

# STATICS USABILITY EVALUATION

---

REPORT ON USABILITY OF  
OLI STATICS COURSE

BILL JEROME & AARON BAUER

OPEN LEARNING INITIATIVE — CARNEGIE MELLON UNIVERSITY

# CONTENTS

<b>Contents</b> .....	<b>2</b>
<b>Executive Summary</b> .....	<b>3</b>
<b>Introduction</b> .....	<b>4</b>
Background .....	4
Methodology .....	4
<b>Interface Evaluation</b> .....	<b>5</b>
Results .....	5
<b>Overall Recommendations</b> .....	<b>16</b>
<b>Further Study</b> .....	<b>21</b>
<b>References</b> .....	<b>22</b>

# EXECUTIVE SUMMARY

A team of two researchers performed a think aloud usability study on the current development version of an upcoming OLI course in Statics. The goal of the study was to examine the interactive elements currently in place by observing their use by real students. Five individuals were observed for approximately one hour each. Users were then given a five question quiz on the material, which they were told would account for a part of their compensation in order to motivate them to learn the material in the time allotted.

After reviewing the recordings and notes from those sessions, a number of specific element *types* stand out as having been difficult to use, and many smaller breakdowns were observed on specific items.

The most important of these findings are:

- The new force application widget saw a large number of breakdowns
  - Users could not tell in which contexts they were interactive and which ones they were not. The order of their use (switching back and forth between interactive and not) in the context of the course added to this confusion.
  - Text boxes for force display proved confusing as users wanted to enter values into them
  - The widget itself is particular about cursor placement and use.
- The angle selection widget (protractor) also caused a lot of confusion and breakdowns in use. One user never succeeded in using it at all.
- Large images that were crowded with many small figures were troublesome for some students.
- The interactive vs. not interactive problem applied to other tutors as well, such as ones where the top half is not interactive and the bottom half is.
- “Show me” interactions may be problematic.
  - Many users did not understand their functionality or purpose at first
  - Upon realizing they did not need to answer and could simply press all the buttons in sequence, behavior changed from engaged to highly passive

This report itemizes many lesser problems identified throughout the study. Recommendations on changes to the materials in order to address these concerns are also included.

# INTRODUCTION

---

## BACKGROUND

---

A course in statics is currently being developed for the Open Learning Initiative (OLI) and incorporates a number of new interface designs to OLI. In an effort to determine how well students can interact with these new interfaces, this study evaluates real users sitting at a computer interacting with the material. Focus was placed on interfaces that include non-standard widgets that were developed to support the teaching of the Statics material. Not all material currently present in the course was evaluated.

---

## METHODOLOGY

---

The primary evaluation technique used was the Human-Computer Interaction (HCI) protocol known as the think-aloud. [John 99]. Participants in the study were shown how to think aloud while performing tasks, shown demonstrations, and run through practice examples of thinking aloud in both mental exercises as well as computer based tasks before being given tasks related to the study.

Participants were selected by posting announcements requesting participants from the CMU student body who had previously taken Physics I. Participants were offered \$10/hour for one hour of participation, plus \$5 for performance on a paper quiz following the session.

### STAGE ONE

After preliminary questions verifying that students had taken an introductory physics course, students were pointed to parts of the course materials to interact with. This stage contained seven tasks. Not all students completed the tasks in the hour provided. At the completion of each task, researchers used their discretion to speed the process while still allowing the students to interact with the most essential components of the material. Students were instructed that they did not need to speak aloud while reading text materials, but that they should articulate any thoughts that come to mind aside from the direct text.

### STAGE TWO

Upon completing the work, students were asked to answer five questions on paper based on the material they interacted with in stage one.

# INTERFACE EVALUATION

---

## RESULTS

---

### INTRODUCTION

After five participants, here are our findings about the usability of the Statics interfaces for novice student use.

Each section below represents a category of error students made frequently, or an area or part of the course. In the recommendations section we address each of these sections individually.

Evidence is presented in a tabular format at the end of each section. Here is an example of the columns and their meaning.

REFERENCE	USER ACTION	USER QUOTE
U0 14:04	Selects answer	"I'm choosing this label for a force"

This would mean user zero, at time 14 minutes and 4 seconds into recording selects an answer in a tutor and speaks "I'm choosing this label for a force." We attempt to present as much evidence as is reasonable in each case. In some cases, we make a case using a lack of examples. It is suggested that the evidence sections be read, as they often clarify and expand what is written in the text. Not every case of evidence is reported for each section. In cases where something happened repeatedly, for example, only some definitive subset may be cited.

### OVERARCHING BREAKDOWNS

This section focuses on findings that span more than one element of the course, or more than one type of element used in the content. Some of these have very specific instances in the content, but are indicative of broader areas of concern.

