Designing and Assessing a Learning Environment to Support Mechanical Reasoning

Ann McKenna, Alice Agogino

Northwestern University (mckenna@northwestern.edu)/
University of California at Berkeley (aagogino@me.berkeley.edu)

Abstract

This paper describes and assesses a learning environment designed to support mechanical reasoning and understanding of simple machines. Based on recommendations from the literature on instructional frameworks and cognitive aspects of mechanical reasoning, SIMALE (the Simple Machines Learning Environment) was designed to support reflection, collaboration, and presentation of concepts from multiple perspectives and contexts. SIMALE was implemented with middle and high school students with three treatment variations: (1) environment with focus on Lego exercises to engage in hands-on physical activities, (2) environment with focus on a web-based computer module, and (3) environment with both the computer module and Lego exercises. Learning was measured in three categories: analytic problem solving, conceptual understanding, and drawing and modeling ability. The assessment found that students significantly increased their understanding in all three categories for all treatment variations within SIMALE. The results revealed unexpected dramatic results in equalizing post-test scores, in spite of large population differences in pre-test scores. A complete description of the study, the assessments and the statistical analyses are presented. Based on these findings we present recommendations for creating educational experiences and environments that support development and application of simple machines concepts.

Introduction

A concept that is encountered in many science classrooms as well as in every day life is how mechanical advantage is gained through simple machines. One is confronted with simple machines and mechanical advantage when playing on a seesaw, changing gears in a car or on a bicycle, or prying open a can. Furthermore, concepts of mechanical advantage, force and motion, and energy principles are common subject matter requirements listed in many state science standards. One example from the New York science standards is given below.

> New York Standard 4-Physical Science. Energy and matter interact through forces that result in changes in motion. Use simple machines made of pulleys and levers to lift objects and describe how each machine transforms the force applied to it[1].

Concepts of mechanical advantage, particularly force, motion, and energy, are also addressed in college physics and basic engineering mechanics classes[2,3]. These basic concepts are used as a foundation to build more formal understanding and analysis of complex mechanisms. Since mechanical advantage and simple machines are common topics encountered in daily life, middle and secondary school classrooms, as well as in fundamental physics and engineering courses, our work provides a learning environment where students can explore, analyze, and integrate these concepts.

We designed the simple machines learning environment (SIMALE) to support students' mechanical reasoning and understanding. Based on recommendations from learning theory and instructional frameworks, the simple machines learning environment was designed to support reflection and collaborative learning, and to engage students in generative learning through multiple representations of concepts and successive experimentation and design activities.

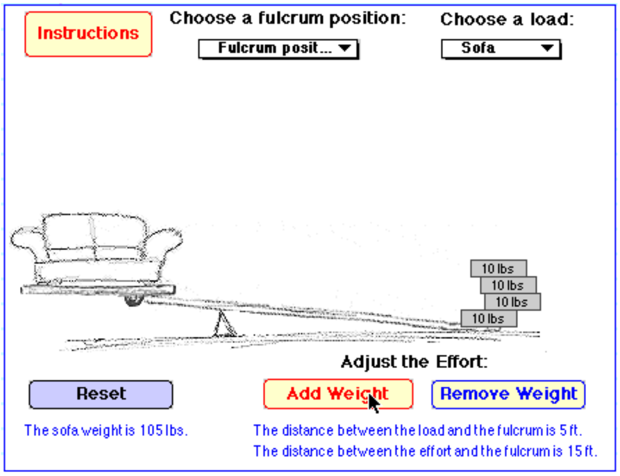
Two key components of SIMALE are an original web-based software tool and hands-on Lego activities. Features of the simple machines computer module allow students to perform simulations of lever and pulley devices and observe the response of a system based on inputs, engage in self-reflection and sharing of ideas, and to interact with concepts from multiple perspectives such as text, simulations, and mathematical plots[4,5]. Students use the computer module and/or Lego sets to test their ideas and investigate simple machines concepts. Based on the information obtained from the computer simulations and Lego experimentation students apply this knowledge to solve lever and pulley design problems.

We designed a research study consisting of three treatment groups to investigate the differential benefits of variations within SIMALE –hands-on and web-based computer activities – on students' analytic problem solving ability, drawing/modeling ability, and conceptual understanding. This paper provides an overview of SIMALE, describes the assessment study, and provides a summary of the results.

Overview of SIMALE

In this section we describe the learning goals of SIMALE as well as the instructional theories upon which the learning environment was built. Two primary learning goals were established in developing SIMALE. First, SIMALE was designed to foster development of simple machines concepts. This includes learning about the names and functions of each of the devices and their parts, and recognizing common uses for each of the simple machines. Middle school, high school, and college students are likely to have experience with simple machines since they are ubiquitous in everyday life. Therefore, this experiential knowledge is a basis upon which students can build more comprehensive and robust understanding of simple machines. Students investigate the lever and pulley through simulations, plotting, Lego building, problem solving and discussions (see examples in Figs. 1 and 2). The combination of these activities familiarizes the students with the basic concepts of simple machines[4,5].

Second, the environment was designed to encourage students to make connections among the physical devices, the mathematical analysis, and appropriate applications. The learning environment provides different representations of the working devices. The different



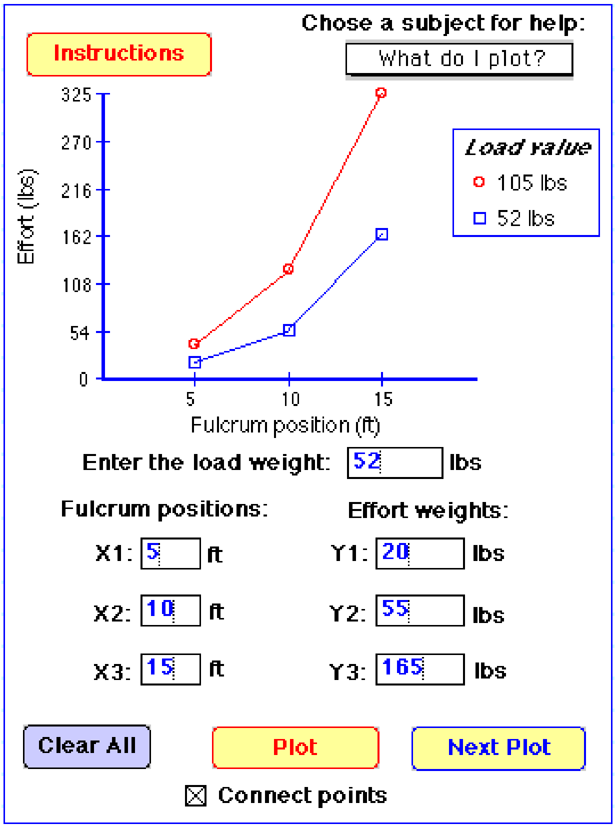
The sofa weight is 105 lbs.

The distance between the load and the fulcrum is 5 ft.
The distance between the effort and the fulcrum is 15 ft.

