Latent Response Times and Cognitive Processing on the FMCE

Scott Bonham

Department of Physics and Astronomy
Western Kentucky University
1906 College Heights Blvd. #11077, Bowling Green, KY 42101-1077

Abstract. Latent response time analysis of students on an electronic version of the Force and Motion Conceptual Evaluation (FMCE) provides information on student reading patterns and the role of mental models in student reasoning. Regression analysis looked at the dependence of response times on characteristics of questions, such as amount of text and inclusion of graphs. Results indicate that students generally read through the question text and instructions when first presented, but do not systematically read through answer choices and graphs. Comparison of average response times between pre- and post-instructional assessment found a drop in response times when students used Newtonian ideas but no change for responses using the main alternative concept. The average response time for students who answered using a mix of Newtonian and alternative concepts was not different from those using primarily one or the other; questions rarely activated both concepts at the same time.

Keywords: Response times, FMCE, student model states
PACS: 01.40.Fk, 01.40.gf, 01.40.Ha

INTRODUCTION

Standardized conceptual assessments such as the Force and Motion Conceptual Evaluation (FMCE)[1] are important tools in Physics Education Research (PER). Computer administration of these provides an opportunity to gather additional information beyond that available in traditional paper administration. An example is latent response times, the time between when a question or task is presented and the student responds. This metric provides information on cognitive processing, such as the Stoop effect[2], in which people take longer to read the written names of colors that are printed in a different color (e.g. red letters forming the word "green") than when there is no conflict between the word and ink color. Similar effects have been seen in physics when students read information that conflicts with their mental models[3].

There are many factors that affect how long it takes to respond to a question. Some are random factors than can be described by a probability distribution such as a shifted log-normal curve[4]. Other factors are systematic, for example the time it takes students to read a question or resolve a cognitive conflict.

The research questions investigated in this work are as follows:

1. How do the reading times depend on question components: the number of graphs and images, and the amount of text in individual questions, that common to a group of questions, and that in the response options?
2. How do the mental models used by students to answer questions affect reading times?

Mental models

An important finding of PER is the use of alternative mental models to reason about physical situations. For example, the idea that the force on an object is proportional to the velocity forms an alternative model (*force ~ velocity model*) to the accepted understanding that force is proportional to the acceleration of the object (*Newton 2 model*). Similar is the belief that when two objects interact, the object that dominates the interaction (bigger/faster/imposes its will on the other) exerts a greater force on the other (*winner model*), as opposed to the accepted *Newton 3* model that the forces are always equal.

Theoretical understandings of this behavior fall into one of two general approaches.[5] The misconceptions approach dominated early thinking in PER. In this view, student understanding is perceived

to be a relatively coherent and stable structure, whether utilizing Newtonian or alternative models. More recently, various versions of the knowledge-in-pieces approach have been proposed, including p-prims[6] and facets[7], where the knowledge structures utilized in utilizing physics concepts are smaller and may be transitory. Latent response times can provide some insight on cognitive processes as students reason about these situations. This can include the time it takes for the stimulus (the question) to activate the appropriate knowledge structures in the mind, and then possibly decide between two ideas if multiple structures were activated (e.g. both the expert framework and a misconception).

Bao's Model Analysis[8] a useful framework for discussing the use of mental models for physical reasoning, describing students' dominate conceptual frameworks for answering physics questions as "states." A student is described being in the "expert state" if they consistently answer questions using accepted physics ideas, the "alternative state" if consistently answer using an alternative mode, or a "mixed state" if use a mix of the two. Students who predominantly answer questions inconsistent with either model are labeled being in a "Secondary model state". Figure 1 illustrates these different regimes.