#### "SHOW ME" INTERACTIONS CONFUSE USERS

A number of users had difficulty in using the "Show me" style interactions, particularly the first time they encounter one. In some cases, users continued to make the same errors on subsequent instances of the interactions. Particularly, users want to interact with the display and often try clicking around the interface in order to make something happen. In the first example of such an interaction, users most frequently began by clicking on the "C" block at the right of the "show me" buttons. Also, some users did not know what they were supposed to be doing with the interface. Multiple users asked aloud if they were somehow supposed to be answering the stated questions.

Additionally, once users determined that the only options available to them were to click the series of buttons, we observed students performing rapid clicks to reach the end rather than reflecting on each step of the interaction.

U1	Tries interacting with block to right of “show me,” fails, but tries again at next step regardless.	”I’m going to try this again because I’m a sheep.”
----	---	--

#### CANNOT DIFFERENTIATE INTERACTIVE FROM DEMONSTRATIVE DISPLAYS

As described above regarding “show me” interfaces, and as detailed below in the specific case of the force application widgets, users often had problems determining what displays were interactive. Users continually tried to interact with “show me” displays, and often were left simply guessing at whether or not they could interact with the force application widgets or how to interact with the protractor widget. This issue was observed for many elements of the content .

U2 40:30		“alright, not interactive this time”
U4 50:03	Tries interacting with non interactive tutor	
U5 52:00		“oh, is it stuck?”
U3 36:05	Does not yet realize they cannot move the spring	“Oh, I can move the spring!”

#### DEFINITIONS NOT LOCATABLE BY USERS

The content does not provide a good reference ability for definitions of the many terms the content provides. Students mentioned on multiple occasions that they were seeking a definition of a term or simple explanation of a term that they could not locate in the content. Most often the exhibited behavior was evident for the terms “sense” and an explanation for labeling.

U3 11:58		“looks like sense is direction but they’re not explaining it”
----------	--	---

#### SUMMING PROBLEM TURNS INTO A GAME

It was often observed that users did not quite understand the N notation for unknown forces. Users were able to often guess the correct answer, not by deriving it, but by being trained that “real weights” were never the right answer unless they were marking force due to earth. Many users articulated, for example, the sum of the weights of blocks on top of them as the force on the upper boundary of the object and became confused when that choice was not available. Rather than learning why such choices were not available, and not having to do so since a very limited set of options was available at any one time, students heuristically began picking the right answer only by identifying the ones provided that they knew to be wrong.

U2 34:56		“well, going from this before, I would before, there’s a rule that I sort of found, which is you’re not going to be labeling actual forces here.”
----------	--	---

U2 26:10		“since there’s no [combo], I’m gonna just click Nb”
U2 33:20		“there’s no 70 pounds, so I’m just gonna click tension”

#### USERS DO NOT READ TUTOR FEEDBACK

Users were not observed to pay attention to the feedback provided by tutors at the bottom beyond noting their correctness. In particular, this allowed students to continue misconceptions they stated when they select the right answer for the wrong reason (by not reading the feedback on a correct selection), and negative feedback for mistakes was also usually ignored in favor of guessing again on all but the most complicated tutors.

U4 10:35	User has misconception of correct answer, does not read feedback	
----------	--	--

#### USERS DO NOT USE HINT BUTTON

Aside from not reading feedback on the bottom of interactions, very few students utilized the hint functionality even when stuck for how to solve a problem. Rather than requesting a hint, users would select bad inputs and read the comments presented. In only a couple of occasions through all five users did a student request a hint on a problem

U1 10:24	Unsure of correct answer, does not ask for hint	“let’s try some random guessing”
----------	---	----------------------------------

#### IMAGES GIVEN PRIORITY OVER TEXT

Overall, it was evident that the users spent more attention on imagery, regardless of interactivity, than on the text presented. This is noted in this report because it is evidence that the quality of both the interactive and the static images in the course should be as well designed and rendered as possible, given that students attempt to do more learning from these elements than from the text. The text is typically seen to be used as a backup for when the image/media fails to answer a question in the student’s mind.

U2 23:28	Spring animations [before reading everything]	“okay, let’s just see what happens”
U2 23:20		“It would be kind of cool if this had like said subsystem in the picture, cuz I usually look at pictures”
U3 33:34		“seems long, so I’m just going to click on the animation” [nothing happens] “let’s read”
U2		“Moving things make me happy”

ANIMATIONS RUN SLOWLY

This particular breakdown was in part due to the recording software used by the researchers to capture the screen as users worked. However, it should be noted that slower computers are likely to display the animations in a slow way as well. In particular, the fading in and out of steps in some tutors took a distracting amount of time, and waiting for the springs to pull back slowly was frustrating to some users. The “configuring” phase of the tutors took upwards of three to five seconds in some cases adding notable delay.