Fig. 1. One representation for a lever exercise[4].

representations range in their level of abstraction from concrete physical devices to abstract mathematical equations and graphs. Students engage in simulation activities and/or Lego building to illustrate the relationship between distance and effort. Results obtained from the simulations are used to make mathematical plots of the data in the computer module (Fig. 2). The simulations and the mathematics provide different representations to illustrate how the concept of mechanical advantage is modeled by each of the devices. Through in-class discussions students are introduced to force equations (an abstract representation) which define the relationship between force and distance for the lever and the pulley. These simulations allow students to solve analytic problems and make predictions about various lever and pulley situations. Worksheets were constructed to focus students on connecting the ideas represented through the repertoire of models presented in the computer module, the Lego building, and the class discussions. The combination of these activities serves to provide multiple opportunities for students to connect the concepts to applications.

Four main findings from instructional and learning theory guided the pedagogy of SIMALE. Specifically, the literature suggests that a learning environment should 1) provide opportunities for students to actively participate (interactivity and collaborative learning), 2) support self-reflection, 3) provide multiple representations of concepts, and 4) cultivate generative learning. Each of these principles is described in more detail.



Enter the load weight: 52 lbs

Fulcrum positions:
X1: 5 ft
X2: 10 ft
X3: 15 ft

Effort weights:
Y1: 20 lbs
Y2: 55 lbs
Y3: 165 lbs

Fig. 2. Simulation and plotting for a lever exercise[4].

In the most general sense, the notion of an interactive environment implies that the learner plays an active role while working within the environment. In contrast to the idea of the learner as a ‘passive recipient of knowledge’, an interactive environment encourages students to actively participate in the learning process. The principle of interactiveness is derived from various frameworks found in the literature. For example, Koschmann et al.[6] describe a principled approach to designing computer based tools for use in collaborative learning. They explain that learning is an active process, requiring mental construction on the part of the learner. Various features were included in SIMALE to promote active engagement of the learner. For example, the simulations and plotting applications in the simple machines web module (Figs. 1 and 2) and the independent design activities foster cognitive initiative and promote interactivity and collaborative learning.

Self-reflection and the development of metacognitive skills are common themes in many instructional frameworks. Many researchers have shown the benefits of self-reflection in order to make "thinking visible" as well as to create a sense of community[7, 8, 9]. For example, self reflection is a central part of the framework for the Knowledge Integration Environment (KIE) developed by Linn and coworkers[10]. Students use tools in KIE to provide evidence and structure arguments, and to provide social support for students as they formulate and defend an argument. These tools implicitly promote self-reflection since students are required to organize, represent, and justify their thinking.

Self-reflection activities are embedded in SIMALE in a variety of ways. The web-based computer software contains a feature called the 'share findings page[4,5]**.** This page provides a space for the students to share their discoveries with the class. Having the students report their findings to the class accomplishes two main purposes: 1) it contributes to the collective knowledge of the class by making each student's thinking visible and, 2) it encourages the students to self-reflect on their learning. Students are encouraged to make their thinking and learning visible and self-reflect throughout the environment by engaging in collaborative activities, performing experiments and discussing results, and creating unique designs for open-ended problems.

Multiple representations are presented throughout SIMALE. For example, the simple machines web-based module was designed to include simulations, animations, sound, video, and feedback to input. The Lego building activities also provide an additional, more concrete, representation of the lever and pulley. Since students may have individual learning styles, the use of multiple representations increases the repertoire of models for students. According to Resnick and Ocko[11] the learning environment should offer multiple paths to learning that allow students the freedom to approach projects from different perspectives. Providing multiple representations also relates to Koschmann's principle of multiplicity. Since knowledge is complex, dynamic, and context sensitive, this principle suggests that multiple representations, perspectives and strategies should be provided and revisited as knowledge matures[6].

The application of the term 'generative' to SIMALE is consistent with the basic definition of the word. Generative implies that something is created, revisited, and capable of being reproduced. In a learning environment, concepts should not be encountered just one time, but revisited and applied in multiple contexts. As Resnick and Resnick[12] note, "for concepts and organizing knowledge to be mastered, they must be used generatively – that is, they have to be called upon over and over again as ways to link, interpret, and explain new information."

The simple machines module (http://socrates.berkeley.edu:7009/simple_machines) provides multiple contexts in which students encounter the concepts. The computer environment provides various levels of abstraction of lever and pulley devices through simulations, and textual and mathematical explanations. In addition, the learning environment presents problem-solving situations where the concepts can be applied. More information on the simple machines module can be found in McKenna & Agogino (1998)[4] . The idea of using grounded, reality-based problems has been researched by the Cognition and Technology Group at Vanderbilt[13]. The Vanderbilt group, through the Jasper series, found that using 'real-life' engaging problems led to student understanding of relevant mathematics concepts. The Jasper series is also an example of how technology can be structured to provide a generative and active learning experience. Based on these design principles, the environment provides multiple opportunities for the student to revisit concepts and, more importantly, to use the concepts to solve 'real life' problems. In the simple machines learning environment, the problem solving tasks can also be completed with the use of physical materials. The physical materials (Legos) provide another opportunity to "link, interpret, and explain" ideas.

In this section we addressed the learning goals of SIMALE and the learning principles that guided the design and pedagogy of the environment. The following section describes the experimental design and procedure of the study.

Experimental Design and Procedure

Two primary resources were used within SIMALE to help students develop mechanical reasoning, an original web-based computer tool and Lego sets. In order to investigate the differential benefits of hands-on and web-based computer activities on students' analytic problem solving ability, drawing/modeling ability, and conceptual understanding three treatment groups were created. One treatment group used just the physical objects, or Lego sets, as a resource to explore the principle of the lever and pulley. Another condition used just the simple machines web-based computer module and the third condition used both the computer module and the Lego sets as resources. All three groups had the same instructional material about simple machines in either hard copy or electronic form, depending on the treatment. The three conditions will be referred to as the 'Lego' group, the 'Computer' group, and the 'Both' group, respectively.

Pre and post-tests were administered and compared among the three conditions. These tests included analytic questions, conceptual questions, and open-ended design problems that required students to produce a model/drawing of their proposed solution. Using these data sets, we compare the performance among the three conditions. In this paper we focus on the results from the overall pre and post-test responses. That is, we do not separate the data along different categories of learning (analytic, conceptual, and drawing ability) in this paper. Future work will describe these findings.

Subjects

Participants for the research study were invited from two summer outreach programs that are offered through the University of California at Berkeley. The two summer programs are the Academic Talent Development Program (ATDP) and the Mathematics, Engineering, and Science

Achievement (MESA) program. Both programs offer science, mathematics, and technology summer courses to middle and high school age students.

A letter was sent to students inviting them to participate in a free Saturday workshop to learn about simple machines. All students who responded to the invitation were accepted and were assigned to one of the three conditions. Since this was a pseudo-random way of assigning students to a condition, significance tests were performed on pre-test scores to ensure no pre-existing differences among students in each condition.

Thirty-seven students participated in the study from the MESA program. Ten students participated in the 'Lego' condition, 16 students in the 'Computer' condition, and 11 students in the 'Both' condition. Students ranged in age from 12-17 years and from grades 7th-12th. Of the 37 students, 59% were male and 41% female. Since the MESA program is an outreach program that serves minorities and under represented groups in science and mathematics, the MESA participants formed a diverse group of students. Approximately 32% of the MESA participants were African-American, 32% Chicano/Latino, and 30% Asian-American.