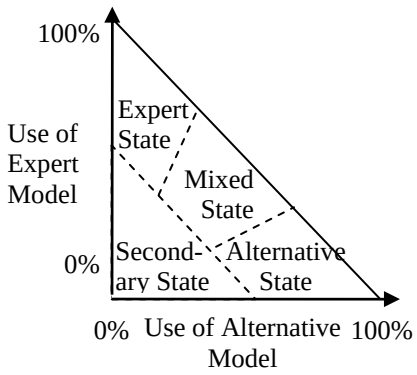

Figure 1: Regions of Student Model states in Model Analysis framework. After Bao[8].

SETTING AND DATA COLLECTION

Students in the first semester calculus- and algebra-based introductory physics at the author's institution took the FMCE. The FMCE was given the first week during class or laboratory time and then again the last or next-to-last week of the semester. Students were told that their performance would not affect their grade but were asked to do their best. On average, students took 17 minutes to complete the FMCE, ranging from 4 to 34 minutes.

Data used in this study was collected during the spring and fall of 2007 and spring 2008. Not all laboratory sections participated due to instructor reluctance or logistics issues (i.e. course related activities occupied the entire final laboratory period). Also, during this time some approximately half of the students answered the FMCE using paper forms as part of a different study to ensure the mode of administration did not affect performance. Paper administration did not include response times and so is not part of this work. Because of the different formats utilized and small numbers, all sessions (sets of responses by a student on an assessment) are included in this data set. Unless otherwise noted, all course levels and pre/post administrations are combined to increase the statistical pool.

Students completed the on-line assessments using a system developed by the author and described elsewhere[9]. Relevant features of this system are that the questions are presented to students one at a time, with the ability to skip ahead or go back, and that by default the system automatically advances to the next question as soon as the student clicks to mark response choice. Also, directions and other text that is the same for a group of questions is clearly delineated from question specific text by its placement in a lightly shaded box at the top of the question.

READING TIMES

Analysis

The FMCE[1] structure is more complex than some assessments. All but one of the questions form groups of three to eight questions with descriptive text, one or more images, and a set of response choices common to all questions in the group.

Linear regression was used to explore how different components of questions affected response times. Component included the number of characters in the direction/question text common to the entire group (n_{common}), characters in text specific to an individual question ($n_{question}$), characters in the text of answer choices ($n_{choices}$), number of graphs in the answer choices (n_{graphs}), and the number of images (n_{images}). With the exception of $n_{question}$, these all describe elements that are repeated within a group of questions, so were split into two variables depending if the question was first in the group. For example, n_{common_first} is the number of characters in common text if the question is the first in the group and otherwise zero, with the opposite for n_{common_later}.

The linear regression analysis was done in two stages because a significant covariance between several of the variables. In the first stage, $n_{question}$ and n_{common_first} were entered, since it would be expected

students should at minimum read the text of the individual question and the common text the first time presented. Both were significant at the at α = .05 level. The second stage consisted of a stepwise regression to determine which of the remaining variables also contributed to the model. Components were entered into the model if the coefficient was significant at α = .05 and removed if it became not significant at α =.10. Only the two image variables entered. The final analysis had a value of R^2 = 0.718 (explaining ~72% of variance in the data). The parameters of the regression are shown in Table 1.

Table 1: Significant components in explaining reading times in linear regression analysis.

Parameter	Value	*p*
constant	7 ± 2	< 0.005
$n_{question}$	0.03 ± 0.01	0.02
n_{common_first}	0.03 ± 0.01	< 0.005
n_{image_first}	11 ± 3	< 0.005
n_{image_later}	5 ± 1	< 0.005

Results and Discussion

The results suggest that students do systematically read question text when it is first presented but not answer choices. Both $n_{question}$ and n_{common_first} yielded values corresponding to reading speeds of about 300 words per minute, perhaps somewhat high, but within a plausible range. In contrast, neither $n_{options}$ variable was found to be significant, indicating that the time spent looking through answer choices is not dependent on to the amount of text present. The n_{graph} variables were also found to not be significant, indicating a lack of systematic looking through graph options. The insignificance of n_{common_later} indicates that repeated text is rarely re-read. The values for the *image* variables are higher than might be expected. This could be a result of the fact that the only groups of questions that do not include images are two groups that ask about force and acceleration of a coin being tossed into the air, a situation that most students would probably find quite familiar. Thus, the time associated with the n_{image} variables may be related to more effort required to understand or visualize the situation.