U5 35:10		“Kind of boring watching those springs load up”
----------	--	---

NEXT VS. START OVER BUTTONS GET CONFUSED

Although we only encountered one experience where a user looped around without even noticing it for a few moments, this style of interaction is present on many interactions, so it is mentioned here. The “Start over” button replaces the “Next” button on many interfaces. Although it is colored differently with different words, it did not immediately convey a different behavior to all users, and one user selected it and completed a previous problem again without noting it.

U3 17:50	User clicks start over thinking it reads next. Doesn't realize duplication until second question.	
----------	---	--

ITEM BY ITEM OBSERVATIONS

DEFINITION OF INTERACTION: WRENCH IMAGE

There are two distinct problems with the wrench image. The first is of course the typo still present in the material that refers to “hand H” which is actually labeled as “G” on the diagram, and the legibility of the diagram and general. Aside from that, however, users had difficulty in bouncing between the paragraph of text and all of its references to the diagram and the diagram itself.

U1	Reading text, has difficulty keeping her place in the text	
U5 4:35-5:00	Pops around a lot between text and image	

CONTACTING BODIES: ANIMATIONS SERVE LITTLE PURPOSE

Users tended to want to interact with the two animations presented on this page. Some users read the material but not the directions and wanted to interact with the bodies presented. It is possible they cause distraction, as users who read the text bounced back and forth between the text and the

animations. No real evidence of reflection on the animations was apparent. Users waited for the animation to complete, and when it was, simply moved on.

U1	Tries to click on drawn bodies	“Anybody in kindergarten knows this by now”
U3 7:53-7:59	Reads content, but not directions, and tries to interact	
U5 12:10	User moves back and forth between reading text and images	

#### CONTACTING BODIES: DIRECTION SELECTOR

This novel widget worked fine for all users.

#### FORCE VECTORS: FIRST SHOW ME DISPLAY

This first example of the “Show Me” interfaces demonstrates most of the failures users had with this interaction type. Right from the beginning most users are not sure what to do, and the interface does not make it clear to them.

<i>Am I supposed to be answering these questions, or just click 'show me'?</i>	U2 10:53
<i>.. you answer in your head, or is there somewhere you can click ?</i>	U2 11:02
<i>So, do I have to answer this, or just see...</i>	U4 12:34
<i>...do I answer it here?</i>	U5 23:00

It appears that the button labeled “show me” is interpreted as a type of hint, or resource to be used when the student cannot answer the posed question.

<i>As I'm clicking through them I guess I'm clicking the show me button because I'm confused, if I'm allowed to click on things here</i>	U2 11:42
--	----------

In an effort to answer the question posed, users attempt to interact with the shown elements. This first appearance of the “show me” interaction has a block redrawn next to the question. This seems to draw the most attention from users seeking something to interact with. One user also had a difficulty when she did not see the block update with the “o” symbol after clicking the first “show me” button. She noticed it later and didn’t understand it. One user did notice the circle appear, and rather than understand what it was, made later attempts to draw the points herself, thinking that she would have been able to draw it and the “show me” was acting like a hint and doing the step she had failed to complete for her.

Some students exhibited a behavior of rapidly clicking through the “show me” buttons once they realized there was nothing for them to interact with. These users may have engaged enough to anticipate the coming steps and then press “show me” but users began to ignore the written text feedback being added to the display, focusing on the images instead. One user articulated their response to this interface like so:

*I supposed I'd think about it, but if I was a student in this class and I was in a rush, I would just click show me through all these things... cuz usually when classes are in session, you don't have much time to ponder things.*

U2 15:30

As one final note regarding this particular instance, the orange coloring of the strings seemed out of place to some users, and one was actually confused by it.

U1	Clicks “show me” and later realizes the point drawn on the block	“not sure what ‘O’ means”
U2 10:53	Examining display	“Am I supposed to be answering these questions, or just click ‘show me?’”
U2 11:02	After clicking “show me”	“you answer in your head, or is there somewhere you can click”
U2 11:42	Examining display	“as I’m clicking through them I guess I’m clicking the show me button because I’m confused, if I’m allowed to click on things here”
U2 12:40	On first one, she clicked show me, because she hadn’t seen before. It drew a little circle in the middle, so she tried to do that on later questions, and is unable to	
U2 15:30		“I supposed I’d think about it, but if I was a student in this class and I was in a rush, I would just click show me through all these things... cuz usually when classes are in session, you don’t have much time to ponder things.”