Forty students participated in the study from the ATDP program. Sixteen students participated in the 'Lego' condition, 9 students in the 'Computer' condition, and 15 students in the 'Both' condition. Students ranged in age from 13-15 years and from grades 7th-9th. Of the 40 students, 58% were male and 42% female. The gender breakdown is very similar for both the MESA and ATDP participants. The ATDP program is an academically competitive program that serves students in the San Francisco Bay Area, Northern California and the California Central Valley. Approximately 5% of the ATDP participants were Native American, 8% African-American, 10% East Indian/Pakistani, 12% Chicano/Latino, 18% Caucasian, and 40% Asian-American.

All of the students in each condition received the same set of classroom instruction and handouts, but the actual delivery varied depending on the treatment. Each class followed the same sequence of activities and each lasted 5 1/2 hours. Pre and post-tests were administered during this time and students were also given a half-hour lunch break. Allowing for tests and breaks, students spent approximately four hours working on the activities for the lever and the pulley.

Overall Lever and Pulley Assessments

Students in all three conditions were given paper-and-pencil pre and post-tests on the simple machines concepts addressed in the class (see Appendix A and B). Both the pre and post-tests were intentionally created to contain identical items. The strength of identical pre and post-tests is that the items are matched so that one can clearly measure improvement across items. Students are asked the exact same questions so we can assume equivalence in terms of interpretation and difficulty.

Questions on the overall lever and pulley tests were based on items from the Test of Mechanical Comprehension[14], literature on mechanical reasoning[15], state standards on simple machines, and the objectives of the curriculum. Each item on the test also asks students to provide an explanation for his or her answer. Since some of the items are forced choice where students select an answer 'a-d', the explanation prompt is used to probe for student reasoning. Conceptual understanding is gauged by the quality of explanation students provide to each question. The data analyses used a combination of quantitative and qualitative methods since the pre and post-tests contained a combination of forced choice and free response questions. We used a blind scoring process such that the researcher did not know to which group the test belonged.

To compare the effectiveness of the resources provided in the three different treatment groups t-tests and analyses of variance (ANOVA) were used. T-tests allow us to test for any significant differences among the three treatment groups, both before and after the intervention. ANOVA were performed to test for effect on post-test scores due to treatment group controlling for three between subject factors: pre-test scores, condition (Lego, Computer, or Both), and gender. The following section presents the results from the overall lever and pulley tests for each condition.

Results

Results from the MESA and ATDP populations are presented for the overall lever and pulley tests. The overall pre and post-test consisted of 8 questions each for the lever and the pulley. Each item was given one point for a correct answer or no point for an incorrect response. The range of scores therefore varied from 0 to 8. The items on the overall test consisted of analytic and conceptual questions with explanation prompts.

Pulley Results Overall: MESA Students

Unpaired sample t-tests were performed to test for differences between the three groups prior to the intervention. All three groups performed equally well on the pulley pre-test and no significant differences were found between the Lego and Computer group, t(22) = 1.066, p = .14, the Lego and Both group, t(17) = 1.18, p = .10, or the Computer and Both group, t(25) = .119, p = .83. Table 1 and Fig. 3 show the comparison of overall pre and post-test pulley scores. The data indicate significant gains in post-test performance for all three groups.

Condition	Pre (SD)	Post (SD)	Mean Diff.	N	t-value	p-value
LEGO	2.0(1.8)	6.2(1.7)	4.2	8	5.19	<.01
Computer	3.1(1.5)	6.2(.78)	3.2	16	7.96	<.01
Both	3.2(1.2)	6.1(1.1)	2.9	11	7.02	<.01
Total	2.8(1.5)	6.2(1.1)	3.4	35	11.31	<.01

Table 1. Comparison of overall pre and post-test pulley scores (max score=8): MESA.

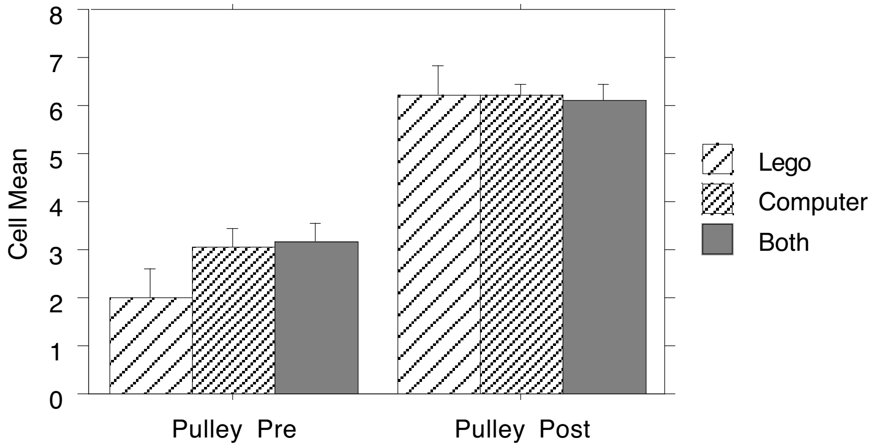


Fig. 3. Plot of pulley pre and post-test scores by experimental group: MESA.

Analysis of variance (ANOVA) was then performed to test for interactions between performance on the pulley post-test and three between-subject factors: treatment (Lego, Computer, Both) and pre-test. The ANOVA on pulley post-test scores revealed no effect of type of treatment within the environment (F=.267, p = .76), gender (F= .02, p = .88), or pulley pre-test (F= 1.2, p = .28), and no interactions were found, see Table 2.

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Treatment	2	.754	.377	.267	.7678	.535	.086
Gender	1	.029	.029	.020	.8879	.020	.052
Pulley Pre	1	1.705	1.705	1.209	.2829	1.209	.175
Treatment * Gender	2	3.814	1.907	1.353	.2783	2.706	.253
Treatment * Pulley Pre	2	.363	.182	.129	.8797	.258	.067
Gender * Pulley Pre	1	.454	.454	.322	.5759	.322	.083
Treatment * Gender * Pulley Pre	2	6.055	3.028	2.148	.1396	4.295	.384
Residual	23	32.423	1.410				

Table 2. ANOVA for pulley post-test scores: MESA.

Lever Results Overall: MESA Students

Unpaired sample t-tests show all three groups performed equally well on the lever pre-test and no significant differences were found between the Lego and Computer group, t(22) = .037, p = .97, the Lego and Both group, t(18) = 1.007, p = .33, or the Computer and Both group, t(22) = 1.14, p = .27. Table 3 and Fig. 4 show the comparison of overall pre and post-test lever scores. Similar to the pulley analyses, the data indicate significant gains in post-test performance for all three groups.

Condition	Pre (SD)	Post (SD)	Mean Diff.	N	t-value	p-value
LEGO	4.3(.95)	6.7(1.2)	2.4	10	5.31	<.01
Computer	4.3(.91)	6.8(1.1)	2.5	14	7.27	<.01
Both	4.7(.82)	6.3(1.2)	1.6	10	4.00	<.01
Total	4.4(.89)	6.6(1.2)	2.2	34	9.57	<.01

Table 3. Comparison of overall pre and post-test lever scores (max score=8): MESA.

