once they have ensured quality, they still are not very likely to start implementing that new material into their classroom right away. They will likely save the changes for the following time that they teach the class. This would mean that advertising to teachers and getting results within the same year is improbable. Therefore, the best time to push this set of content is probably over the summer when teachers are on vacation and will have the time to look over the content. Also, the summer time is typically when they are looking to make changes to their intended course plan for the following year. However, it is also important to make sure that ASSISTments doesn't get forgotten about, so emails shouldn't be sent at the beginning of the summer but instead, closer to the end.

In summary, the best way the group could theorize to get teachers on board with the content would be to target low budget schools that participate in AP physics 1 exams. Then, send emails to their science teachers about the content during either the summer or spring time to get them interested in using it in their classrooms for the following year.

In retrospect, the possibility of being able to create content and have it tested within the same school year is a difficult task. In the talks with teachers and professors that happened in the time of this IQP, it could be said that the content created and the method with which it is available has potential to be used by a decent group of teachers and to get more widespread feedback over the course of next school year. More advertising would be needed in order for a substantial group of people to use the content more frequently. However, teachers at AweSTEM seemed to be pretty excited about the concept of the content and if extrapolated further, it is likely more teachers would get excited about ASSISTment physics content.

6.3 RCT Development

Some things that could potentially be built upon in the future with respect to ASSISTments physics content would be the incorporation of graphs and figures. At the current state of the content no graphs or figures are present for the students to interpret and that would be a nice addition in the future to enhance the problems. There are certainly more iterations of RCTs that would be useful using the ASSISTments platform. Just a few that the group had discussed at the time of writing the current RCT's that have already been mentioned would have been video explanations and problems that revolved around figures. In addition some others that were discussed were a more comprehensive question where the system provided pop up hints to certain common errors for that form of problem. Another option was to move exclusively to multiple choice for a problem set and then see if the students who underwent only multiple choice questions had the same grasp of the material as those who had to answer the open response. Another interesting idea for a problem set that could test retention would be to take the questions that were designed by the group and put them against questions designed by another source. The other source could range from being physics teachers to being questions taken off the AP exam and conducting a more in depth post test to get a better feel for whether students benefit from more complex questions being asked of them.

These were just a few ideas that the group had considered briefly and there are many other RCTs that could be conducted that would be smaller in scope, that would include just feature changes throughout the problems. The three that the group decided to go with will be discussed further in the next section to go over what their findings were.

References

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: lessons learned. *The Journal of the Learning Sciences, 4*(2), 167-207. Retrieved from http://www.jstor.org/stable/pdf/1466690.pdf

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Artificial Intelligence in Education Society*. Retrieved from https://www.researchgate.net/profile/Neil_Heffernan/publication/266378101_The_ASSISTments_Ecosystem_Building_a_Platform_that_Brings_Scientists_and_Teachers_Together_for_Minimally_Invasive_Research_on_Human_Learning_and_Teaching/links/56894dec08ae1975839a93a5.pdf

VanLehn, K., Lynch, C., Schulze, K. Shapiro, J. A., Shelby, R., Taylor, L., Treacy, D., Weinstein, A., & Wintersgill, M. (2005). The Andes physics tutoring system: Five years of evaluations. In G. McCalla & C. K. Looi (Eds.), Proceedings of the 12th International Conference on Artificial Intelligence in Education. Amsterdam: IOS Press. Retrieved from http://www.public.asu.edu/~kvanlehn/Stringent/PDF/05AIED_KVL_CL_KS_JAS_RS_LT_DT_AW_MW.pdf

Appendix A

What do you like about Mastering Physics (Please be specific with your answers)?
-(5)Hints/explanations
-(2)It understood when answers were close and informed you
-Easy to keep track of assignments
-(7)Easy to submit homeworks/Ease of use
-Good question bank
-(2)Pictures and animations help visual learners
-(7) instant feedback
-(2)Option for additional practice
What do you dislike about Mastering Physics?
-(7)Cost
-Doesn't teach effectively
-Lack of interactive work
-(19)Poor answer boxes/system fails to read answers correctly
-(12)Poor hints/explanations
-(7)hints deduct points
-(8)Answers are available online so students just look up the answer instead of doing the work
-(3)unrelated to class
-(1)The exams felt like a large jump from mastering physics
Please rate how you feel Mastering Physics helped to enhance your understanding of the material provided by the professor
Comment on your reasoning for your above rating:
-Good practice(3)
-Gives where you went wrong(3)
-It was like normal Hw (2)
-(2)Only source of homework(2)
-Giving only a numeric answer, hurts the need for procedures (1)
-The test questions are very different than those on mastering physics(1)
-They were basically just textbook questions (2)
-Questions were unrelatable or inapplicable(1)
-(3)questions were not related to class(1)
What do you think of the way problems are presented?
-(7)Poorly worded
-(11)Easy to read
-(3)Lack of organization
-(2)Ambiguous questions
-Liked that the problem was broken up into parts
-Disliked when it did not specify which figures were required in the answer
-The questions are similar to that of the textbook and detract from the use of the textbook
Any additional comments that you would like to add about Mastering Physics:
-(2)Cheating is very frequent

-Connect with canvas
-Looking for the answer online in order to not lose credit
-should be an optional service

Survey question responses	
Question 3	Question 6
4	3
5	2
3	1
4	2
3	1
3	2
2	2
3	1
5	3
2	2
3	1
2	1
2	2
4	3
3	2
3	2
4	2
4	2
3	2
2	1
3	2
1	1
2	2
1	1
2	1
1	2

3	1
2	1
3	2
4	2
2	1
2	1
2	1
2	1
2	1
5	2
3	2
2	1
4	3
Avg 3	Avg 6
2.820512821	1.666666667

Appendix B

Class Year:

What classes have you taken, that have used Mastering Physics by Pearson?