U3 10:58	Quickly clicks through steps	
U4 14:34		“So, do I have to answer this, or just see...”
U5 23:00	After attempting interaction	“do I answer it here?”
U5 26:12	Examining display	“what is this orange thing?”

#### PURPOSE OF DRAWING FBDS: IMAGE

The purpose of the image at the top of this page seems to be lost on users who stare at it briefly and move on. The image toward the middle of the page is far too cramped, which lead users to attempt to enlarge the image, including one who thought about saving it and opening it in paint in order to enlarge it, but did not.

U2 21:33	Examining image, clicking it (causing popup)	“this stuff here is kind of small to read. I wonder if I can click on it to make it bigger? Not so much.”
U3 21:22	Examining image, clicking it (causing popup)	“I’m trying to make this picture bigger”
U4 33:04	Examining image	“and it could be bigger”
U5 34:04	Examining image, clicking it (causing popup)	

#### PROCEDURE OF DRAWING FBDS: DID I GET THIS?

Some issues surrounding what is and is not clickable were apparent on this display. Some users attempt to interact with the top as well as the bottom portions of the tutor. Additionally, some users do not understand which elements of the copy they are supposed to be interacting with are selectable. One user never realized the circles were selectable and was stuck on the problem, though they did not ask for a hint. This leads to guessing.

U1 10:24	Unsure of correct answer, does not ask for hint	“let’s try some random guessing”
U5 43:50	Doesn’t understand circles are selectable until prompted by researcher	
U4 10:22		“let’s try randomly guessing then. Nope, didn’t think so”

#### PROCEDURE OF DRAWING FBDS: PULLEY DIAGRAM

Graphical appeal aside, this image is very large, contains many elements which then become too small, and is difficult for users to grasp. Readers moved frequently between this image and the above

material. Due to the size of the graphic, this causes readers to scroll back and forth in an effort to understand what is being taught.

U2 31:31		“The way I like to read things, I’d prefer that it [the text and image] was sort of like, next to each other, so you can see when they’re talking about subsystem 1 and subsystem 2, cuz... scroll up and down...find place again”
U3 27:15	User reads material, has not seen image, becomes confused	“which one is this talking about”
U1 13:15	Scrolls back and forth between text and image	“where are the subsystems?”

#### PROCEDURE OF DRAWING FBDS: QUIZ: INTERACTION TOO SMALL

The closeness of the interaction areas on parts of this quiz display cause students to accidentally select the wrong response. Every user that completed all of the steps made at least one accidental selection on this interface.

U1 18:34	Accidentally selecting the wrong line	“oh no, that’s not what I meant to click on”
U2 33:58	Accidentally selecting the wrong input	“Oops, I accidentally clicked that one, because they’re all rather close together”
U3 32:00	Accidentally selecting the wrong line	“I’m just going to try the circle”

#### TYPES OF MOTION: ANIMATION VS. INTERACTION

Users had difficulty in determining which of the displays required user interaction and which ones simply animated at the press of the “start motion” button. The methodology used by the students was to first attempt to interact, and if that failed, fall back to clicking the start button to see if they really didn’t need to do anything.

#### TYPES OF MOTION: GENERAL MOTION DEFINITION

There seemed to be some confusion over the definition of general motion. Aside from the somewhat dense phrasing of that definition when compared to the others, the listed definition is “general planar motion” which is not the way the term is used anywhere else in the content. These core definitions are also rather oddly placed in the content making the terms harder to locate for reference after being read (or passed over).

#### DEPENDENCE OF MOTION ON FORCE ATTRIBUTES: SPRINGS (I)

Users had particular trouble with the spring/force application widget. Some of the issues that aided confusion are separated out below, specifically, the use of text boxes and affordances of the widget itself. Aside from those aspects of the tutors, the widgets themselves proved difficult to use and understand. All users, even if successful on the first of a sequence of three tutors, had breakdowns on the third due in part to the particular sequencing of elements. In the first of three, the springs can not be set by the user by direct interaction. In the second, they can be manipulated, and in the third,

they again cannot be interacted with. This back-and-forth of available interaction left users stumped each time as to what they were supposed to do, distracting from what they are supposed to observe as learning material. Instead, they became focused on just getting them to function. One user began clicking like mad all over the third interface believing they needed to set the values on their own.

U2 40:15	After correctly using first two, clicks all over interface on third	“Having issues with this now”
U2 40:30		“alright, not interactive this time”
U4 50:03	Tries interacting with non interactive tutor	
U5 52:00		“oh, is it stuck?”
U3 36:05	Does not yet realize they cannot move the spring	“Oh, I can move the spring!”