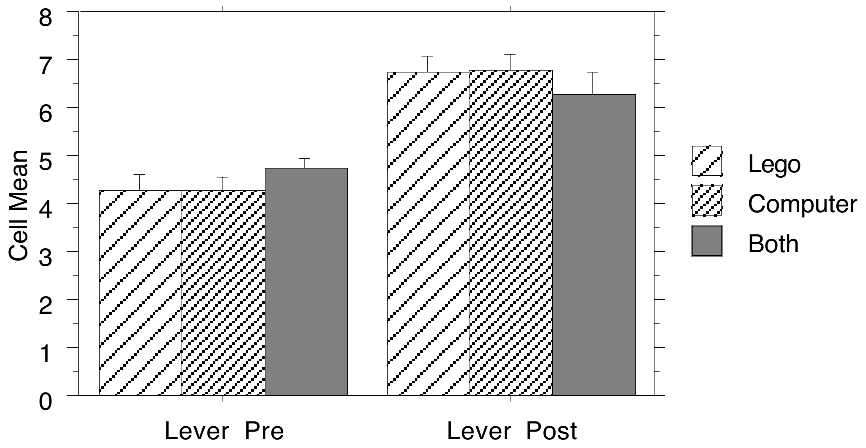


*Fig. 4. Plot of lever pre and post-test scores by experimental group: ME*SA.

The ANOVA on lever post-test scores also revealed no effect of type of treatment within the environment (F = .061, p = .94), gender (F = .182, p =.67), or lever pre-test (F = .42, p =.52), and no interactions were found, see Table 4.

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Treatment	2	.186	.093	.061	.9407	.123	.058
Gender	1	.276	.276	.182	.6739	.182	.069
Lever Pre	1	.632	.632	.417	.5253	.417	.093
Treatment * Gender	2	.437	.219	.144	.8667	.288	.069
Treatment * Lever Pre	2	.354	.177	.117	.8905	.233	.065
Gender * Lever Pre	1	.136	.136	.089	.7678	.089	.059
Treatment * Gender * Lever Pre	2	.956	.478	.315	.7330	.630	.093
Residual	22	33.396	1.518				

Table 4. ANOVA for lever post-test scores:MESA.

The ANOVA data suggest some interesting and informative findings. There is no effect (on lever and pulley post scores) due to type of treatment within the environment. This suggests that the students who used just the computer module did as well as students who received both legos and the computer, or just lego materials within the simple machines environment. These data also indicate that there is no effect due to pre-test scores. One might have expected to see an effect due to prior knowledge but this finding contradicts this assumption. The fact that there is no effect of pre-test scores suggests that the simple machines environment, with each treatment group, is similarly beneficial to students regardless of prior knowledge.

Pulley Results Overall: ATDP Students

Unpaired sample t-tests were performed to test for differences between the three ATDP groups prior to the intervention. All three groups performed equally well on the pulley pre-test and we found no significant differences between the Lego and Computer group, t(23) = -.842, p = .408, the Lego and Both group, t (29) = .469, p = .64, or the Computer and Both group, t(22) = 1.187, p = .249. Table 5 and Fig. 5 show the comparison of overall pre and post-test pulley scores. These data indicate significant gains in post-test performance for all three groups. These results are consistent with the MESA data as shown in Table 1.

Condition	**Pre (SD)**	**Post (SD)**	**Mean Diff.**	**N**	**t-value**	**p-value**
LEGO	3.8(1.9)	6.6(1.7)	2.8	16	5.12	<.01
Computer	4.4(1.5)	6.5(1.7)	2.1	9	4.36	<.01
Both	3.5(2.2)	6.9(1.1)	3.4	15	5.59	<.01
Total	3.8(1.9)	6.7(1.4)	2.9	40	8.55	<.01

Table 5. Comparison of pulley pre and post-test scores (max score=8): ATDP.

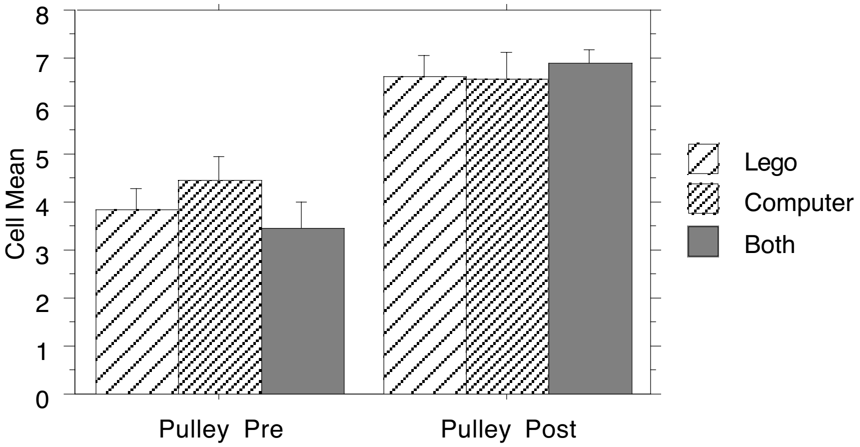


Fig. 5. Plot of pulley pre and post-test scores by experimental group: ATDP.

An ANOVA was also performed for the pulley scores. The ANOVA on ATDP pulley post-test scores revealed no effect of treatment (F =1.19, p = .318), gender (F= .087, p = .77), or pulley pre-test (F = 3.64, p = .06), and no interactions were found, see Table 6.

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Treatment	2	5.739	2.869	1.194	.3180	2.388	.232
Gender	1	.209	.209	.087	.7702	.087	.059
Pulley Pre	1	8.758	8.758	3.644	.0666	3.644	.440
Treatment * Gender	2	2.808	1.404	.584	.5642	1.169	.134
Treatment * Pulley Pre	2	2.733	1.367	.569	.5727	1.137	.132
Gender * Pulley Pre	1	.032	.032	.013	.9085	.013	.051
Treatment * Gender * Pulley Pre	2	3.003	1.502	.625	.5427	1.250	.140
Residual	28	67.290	2.403				

Table 6. ANOVA for pulley post-test scores: ATDP.

Lever Results Overall: ATDP Students

Unpaired sample t-tests were performed to test for differences on the overall lever pre-test. All three groups performed equally well on the lever pre-test and no significant differences were found between the Lego and Computer group, t (22) = .777, p = .44, the Lego and Both group, t (29) = -.208, p = .83, or the Computer and Both group, t (21) = -.902, p = .37. Table 7 and Fig. 6 show the comparison of overall pre and post-test lever scores. Using a significance level of p < .05, the data indicate statistically significant gains in post-test performance for all three groups. These results are consistent with the MESA data given in Table 3.

Condition	**Pre (SD)**	**Post (SD)**	**Mean Diff.**	**N**	**t-value**	**p-value**
LEGO	5.6(1.4)	7.1(1.0)	1.5	16	3.98	<.01
Computer	5.1(1.6)	6.6(1.4)	1.5	8	2.81	.026
Both	5.7(1.5)	7.0(1.2)	1.3	15	4.75	<.01
Total	5.6(1.5)	7.0(1.2)	1.4	39	6.71	<.01

Table 7. Comparison of lever pre and post-test scores (Max score=8): ATDP.