Rate your overall satisfaction with Mastering physics

- ❏ **1 Highly Dissatisfied**
- ❏ **2 Dissatisfied**
- ❏ **3 Indifferent**
- ❏ **4 Satisfied**
- ❏ **5 Highly satisfied**

What do you like about Mastering Physics (Please be specific with your answers)?

What do you dislike about Mastering Physics?

Please rate how you feel Mastering Physics helped to enhance your understanding of the material provided by the professor

- ❏ **1 Less than other homework (such as problems from a text book)**
- ❏ **2 As much as other homework**
- ❏ **3 Even more than other homework**

Comment on your reasoning for your above rating:

What do you think of the way problems are presented?

Any additional comments that you would like to add about Mastering Physics

Appendix C

Data Record for PSA3Q3X - Logs prior to April 26, 2017

Dear Researcher,
Welcome to the data record for problem set PSA3Q3X. You have received this record based on your recent data request. Automated data analysis is featured below, offering a preliminary overview of your sample and a selection of analyses for your consideration. The latter portion of this report contains the raw data files from which you can conduct your own thorough analyses. When publishing your work, please reference this report as a stable location for readers to access your data for review and replication.

By clicking any link to download content from this page, you are agreeing to our Terms of Use.

Automated Data Analysis

Completion Rates
Students that have started PSA3Q3X : 18
Students that have completed PSA3Q3X : 7

Bias Assessment
Before analyzing learning outcomes, we suggest first assessing potential bias introduced by your experimental conditions (i.e., examine differential dropout). The table below reports the number of students that have completed PSA3Q3X, split out by experimental condition.

Conditions	Students who started the problem set	Students who finished the problem set	Percent of students completed

control	18	7	39
treatment	7	1	14
N	25		
Degrees Of Freedom	1		
Chi-Squared	0.5		
P-Value	0.48		

Mean and Standard Deviation of Posttest Score by Condition
To examine learning outcomes at posttest, an analysis of means was conducted across conditions. The tables below reports mean post-test score and standard deviation for each condition. This information was sourced from our automated post-test subreport.

	Students who completed the problem set	Percent of students completing the problem set	PostTestScore*
control	0	0	NaN (0.0)
treatment	0	0	NaN (0.0)

* Presented as Mean (SD)

Appendix D

Data Record for PSA3QYY - Logs prior to April 26, 2017

Dear Researcher,
Welcome to the data record for problem set PSA3QYY. You have received this record based on your recent data request. Automated data analysis is featured below, offering a preliminary overview of your sample and a selection of analyses for your consideration. The latter portion of this report contains the raw data files from which you can conduct your own thorough analyses. When publishing your work, please reference this report as a stable location for readers to access your data for review and replication.

By clicking any link to download content from this page, you are agreeing to our Terms of Use.

Automated Data Analysis

Completion Rates
Students that have started PSA3QYY : 13
Students that have completed PSA3QYY : 3

Bias Assessment
Before analyzing learning outcomes, we suggest first assessing potential bias introduced by your experimental conditions (i.e., examine differential dropout). The table below reports the number of students that have completed PSA3QYY, split out by experimental condition.

Conditions	Students who started the problem set	Students who finished the problem set	Percent of students completed

control	9	1	11
treatment	3	2	67
N	12		
Degrees Of Freedom	1		
Chi-Squared	1.33		
P-Value	0.25		

Mean and Standard Deviation of Posttest Score by Condition
To examine learning outcomes at posttest, an analysis of means was conducted across conditions. The tables below reports mean post-test score and standard deviation for each condition. This information was sourced from our automated post-test subreport.

	Students who completed the problem set	Percent of students completing the problem set	PostTestScore*
control	1	11	0.17 (0.35)
treatment	2	67	0.33 (0.58)

* Presented as Mean (SD)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)

Condition	1	0.06	0.06	0.37	0.55
Residuals	10	1.67	0.17	NaN	NaN

Appendix E

Data Record for PSA3HGT - Logs prior to April 26, 2017

Dear Researcher,
Welcome to the data record for problem set PSA3HGT. You have received this record based on your recent data request. Automated data analysis is featured below, offering a preliminary overview of your sample and a selection of analyses for your consideration. The latter portion of this report contains the raw data files from which you can conduct your own thorough analyses. When publishing your work, please reference this report as a stable location for readers to access your data for review and replication.

By clicking any link to download content from this page, you are agreeing to our Terms of Use.

Automated Data Analysis

Completion Rates
Students that have started PSA3HGT : 10
Students that have completed PSA3HGT : 1

Bias Assessment
Before analyzing learning outcomes, we suggest first assessing potential bias introduced by your experimental conditions (i.e., examine differential dropout). The table below reports the number of students that have completed PSA3HGT, split out by experimental condition.

Conditions	Students who started the problem set	Students who finished the problem set	Percent of students completed

control	1	0	0
treatment	3	0	0
N	4		
Degrees Of Freedom	1		
Chi-Squared	NaN		
P-Value	NaN		

Mean and Standard Deviation of Posttest Score by Condition
To examine learning outcomes at posttest, an analysis of means was conducted across conditions. The tables below reports mean post-test score and standard deviation for each condition. This information was sourced from our automated post-test subreport.

	Students who completed the problem set	Percent of students completing the problem set	PostTestScore*
control	0	0	NaN (0.0)
treatment	0	0	NaN (0.0)

* Presented as Mean (SD)