#### DEPENDENCE OF MOTION ON FORCE ATTRIBUTES: SPRINGS (II): TEXT BOXES

Using text boxes to display the amount of force being applied confuses users into thinking that they can directly affect the values inside of those boxes. In no case does selecting the text (which is possible in the text boxes but not elsewhere) allow the student to enter a value. It is even more confusing when in some displays which are primarily animations which never let the user change the force values, even indirectly. This compounds the problems users had in determining how and when to interact with the force displays.

U2 38:30	Tries typing into text box	“When I first saw it, I thought I was supposed to type in the force that I wanted and see what happened”.
U2 39:10		“over here is says 0 and 0 and it had been talking about 1 and 2, so I thought I was supposed to like, enter in...”
U3 37:10	Tries to click apply forces on second animation. Nothing happens. Clicks on it (doesn't hold mouse, so changes with no force applied) Then tries text box again	“Oh, I have to click on this?” ... . “No?”
U3 37:40	Tries textbox again	

#### DEPENDENCE OF MOTION ON FORCE ATTRIBUTES: SPRINGS (III): AFFORDANCES

Some of the visible affordances on the force application widget are misleading about the available interactions. While the “+” symbol inside of a circle tends to afford interaction, it also implies opening, which isn't quite right, rather than grabbing. When the interaction is not available to users, it appears as a grippable rectangle, leading users to believe they can interact with it.

U3 37:10	Tries to click apply forces on second animation. Nothing happens. Clicks on it (doesn't hold mouse, so changes with no force applied) Then tries text box again	"Oh, I have to click on this?" ... . "No?"
U3 37:40+		"I can't drag this"

#### DEFINITION OF MOMENT IN 2D (PAGE 2 OF 2): PROTRACTOR

The protractor-looking widget for selecting an angle caused many problems for students. In some cases, users thought they had completed the interaction without ever changing the angle, and one user in fact imagined to herself that she could actually see the one that was supposed to move faster doing so, even though the angles and speeds on display were identical. Only one user actually read all of the use instructions first, and was fairly successful in using the tutor although he still had difficulty in making it work at first.

One concern raised for the researchers by this imagining of results is that this interface in particular allows 'mistaken' results. For example, one wheel may be left at zero degrees at "half force" and the other one turned at an angle and set to "full force" resulting in a wheel that spins faster. Although it is probably valuable for these interactions to be possible, in the narrow context of the single question, students may accidentally observe the wrong behavior without understanding why.

U3 43:55	Clicks force bat instead of angle	"dragging round the circle... what?"
U4 51:00	Never uses angle selection	
U5 53:49	Tries, fails, starts on same angle. Reads, resets, and correctly uses angle selection	
U1 28:16	First seeing interaction display	"That's kind of a confusing picture"
U1 28:40	Tries and fails to grab angle selector	
U4 52:30	User imagines right wheel is "correctly" moving faster though they are identical, and recounts what is supposed to be showing	"So basically, it shows..."

#### DEFINITION OF MOMENT IN 2D (PAGE 2 OF 2): WAITING FOR ANSWER

Some users, upon selecting their answer on some of the tutor interfaces chose not to wait for the animation that is supposed to allow them to see the correct answer for themselves. It is not certain what the cause of this behavior is. Perhaps lack of immediate feedback which they are now used to seeing leads them to believe they were correct, or perhaps the simple speed problem made them skip the elements.

### PAPER TESTING RESULTS

Below we chart successfulness of students on the paper problems they worked. All students had some level of exposure to the topics quizzed prior to answering the questions.

	Paper 1	Paper 2	Paper 3	Paper 4	Paper 5
U1	✓	✓	✓	✓ <sub>2</sub>	✓
U2	✓	✗	✓	✗	✓ <sub>2</sub>
U3	✓	✓	✓	✓	✗
U4	✓ <sub>1</sub>	✓	✗	✗	✗
U5	✓ <sub>1</sub>	✗	✗	✗	✓

1 : Did not answer all parts of question

2: Minor error

Many small errors were due to mislabeling of forces or omission of forces on the last two questions. The most mistakes seemed to be made on the questions requiring the drawing and labeling of a subsystem (question four) and labeling the external forces on a body (question five). Question four also asked users to articulate the steps for drawing a FBD. It is worth noting that no student seemed to articulate the sequence of steps in the ways that the content presents them. Whether or not this is of concern is up to the instructional designers. Most students condensed the sequence of steps into shorter sequences. One student separated finding forces due to earth and then all other bodies.

Some students failed to correctly determine which wheels spin fastest on question three.

Most students performed well on questions one and two, at least for all parts they chose to answer.