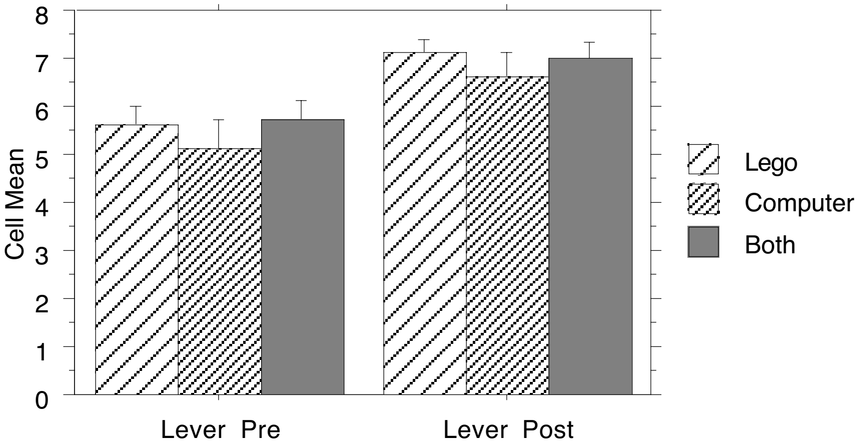


Fig. 6. Plot of lever pre and post-test scores by experimental group: ATDP.

An analysis of variance was then performed to test for interactions between performance on the lever post-test and three between-subject factors: treatment (Computer, Lego, Both), gender, and pre-test. Similar to the analyses of the MESA data, the ANOVA on lever post-test scores revealed no effect of treatment (F = 2.336, p = .12), or gender (F = 2.113, p =.157), and no interactions were found, see Table 8. However, the ANOVA did reveal a statistically significant effect of lever pre-test (F = 17.94, p < .01). This indicates an effect of prior knowledge on lever post-test performance. More on this will be discussed in the MESA vs. ATDP Comparison section.

	DF	Sum of Squares	Mean Square	F-Value	P-Value	Lambda	Power
Treatment	2	4.232	2.116	2.336	.1160	4.672	.421
Gender	1	1.913	1.913	2.113	.1576	2.113	.274
Lever Pre	1	16.254	16.254	17.945	.0002	17.945	.991
Treatment * Gender	2	.920	.460	.508	.6075	1.015	.122
Treatment * Lever Pre	2	3.793	1.897	2.094	.1428	4.188	.381
Gender * Lever Pre	1	.424	.424	.468	.4998	.468	.098
Treatment * Gender * Lever Pre	2	.730	.365	.403	.6722	.806	.107
Residual	27	24.456	.906				

Table 8. ANOVA for lever post-test scores:ATDP.

MESA vs. ATDP Comparison: Pulley

The ATDP overall pulley results were compared to the MESA results. MESA students scored consistently lower on the pulley pre-test than the ATDP students. T-tests revealed a statistically significant difference between the MESA and ATDP students on the pulley pre-test, t (73) = -

2.39, p = .02. However, post-test scores for the MESA and ATDP students were similar and t-tests revealed no significant differences in student post-test performance, t (73) = -1.66, p = .10. Figs. 7 and 8 show scatterplots of pre and post-test pulley scores, respectively, sorted by population.

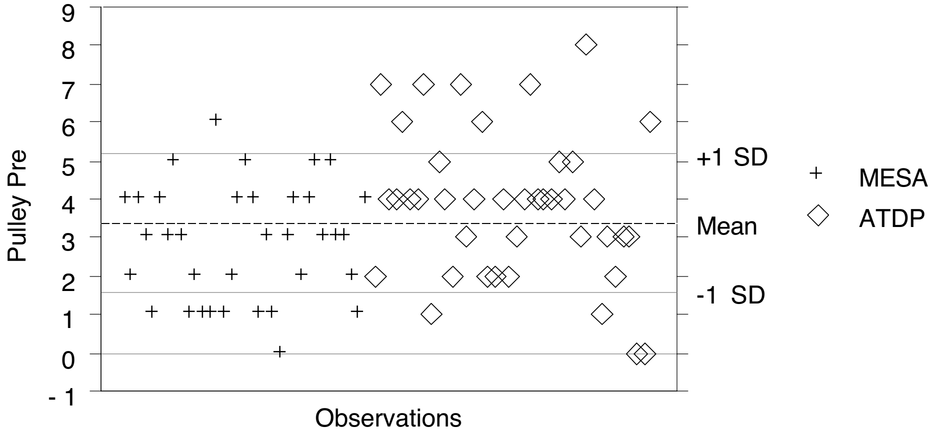


Fig. 7. Scatterplot of pulley pre-test scores sorted by program (MESA, n =35, ATDP, n =40).

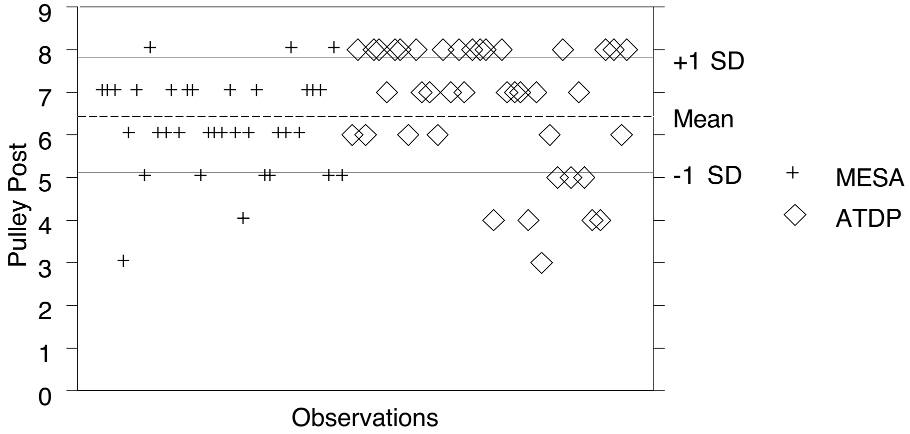


Fig. 8. Scatterplot of pulley post-test scores sorted by program.

These two scatterplots provide a visual display of the differences in scores between the two student populations. On the pre-test we see that some ATDP students approached or achieved the maximum score while none of the MESA students achieved this level. However, Fig. 8 shows a comparable distribution of post-test scores for the two populations, with both groups including students who approached or achieved the maximum score. The distribution of scores in Fig. 8 indicates that more ATDP students than MESA students achieved the maximum score but the post-test data are equivalent otherwise. This is a dramatic result that is replicated in the Lever analysis as well.

MESA vs. ATDP Comparison: Lever

The lever results were also compared between the ATDP and MESA data. The average MESA lever scores were consistently lower than the ATDP students' scores on the pre-test. T-tests revealed a significant difference between the MESA and ATDP students on the lever pre-test, t (71) = -3.98, p <.01. While there was a statistically significant difference between ATDP and MESA students on the lever pre-test, post-test scores for the MESA and ATDP students were similar and t-tests revealed no significant differences on student post-test performance, t (71) = -1.3, p = .19. The data from the overall pulley and lever analyses indicate that MESA students entered with less prior knowledge about simple machines but achieved greater gains in performance compared to ATDP students.