MENTAL MODELS

The relation of response times to the mental models utilized is a more challenging but valuable question. One way to organize data is by the mental model used on a particular question. A second way is the dominant model state[8] of the student during the assessment (see Figure 1). Both approaches were carried out for both Newton 2nd related questions (#1-21) and Newton 3rd related questions (#30-39)

First, response times were normalized to remove much of the variation of response times from question to question. This was done by dividing individual

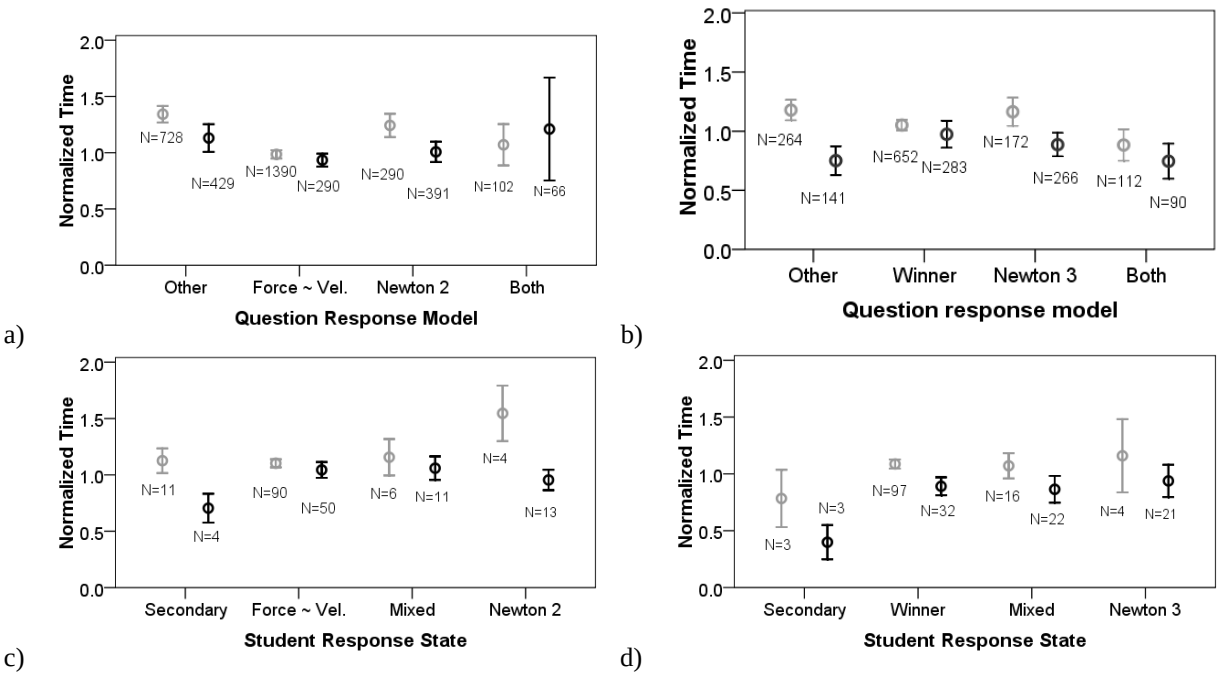

Figure 2: Mean normalized response times by response model and student state for Newton 2 questions (a, c) and Newton 3 questions (b, d). These are grouped by individual question responses and predominant model demonstrated by the individual students. Gray markers represent pre-instruction administration of the FMCE, and black markers represent post-instruction administration. Error bars represent 95% confidence level.

response times by the predicted response time calculated for each question utilizing the parameters shown in Table 1. It was found that this normalization did not differ greatly from dividing individual response times by the average response time for that particular question.