# OVERALL RECOMMENDATIONS

Here we describe *possible* solutions to each breakdown cited in the results section above. Please note that these suggestions are just that. It is possible that some of these recommendations may be difficult or time consuming to implement. It is up to the developers to assess the resources involved in implementing each of these recommendations and value that against the severity of the issue being resolved, as well as to evaluate other possible remedies.

## OVERARCHING BREAKDOWNS

### “SHOW ME” INTERACTIONS CONFUSE USERS

The tutors that use the “show me” interactions appear to be designed in a way that lead students to believe that they should be answering questions. They therefore look at the “show me” button as a hint that should be used if they cannot figure out the answer to the question. When they cannot find a way to answer the question, then they click show me.

There are two potential solutions to this problem. The phrase “Show Me” may be what is leading users to conceptualize the button as a hint. Replacing “Show Me” with another term, such as “Play” (or utilizing a play icon), may be appropriate. Using the question text to explain the interaction may help, but it may not be read by students.

However, students may abuse any such button, clicking through too quickly to fully process the material. One student mentioned this possibility outright, and others appeared to click-through more quickly than they could read. Giving students the option of actually answering questions [for example, clicking on the “point of application” may be appropriate, depending on learning objectives. Another solution would be to add a timeout between pressing one “show me” button and the activation of the next.” The danger with this solution is that any lengthy delay may either not be understood by students, or it may discourage them from using the rest of the interaction completely.

### CANNOT DIFFERENTIATE INTERACTIVE FROM DEMONSTRATIVE DISPLAYS

Though the interaction technique [press play, drag] was often described in text, many students attempted to complete the animation prior to reading the text, leading to confusion. Some students who read the text before doing the animation skipped over instructions. This makes it clear that the interaction must be clear in the interface. In many cases, it was not.

Interactive elements must be clearly distinguished from non-interactive elements. At a minimum, the interactive elements themselves (not just the cursor) should change when they are moused-over. Color coding is one possible solution. Fading in and out is another potential solution, though this should be a very subtle effect. Elements that appear interactive but in fact are not [the biggest example being text boxes, in which students want to enter information] should be avoided.

Students appear to want to interact, rather than simply clicking buttons. For example, allowing the student to slide the block in the force application widget may be appropriate.

Should it be necessary to prevent student interaction in some interfaces in favor of simple display, those interfaces that do not allow interaction could possibly be globally identified as demonstration only. For example, a header across the top that clearly defines the interface as non-interactive may be appropriate.

#### DEFINITIONS NOT LOCATABLE BY USERS

New words should not be used as examples prior to their explanation. The most prominent example of this is “sense,” which was never fully explained in text. Another example is “general motion,” which did not cause as many problems.

#### SUMMING PROBLEM TURNS INTO A GAME

When deciding on forces between two bodies in contact, students often attempted to add the weights of elements on top. Though we are unsure whether this is a misconception, it is clear that this is common and not addressed in the learning materials. Due to this fact, students appear to learn the “unknown force” rule on a very surface level. Perhaps this should be addressed either in the learning content, or by allowing the user to answer in this form, and then giving feedback. Increasing the number of available choices somehow, early in the presentation of the material, may also help address this apparent misconception.

#### USERS DO NOT READ TUTOR FEEDBACK

Feedback often starts with a quite generic statement, such as “That is not quite right” or “Very Good.” Students appeared to stop reading there, instead clicking “next” or giving another answer very quickly. On several occasions we saw students answer correctly through incorrect reasoning, which would have been addressed in the feedback they did not read.

We see two potential solutions. The generic statement could be eliminated, or placed at the end of the feedback. However, students do use the color coding [so may not read feedback at all]. Therefore it may be appropriate to allow a short timeout [2-3 seconds] before the next button is activated. However, fading in of the feedback should be quicker than it currently is in order to make it more likely to be read.

#### USERS DO NOT USE HINT BUTTON

The hint button is often located far away from where the students are actually working, as it is located above the question introduction, and students, having read the problem statement, are now working below that portion of the display. It may be appropriate to reposition the hint button to the area in which students answer questions.

It may be possible to determine automatically if a student is engaged in random guessing behaviors in order to solve a problem (or view feedback associated with wrong guesses) and guide users to the hint functionality.

It is important to note that the hint facility was hardly used at all, and it should be determined if the tutors that place the button far away from student interaction are the only ones that need to be addressed.

#### IMAGES GIVEN PRIORITY OVER TEXT

When designing material, it is important to realize that many students appear to view images, and especially animations, prior to reading text. This means that you cannot depend on text to explain the interaction technique; it must be evident in the interface. This is discussed in other locations in the document.