Figs. 9 and 10 show scatterplots of pre and post-test lever scores, respectively, sorted by population. On the pre-test we see that some ATDP students approached or achieved the maximum score while none of the MESA students achieved this level. However, Fig. 10 shows an equivalent distribution of post-test scores for the two populations, with both groups including students who approached or achieved the maximum score. The distribution of pre-test scores in Fig. 9 show a similar trend as the pulley data in Fig. 7. Some ATDP students approached or achieved the maximum score while none of the MESA students achieved this level. This is most likely the main factor which contributed to the statistical difference found between MESA and ATDP on pre-test scores.

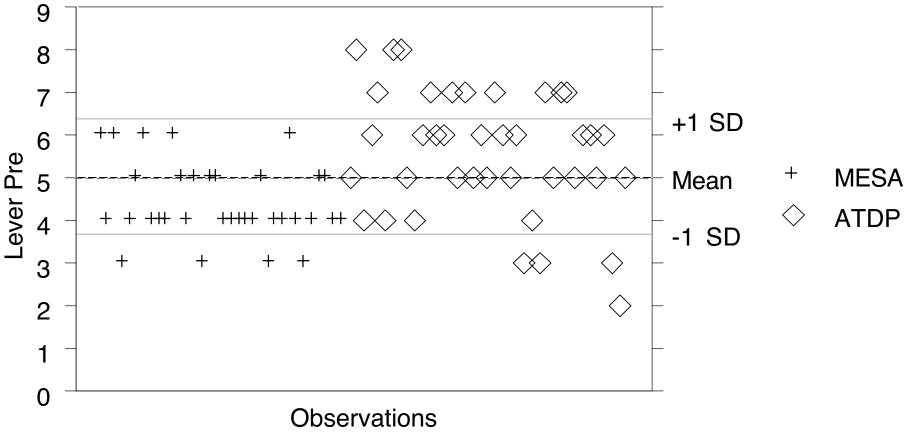


Fig. 9. Scatterplot of lever pre-test scores sorted by program (MESA, n =34, ATDP, n =39).

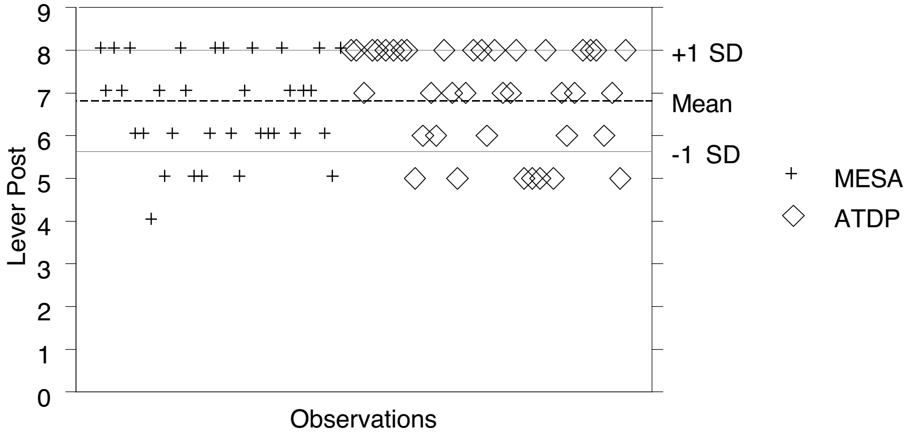


Fig. 10. Scatterplot of lever post-test scores sorted by program.

One might argue that there is a ceiling effect on the lever and pulley post-tests. Based on the distribution of post-test scores displayed in Figs. 8 and 10, some students did achieve the maximum score of eight. However, the majority of scores remained below the maximum. This may suggest a small ceiling effect but since both groups performed equally well on the post-test it would apply to both the MESA *and* ATDP students. Perhaps both groups of students approached the maximum level of performance that the post-test was capable of measuring. A possible ceiling effect does not diminish the compelling result that MESA students achieved such gains in performance.

Discussion

A review of the data from the MESA and ATDP overall lever and pulley pre and post-tests show a consistent trend. Students performed significantly better on post-test measures, regardless of treatment group. Given these overwhelming positive findings it is useful to re-examine the features of SIMALE that may have contributed to these results. During the implementation of SIMALE, regardless of treatment group, students engaged in collaborative and reflective learning and were provided multiple opportunities to interact with concepts. Students encountered simple machines concepts through multiple representations and perspectives and engaged in iterative experiments to test ideas. Based on the results from the study the combination of these activities led to significant improvement in performance in mechanical reasoning and problem solving ability. The improvements, regardless of treatment group and focus (computer simulation, Legos, both), show the strength of the environment and the associated instructional design, leaving much flexibility for the instructor in its implementation depending on teaching preferences and classroom infrastructure.

The mechanism by which students refine their understanding, and develop a more robust understanding of the concepts is what we seek to explore. Based on the instructional framework developed for SIMALE we can extract general principles that foster a productive learning experience. Within SIMALE, collaborative activities served multiple purposes. Students worked with partners to share and test ideas, clarify perspectives, and discuss points of confusion. This process encouraged students to reflect on their understanding and enabled them to make their thinking explicit. Students revealed their thinking through verbal discussions, graphical representations, and written explanations. These tangible expressions of tacit understanding were visible products that could be debated and refined within SIMALE.

Summary and Future Work

This paper provided an overview of SIMALE by describing the learning goals of the environment and discussing the theoretical principles that guided the design of the pedagogy. SIMALE was designed to help students develop concepts about simple machines and make connections among the physical devices, the mathematical analysis, and appropriate applications. The environment included activities to provide opportunities for students to actively participate, support self-reflection, provide multiple representations of concepts, and cultivate generative learning. These four principles, taken from instructional and learning theory, embody the pedagogy of SIMALE.

SIMALE included two primary resources for students to investigate the lever and pulley, the simple machines web-based module and Lego sets. A study was conducted to investigate the benefits of these two resources in supporting students' mechanical reasoning. Three treatment groups were created, one used just the computer module, one used just the Lego sets, and one used both resources within the simple machines environment. Subjects were recruited from two outreach programs at the University of California at Berkeley. The subjects in this study formed a diverse population with respect to gender, ethnicity, and academic achievement.

Statistical analyses showed significant gains in lever and pulley post-test performance for both the MESA and ATDP populations, in spite of major differences in pre-test scores. ANOVA revealed no effect of type of treatment within SIMALE for both tests (lever and pulley) and both populations (MESA and ATDP). That is, students who used the computer module within SIMALE performed as well as those that received hands-on activities or both resources within SIMALE. Results from this study show SIMALE supported a collaborative, reflective, and generative learning environment. Furthermore, SIMALE clearly contributes to students' mechanical reasoning and understanding of simple machines concepts for a diverse population of students. The results are quite promising in terms of showing the strength of the environment and the associated instructional design, which leaves much flexibility for the instructor in its implementation depending on teaching preferences and classroom infrastructure. As a consequence, SIMALE has been used effectively in a range of instructional settings with a diversity of instructional styles.

The current paper discussed results from the overall pre and post lever and pulley tests of the simple machines learning environment. Future work will address analyses along different categories of learning: analytic problem solving ability, conceptual understanding, and drawing/modeling ability. Gender differences along these dimensions will also be examined.