The response model and student model states where determined as follows. Each response was marked as to being consistent with using the expert model, the alternative model, both or neither. Data was then aggregated by session (student) to determine the number or responses consistent with both the expert model and the alternative model for both the Newton 2 and Newton 3 questions. The student model state was determined from this. Students who answered 50% or more of the questions inconsistent with either the expert or alternative model were classified in the secondary state. The remaining three states were determined by the relative frequency of using the expert vs. alternative models. The expert state was defined as using the expert model at least 2/3 of the time relative to the alternative model, and the alternative state as using that model 2/3 of the time relative to the expert model. The remaining students were considered to be in the mixed state.

Results and Discussion

Figure 2 shows plots of average normalized response times for the model used on individual question and by student model states. Results for both pre- and post-instructional administrations of the FMCE are shown. Several observations can be made.

First, on the question response data, there is a significant drop in response times between pre- and post-instructional administration on the expert and other models, but not on the alternative or both. The drop on the expert model would reflect that fact that this model would have been utilized and practiced during instruction, so at the end of the semester it is easier to activate it, while no change is made with respect to the alternative model, which would have been discouraged.

Second, with respect the student response state, the average response times for students in the mixed state differ little from those in the alternative and expert states. This shows that it is uncommon for students in the mixed state to have both expert and alternative ideas activated, requiring them to decide between them. This provides some support for the knowledge-in-pieces theoretical model; in the misconceptions framework a mixture of expert and alternative responses would suggest the student possesses both alternative and expert-like frameworks, so it might be expected that both would sometimes be activated.

A notable drop in the average response time on the secondary model state can be seen in both data sets. This may suggest that those in this category on the post assessment tend not to take the test seriously, However, only one individual on the post-assessment fell into this category on both Newton 2 and Newton 3 questions, and the small numbers make it difficult to draw any conclusions.

CONCLUSIONS

Latent response times can provide valuable, complementary information for analyzing student performance on assessments like the FMCE. An investigation as to how response times are related to text, images and graphs in questions indicates that students generally read systematically through question text when first presented to them, but not through text and graphs in the answer choices. Analysis with regard to mental models finds a drop due to instruction in times on expert-like responses but not on alternative ideas, presumably with as a result of practicing with expert ideas over the semester. Although students in the mixed state used two different models for answering questions, no evidence is seen both models are activated at the same time, as might happen under the misconceptions frame work for understanding student reasoning.

REFERENCES

1. Ronald K. Thorton and David R. Sokoloff, *American Journal of Physics* **66** (4), 338-352 (1998).
2. J. Ridley Stoop, *Journal of Experimental Psychology* **18**, 643-662 (1935).
3. D. T. Brookes, J. Mestre, and E. A. L. Stine-Morrow, "Reading Time as Evidence for Mental Models in Understanding Physics" in *Proceedings of the 2007 Physics Education Research Conference*, edited by L. Hsu, C. Henderson, and L. McCullough, AIP Conference Proceedings 951, Melville, New York, 2007, pp. 65-68.
4. Eric Jan Wangenmakers and Scott Brown, *Phsychological Review* **114** (3), 830-841 (2007).
5. Rachel E. Scherr, *American Journal of Phyiscs* **75** (3), 272-280 (2006).
6. John P Smith, Andrea A. diSessa, and Jeremy Roschelle, *The Journal of the Learning Sciences* **3** (2), 115-163 (1994).
7. Jim Minstrell, *Facets of Students' Thinking*, (University of Washington, 1996), *http://depts.washington.edu/huntlab/diagnoser/facetcode.html*
8. L. Bao and E. Redish, *Phys. Rev. ST Phys. Educ. Res.* **2** (1), 010103 (2006).
9. Scott Bonham, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010106 (2008).