This also brings to light another issue with the material. Very few images, figures, or animations have titles above or labels underneath. This is symptomatic of a larger lack of structural organization within content-pages. Labeling images, and including small descriptions (i.e. “Figure 1: Interaction and Direct Contact. Note that ...interact, while.... Do not) will make them easier for students to interpret, and perhaps use. Adding additional structural elements, such as example or section titles, may also help accommodate student scanning behavior.

#### ANIMATIONS RUN SLOWLY

Though fading can be a very good technique, slow animations may cause students’ attention to wander. This was especially evident in our study, where screen-capture software slowed rendering even further than normal on a high speed desktop computer. However, even at full speed, most animations appeared slow, from the appearance of feedback to the loading of springs to the rotation of the wheel. These animations should be sped up quite a bit. Delays such as the load/configure process and fading between questions and steps became especially aggravating during testing.

#### NEXT VS. START OVER BUTTONS GET CONFUSED

We observed one student going into a loop, in which he clicked the start over button thinking it was a next button. He did this twice. This most likely occurred because expectations had been set as to what the button at that location did, so he did not bother reading it. A simple solution would be to relocate the “start over” button, perhaps to the center of the screen.

#### ITEM BY ITEM OBSERVATIONS

##### DEFINITION OF INTERACTION: WRENCH IMAGE

If the wrench image is retained, the first task is to correct the “G in image, H in text” issue. However, the image appeared to have problems beyond simple mislabeling. It is a complicated image that appeared to require students to scroll back and forth between the image and the text. It may be useful to replace the image with a simpler image, that would require less review of text. If the image is retained, it may be appropriate to have mouseover areas explaining the relevant material, so that students can read the relevant text without having to leave the image. It may also be possible to split the key ideas (what does interact, what does not) into multiple versions of the image and split the text between them to minimize the need for students to move between text and image.

##### CONTACTING BODIES: ANIMATIONS SERVE LITTLE PURPOSE

Several students appeared bored with the “Contacting Bodies” animations, and found the content trivial. This may have been compounded by the animations’ lack of interaction, which students appeared to expect. If the instructional designers feel the animations should be retained, allowing interaction [i.e. student drags blocks instead of pressing “Go”] may increase students’ appreciation of said animations.

##### CONTACTING BODIES: DIRECTION SELECTOR

The direction selector in the “Did I Get This” in “Contacting Bodies” was used as intended by students. This interaction should be retained.

##### FORCE VECTORS: FIRST SHOW ME DISPLAY

This is the first experience students get with the “Show Me” interaction. The design of this interface appears to suggest to students that they answer the questions. Though a wording change may be appropriate [i.e. “Show Me” to “Play” or play button iconography], it may be better to tell the

students in the question to try to figure it out in their heads and then get the answer. A better solution may be to allow students to actually answer the question by interacting with the elements.

As mentioned above, abuse of any type of “Show Me” button may be reduced by extending the timeout between pressing one button and activating the next. Also, it may help to change the color of the additional text away from green, which by now students have learned to mean “correct” and therefore, not necessary to read regardless of length.

Finally, the orange strings appear to be confusing. At a minimum, borders should be added as in the other strings. However, the color itself should probably be changed to something less bright.

#### PURPOSE OF DRAWING FBDS: IMAGE

The drive system did not appear to grab much attention, and its purpose appears to be unclear, especially as it has no label. It should be explained further both in text and through labeling, or eliminated. Lack of labeling is a problem throughout this course, both through labels and titles.

The image of building free-body diagrams is far too small, the text is not readable, as it is both too small and sometimes overlaps other elements. If increasing the size of the image as is creates an image with too much horizontal length, the image should be split into its three elements and presented vertically with labels. This may be preferable to simply increasing the size of the image, as the labeling could be highly instructive.

#### PROCEDURE OF DRAWING FBDS: DID I GET THIS?

It is not clear what students are supposed to interact with. As mentioned above, these elements should be changed (for example their borders could change color, they can pop-out slightly) when moused-over. The presence of two images may be distracting as well. Perhaps students can interact with the one in the problem statement. Selectable items should be distinguished from non-selectable items in other ways, mentioned above.

#### PROCEDURE OF DRAWING FBDS: PULLEY DIAGRAM

Adding labels to the pulley diagrams (even as simple as “Figure X”) may solve some of the confusion displayed by students. There is still too much information split between text and tiny images. Splitting up the images and text may allow students to better concentrate on the learning materials.

#### PROCEDURE OF DRAWING FBDS: QUIZ: INTERACTION TOO SMALL

The small size of the images led to many inadvertent clicks. All images should be increased in size. As above, it may also be appropriate to have only one body-diagram that serves as the problem statement and the interactive element.

#### TYPES OF MOTION: ANIMATION VS. INTERACTION

It may not be necessary for this animation to be interactive at all. Would anything be lost by having the bars moving at all times, with the “translation” and “general motion” bars going “around the world?”