References

1. NY Learning Standards for Mathematics, Science, and Technology (1996). http://www.emsc.nysed.gov/ciai/pub.html#cat4

2. Beer, F. P. and Johnston, E. R. (1988). *Vector Mechanics for Engineers*, 5th Edition, NY: McGraw-Hill.

3. Meriam, J. L. and Kraige, L.G. (1997). *Engineering Mechanics, Dynamics*, 4th Edition, NY: John Wiley & Sons, Inc.

4. McKenna, A. and Agogino, A. (1998), A Web-Based Instructional Module for Teaching Middle School Students Engineering Design with Simple Machines, *Journal of Engineering Education*, 87 (4), 437-444.

5. McKenna, Ann Frances (2001). *Designing Instruction to Support Mechanical Reasoning: Three Alternatives in the Simple Machines Learning Environment*, Doctoral Dissertation, University of California, Berkeley.

6. Koschmann, T. D., Myers, A. C., Feltovich, P. J., and Barrows, H. S. (1994). Using Technology to Assist in Realizing Effective Learning and Instruction: A principled Approach to the Use of Computers in Collaborative Learning. *The Journal of the Learning Sciences*, 3(3), 227-264.

7. Linn, M. C. & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*, Mahwah, NJ: Lawrence Erlbaum Associates.

8. Scardamalia, M., and Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *The Journal of the Learning Sciences*, 1 (1), 37-68.

9. White, B.Y., and Frederiksen, John R. (1998). Inquiry, modeling, and metacognition: Making Science Accessible to all Students. *Cognition and Instruction*, 16(1), 3-118.

10. Linn, M. C., Bell, P., and Hsi, S. (1998). Using the Internet to enhance student understanding of science: The knowledge integration environment. *Interactive Learning Environments*, 6(1), 4-38.

11. Resnick, M., and Ocko, S. (1991). LEGO/Logo: Learning through and about design. In I. Harel and S. Papert, (Eds.), *Constructionism* (pp. 141-150), Norwood, NJ: Ablex Publishing Corporation.

12. Resnick, L. & Resnick, D. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. Gifford & M. O’Connor (Eds.), *Cognitive Approaches to Assessment*. Boston: Kluwer-Nijhoff.

13. Cognition and Technology Group at Vanderbilt (1993). Designing learning environments that support thinking: The Jasper series as a case study. In T. M. Duffy, J. Lowyck, and D. H. Jonassen, (Eds.), *Designing Environments for Constructive Learning* (pp. 9-36), Berlin: Springer-Verlag.

14. Bennett, G.K. (1940). *Test of mechanical comprehension*, New York: The Psychological Corporation.

15. Ferguson, E. L. & Hegarty, M. (1995). Learning with real machines or diagrams: Application of knowledge to real-world problems, *Cognition and Instruction*, 13(1), 129-160.

Biographical Information

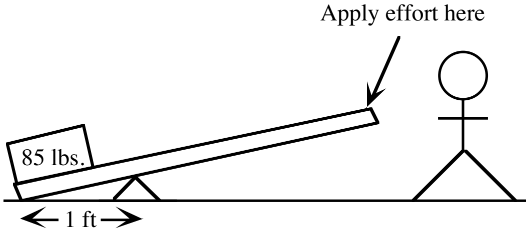
ANN FRANCES MCKENNA
Ann McKenna is currently a Research Assistant Professor in the School of Education and Social Policy at Northwestern University. She received her B.S. and M.S. degrees in Mechanical Engineering from Drexel University in Philadelphia, Pennsylvania and a Ph.D. in Science and Mathematics Education from the University of

California at Berkeley. Dr. McKenna has extensive experience in engineering education research, spending several years as the Berkeley assessment coordinator for the Synthesis coalition. She currently serves as the learning science and assessment consultant within the VaNTH (www.vanth.org) Engineering Research Center.

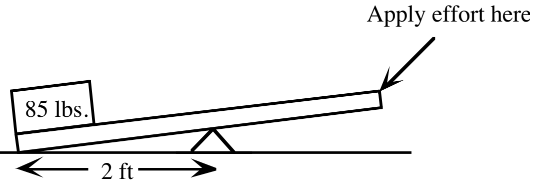
ALICE M. AGOGINO
Alice M. Agogino is the Roscoe and Elizabeth Hughes Chair of Mechanical Engineering. She directs the Berkeley Expert Systems Technology (BEST) Laboratory, the Berkeley Instructional Technology Studio (BITS) and the BITS Multimedia Classroom. She served as Director for Synthesis, an NSF-sponsored coalition of eight universities with the goal of reforming undergraduate engineering education, and continues as PI for the NEEDS (http://www.needs.org) digital library of courseware in science, mathematics, engineering and technology (http://www.smete.org). She has supervised 52 MS projects/theses, 23 doctoral dissertations and numerous undergraduate researchers. Dr. Agogino received a Ph.D. from the Department of Engineering-Economic Systems at Stanford University (1984).

Appendix A: Lever Overall Pre/Post Test

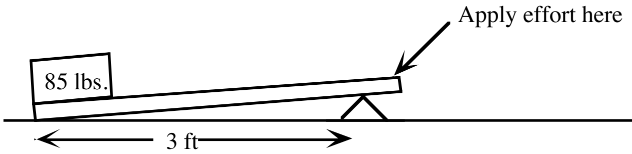
1.



(A)



(B)



(C)

Doesn't Matter.

(D)

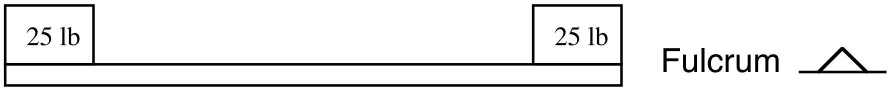
Which lever given above A, B, C, or D would be the <u>easiest</u> to lift the 85 lb. box? Why?

Answer:
Explanation:

2. Given the same picture above, which lever A, B, C, or D would be <u>the most difficult</u> to lift the 85 lb. box?

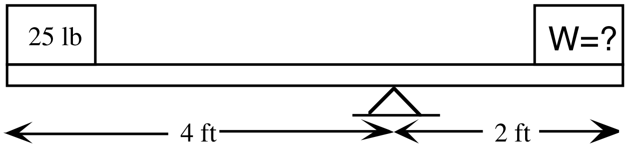
Answer:
Explanation:

3. If there are equal weights on both sides of the lever, where should you put the fulcrum so that the lever will balance?



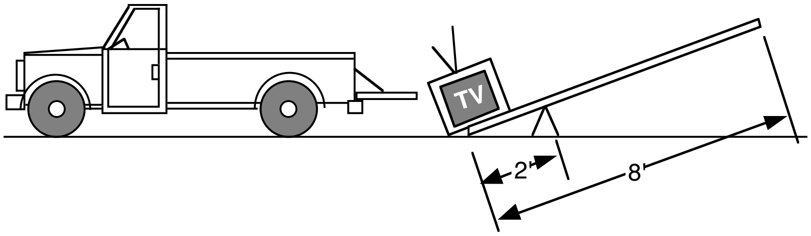
(A) Closer to the right side
(B) Closer to the left side
(C) In the middle
(D) Doesn't Matter
Answer:

4. If there is a 25 lb weight on the left side of the lever, how much weight (W) do you need on the right side to balance the lever given below?