Once again, labeling elements would be appropriate.

#### TYPES OF MOTION: GENERAL MOTION DEFINITION

“General motion” should be defined more clearly, perhaps before the animation. Consistency in the use of terms is very important. If “general planar motion” is a subset of “general motion”, this should be explained. Using some method of making definitions stand out in the text will also help students attempting to refer back in the material for them.

#### DEPENDENCE OF MOTION ON FORCE ATTRIBUTES: SPRINGS (I)

Students are not sure with what to interact. This is not surprising, as the affordances offered are confusing and not consistent. The grippable box in the spring in the top interaction generally afford “grabbing.” The “plus” does not. In addition, it is not clear why students are allowed to interact with some springs, but not others.

There should be a consistent interaction. It may be enough to simply offer the grippable box in all animations. However, we suggest including a slider next to each spring, which allows students to set that spring’s force, rather than having to learn where to grab the spring. It is generally preferable to use standard widgets whenever possible.

#### DEPENDENCE OF MOTION ON FORCE ATTRIBUTES: SPRINGS (II): TEXT BOXES

Text boxes are generally interactive, though in this case they are not. Boxes should not be placed around the Forces. The text should not be selectable.

#### DEPENDENCE OF MOTION ON FORCE ATTRIBUTES: SPRINGS (III): AFFORDANCES

See above: “Springs(I)”

#### DEFINITION OF MOMENT IN 2D (PAGE 2 OF 2): PROTRACTOR

This interaction is understandably difficult, as there are many things to manipulate. At this point, students appear fairly comfortable manipulating springs, so whatever interaction is chosen for that purpose can be maintained. However, the angle widget caused many problems, as the affordances are not clear. At the very least, the bottom segment of the spring should be “grippable” [the three-line box]. However, this would still cause problems, as two interactive elements would be located side-by-side. It may be preferable to have students click on locations in the protractor circle in order to set angle. These hotspots should be made clear through coloring.

It is also reasonable to move the selection outside of the current working area by placing a slider under each one with the number value always visible. Changing the value of this slider can also visibly move the protractor, without the user having to manipulate it directly. It is usually preferred to use standard existing interface widgets wherever possible.

Finally, it may be appropriate to set the default (start) position of one of the springs to a different angle. This would reduce the problem described above in which students never change the angle, and thus and misinterpret the results.

#### DEFINITION OF MOMENT IN 2D (PAGE 2 OF 2): WAITING FOR ANSWER

Some students appeared to not wait for the animations to be activated. Speeding the animations up may reduce this problem. Another solution would be to add feedback [“Wheel X is rotating faster than Y because Z”].

## FURTHER STUDY

It is suggested that some portion of the findings in this document be investigated further by another study or set of studies in the future. Use of the interactive elements is currently logged. Using log data of actual student use will be key to understanding how students are using the tutors and demonstrations. There are specific questions that could be answered with this data. For example, determining if students spend time between steps on “show me” displays, or if they rapidly press the sequence of buttons (if they use them at all) may help shape the experience better.

It may also be worth investigating the same questions of other interactive elements. The interfaces that have students select an answer before having access to the animation that shows the answer may also reveal just how students use those interactions after they answer the question (and whether or not there is any correlation to the correctness of the answer they choose).

The lack of student use of feedback and hints should also be examined. Placing key pieces of information in the ‘correct’ feedback blocks on traditional tutors does not seem to be effective. Whether or not the positioning of the hint button on tutors where all other interaction takes place on the bottom half of the display could be examined. (i.e. do those tutors see a lower use rate of hints than others).

One idea worth investigating is how, through data logging, it may be possible to track student difficulties with the interfaces. Being able to tell that students have failure/success ratios on certain interfaces through some automatic means could be useful, even if it does not automatically determine source of user errors.

# REFERENCES

Gilmore, D. J. (1996) **The Relevance of HCI Guidelines for Educational Interfaces.** *Machine-Mediated Learning*, 5(2), 119-133.

John, B. E. (1999) Carnegie Technology Education Course #SSD4 User-Centered Design and Testing, Unit 3 Think-aloud Usability Testing.

Martinez Miron, Erika A., du Boulay, B., Luckin, R. (2004) **Goal achievement orientation in the design of an ILE.** *Proceedings of Workshop on Social and Emotional Intelligence in Learning Environments: Seventh International Conference, ITS2004, Maceio, Alagoas, Brazil, September*, 72-78.

Roll, I., Baker, R. S., Aleven, V., Koedinger, K. (2004) **What goals do students have when choosing the actions they perform?.** International Conference on Cognitive Modeling.