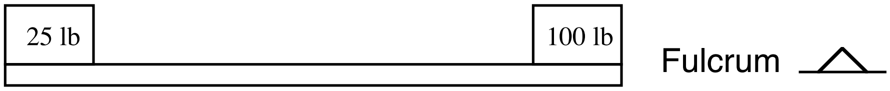
Answer:
Explanation:

5. Your friend is moving into a new apartment and she has asked you to help her move. She needs to load her TV set onto the truck but it is too heavy to lift. You have decided to use a lever to help you lift it onto the truck. If the TV set weighs 50 lb., how much effort (force) is needed to lift it up.



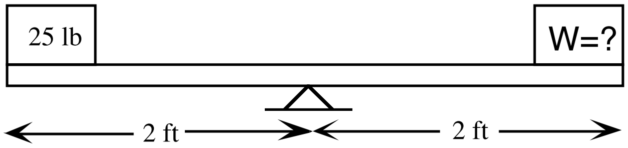
Answer:
Explanation:

6. If there are two different weights on both sides of the lever, where should you put the fulcrum so that the lever will balance?



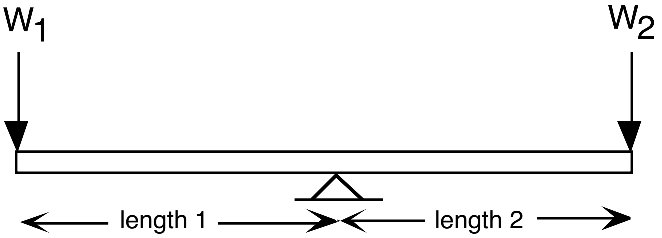
(A) Closer to the 100 lb side
(B) Closer to the 25 lb side
(C) In the middle
(D) Doesn't Matter
Answer:

7. If there is a 25 lb weight on the left side of the lever, how much weight (W) do you need on the right side to balance the lever given below?



Answer:
Explanation:

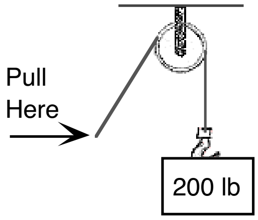
8. You are given the lever below where W_1 is the weight on the left side, W_2 is the weight on the right side, and the length between the fulcrum and the weights are length 1 and length 2. Using W_1, W_2 , length 1, and length 2, write an equation that will balance the lever.



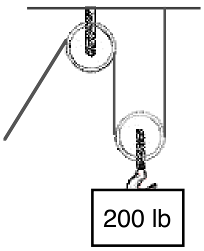
Answer:
Explanation:

Appendix B: Pulley Overall Pre/Post Test

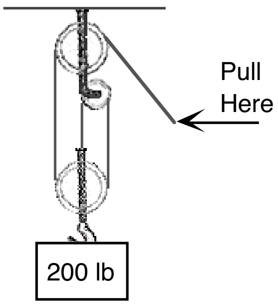
1.



(A)



(B)



(C)

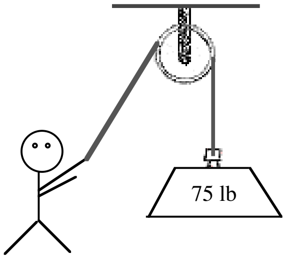
Which of the three pulley arrangements would be <u>the easiest</u> to lift the 200 lb weight, (A), (B), or (C)? Why?

Answer:
Explanation:

2. Which of the three pulley arrangements would be <u>the hardest</u> to lift the 200 lb weight, (A), (B), or (C)? Why?

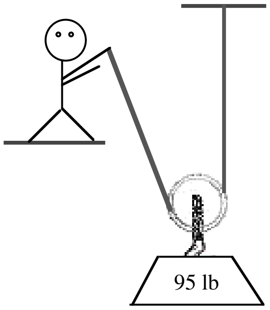
Answer:
Explanation:

3. You are given the pulley system as shown below. With how much force do you need to pull to hold the 75 lb. weight?



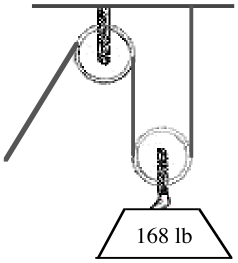
Answer:
Explanation:

4. You are given the pulley system below. How much force do you need to hold the 95 lb. weight?



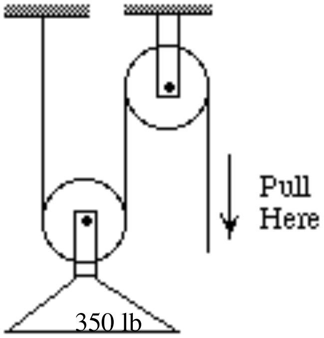
Answer:
Explanation

5. Given the pulley system shown below, how much force is needed to hold the 168 lb. weight?

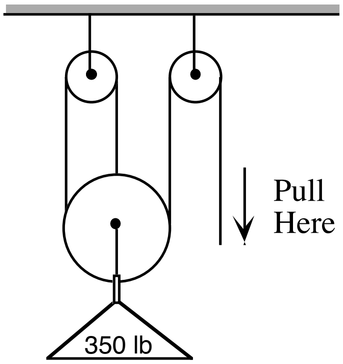


Answer:
Explanation:

6. For which one would you have to pull harder A or B? Why?



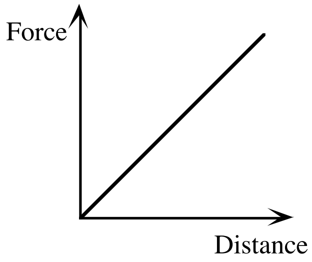
(A)



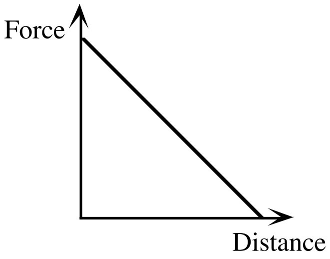
(B)

Answer:
Explanation:

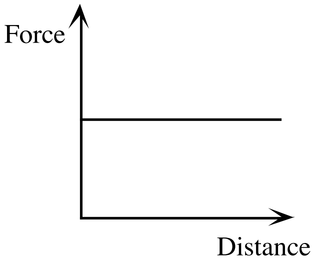
7. Below are three plots a, b, and c. Which one of these plots shows that an increase in distance corresponds to an increase in force? (Circle your response)



(a)



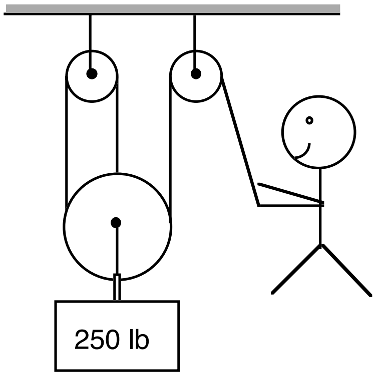
(b)



(c)

Answer:

8. Given the pulley system shown below, how much force is needed to hold the 250 lb. weight?



Answer:
Explanation: