Time to completion reveals problem-solving transfer

Rasil Warnakulasooriya & David E. Pritchard

Dept. of Physics & Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139.

Abstract. Two equally skilled groups of students taking introductory mechanics solve related physics problem pairs in reverse order with respect to each other, using the web-based Socratic tutor, MasteringPhysics. In tutorial problems containing help in the form of requestable hints, descriptive text, and feedback, twice as many students were able to complete problems correctly in real-time compared to problems that did not provide any help (end-of-chapter problems). The prepared group in a given related pair was able to solve it in ~15% less time on average compared to the unprepared group. Furthermore, the prepared group requests ~7% fewer hints on average than the unprepared group. We conclude that shorter completion times and problem-solving transfer are facilitated through tutorial problems.

INTRODUCTION

In this paper we provide evidence of learning and problem-solving transfer in physics as seen by time to completion curves using the web-based homework tutor, MasteringPhysics. The time to completion curves depict the fraction of students who have completed a given physics problem correctly versus the logarithmic time for the completion of that problem.

A study of time to completion of physics problems is enabled by the general mastery learning [1] pedagogy of MasteringPhysics and the availability of the log files it generates. Such studies are crucial for maximizing learning for a given amount of student time on task. In MasteringPhysics mastery learning is implemented within a Socratic dialogue where students are provided with hints and simpler sub-problems upon request, and are given feedback when wrong answers are proposed.

THE STUDY

The data in this study involve students from the fall 2003 semester taking the introductory Newtonian mechanics course at the Massachusetts Institute of Technology (MIT). The students (~ 430 in total) were divided into two equally skilled groups based on their MasteringPhysics homework assignment grades for the first six weeks of the semester. Our study involves these two groups solving related problem pairs in opposite orders relative to each other with no intervening problems. That is, the "first" problem in a given pair to one group is the "second" to the other and vice versa. We will call the group that solves a given problem in a pair first, the unprepared group and the group that solves the same problem second (having solved its related problem first), the prepared group.

The problems in a pair are related in the sense that they both involve the same concepts and methods. Seven problem pairs in the concept domains of torque, linear momentum conservation, energy conservation, friction, angular kinematics and dynamics, rigid body rotation, and Newtonian gravitation were studied. Since these problems involve requestable hints and wrong answer responses, we will call them problems with help or tutorial problems. Five end-of-chapter (EOC) problems [2] that do not contain any hints or wrong answer responses were also administered through MasteringPhysics in the domains of energy conservation and rigid body rotation.

The main variable of interest in this study is time to completion, which is defined as the time interval between a student's first opening of a problem and his/her submission of the completed problem. No time intervals for interactions in between (even for logins and log-offs) are accounted for. Completion is defined as finishing all the main parts of a given problem correctly (finishing sub-parts when hints are requested is not required). We chose to bin the data in time intervals of 0.5 using a logarithmic time scale (a factor of 1.6 per bin in real time).

RESULTS AND CONCLUSIONS

1. **Three groups:** We identify three major groups of students in completing a given problem by plotting the rate of completion against logarithmic time (see Fig. 1). The three groups occur, consistently separated by a local minimum, in all fourteen tutorial problems.

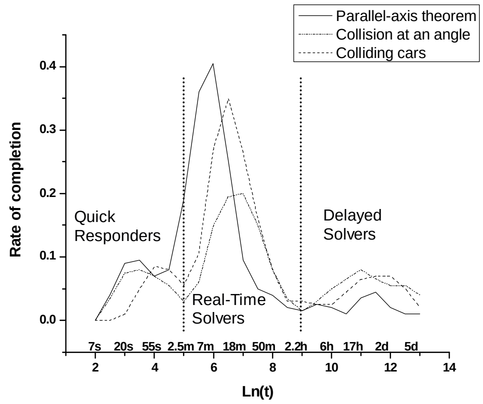

FIGURE 1. The rate of completion graphs for three representative tutorial problems reveal three distinct peaks. In the order of increasing time, these peaks correspond to the quick responders, the real-time solvers, and the delayed solvers.

The "quick responders" (~10% of the students in the course) are able to finish the problem within 2.5 minutes. These students typically do not request any help (hints) or submit wrong answers. The peak that occurs after about 2.2 hours, we identify as "delayed solvers." These students request help within MasteringPhysics and submit incorrect answers (see Fig. 2). We believe that the majority of both the quick responders and the delayed solvers seek assistance outside MasteringPhysics. This may indicate intellectual dishonesty on the part of quick responders and/or help from human tutors for the delayed solvers. Indeed, further analysis show that the more quick responses a student submits, the more his/her score is below average in the paper-based final examination.

The most interesting group from the view point of problem-solving, we believe, is the "real-time solvers" – they complete the problem in the interval from 2.5 minutes to 2.2 hours, generally using hints and feedback (For two gravitation problems this peak is shifted towards longer times and occur between 7 minutes to 6 hours). This group corresponds to the prominent central peak in rate of completion graphs. Since these students seem to struggle with the problem by giving wrong answers and requesting hints within reasonable time intervals we believe that they solve problems in real-time (see Fig. 2).

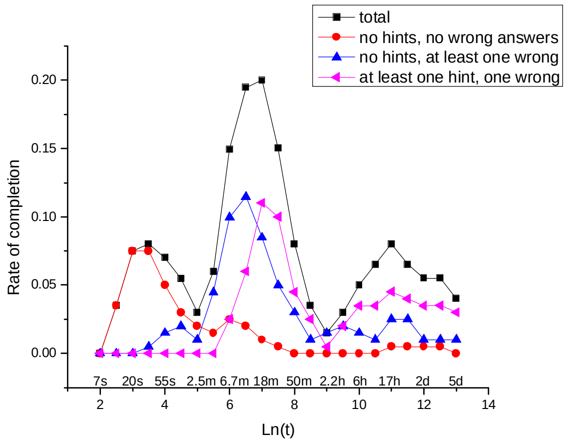

FIGURE 2. The break-down of the rate of completion graph for a typical tutorial problem. The real-time and the delayed solvers make mistakes and ask for help.

The PER research community (and others) defines "problem" (as opposed to "exercise") fairly consistently – the route to solution is not immediately apparent to the student. We therefore feel these questions are "problems" for the real-time solvers since the central peak mainly consists of students who give wrong answers and/or request hints (at the expense of a penalty of their homework grade), and only a small minority answer without either of these indications of struggle.

2. **Sigmoidal shape for "real-time solvers":** The time to completion graphs for real-time solvers (which are the integral curves of the rate of completion curves) yield S-shaped (sigmoid) curves for all problems with help (see Fig. 3). This observation is interesting since the sigmoid curve plays a prominent role in growth models [3] and time to completion of other psychological tasks [4,5] from sentence completion [6] to pigeon learning [7]. These curves are important for short term and long term objectives. Short term, they indicate whether certain instructional strategies will help more students obtain the correct solutions with less expenditure of time, and enable measurements of knowledge transfer per unit time. Longer term, they will seriously constrain models of human problem solving which must be able to predict the detailed shape observed (as is already the case for pigeon learning [7]).

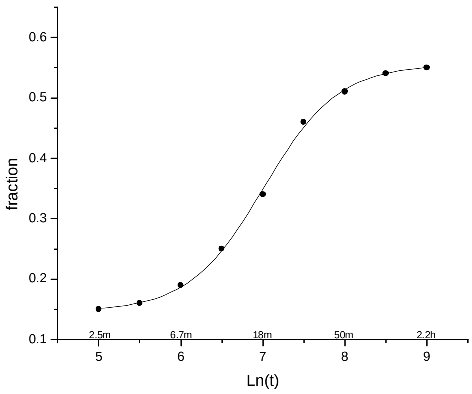

FIGURE 3. The integral of the rate of completion curve of a typical tutorial is sigmoidal or S-shaped. This shape is a characteristic of psychological tasks.

3. **Logistic and Boltzmann fit best:** The high signal to noise ratio of our S-shaped time to completion curves for real-time solvers invites a more detailed fit. For this, we picked three sigmoidal functions even though we have no a priori reason (e.g. they do not result from plausible "growth models" of problem solving):

Logistic $$f_L(x) = \frac{a-b}{1+(\frac{x}{x_C})^p} + b$$

Boltzmann $$f_B(x) = \frac{a-b}{1+e^{(x-x_C)/d}} + b$$

Gompertz $$f_G(x) = ae^{-e^{-k(x-b)}}$$

These were considered as both functions of time (x = t) and functions of logarithmic time [x=Ln(t)]. We fitted the differentials of the above functions to the rate of completion curves in order to eliminate the interdependency among the data taken from different times, and to remove the effects of the quick responders and delayed solvers. This also allowed us to use the reduced chi-square (χ^2/ν) goodness-of-fit test as well as the more sophisticated second order Akaike information criterion (AIC_c) [8,9]. We find that the Logistic and the Boltzmann functions using x=Ln(t) are both good fits (χ^2/ν=1.15) and are preferred for our data (i.e. ΔAIC_c >4) relative to the Gompertz or to any function of linear time.

In order to see if our finding of the preference for x=Ln(t) has precedents, we reanalyzed the human word recognition data in sentence completion reported by Tulving et al.[6]. These data, when plotted against a logarithmic time, were generally fitted better by a linear model (evidence ratio > 100) than with any of the above sigmoidal functions. We also find, when considering the data in linear time (as originally reported by Tulving et al.), that the Gompertz model seems more likely (ΔAIC_c >4) than the linear, the Logistic, or the Boltzmann. The only precedent we have found for the superiority of Ln(t) is in computer adapted testing [10] where a reanalysis of the data seems to favor this independent variable. In this context we note that it is for complex tasks like the problems given in our study but unlike the simple tasks in [4-7].

4. **Lack of hints reduces "real-time solvers."** We find clear differences in time to completion curves between tutorial problems (which contain hints and feedback to specific wrong answers) and the EOC problems (which do not). The S-shape typically becomes linear in EOC problems for the real-time solver interval. In Fig. 4 we provide an example of this effect for an EOC problem and compare with a tutorial problem both involving the concept of energy conservation. We find that the fraction of real-time solvers is 67±3% for 12 problems with help but is only 33±5% for five EOC problems (see Fig. 4). We feel that these problems are sufficiently equal in difficulty that this factor of two difference almost certainly results from the presence of tutoring.

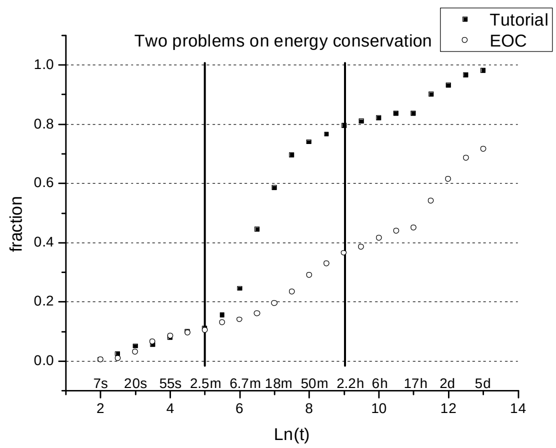

FIGURE 4. Twice as many students are able to solve a given tutorial problem with hints and wrong answer responses in real-time (2.5 min. – 2.2 hrs) compared to an end-of-chapter (EOC) problem.

5. **Problem-solving transfer to related problems:** We find that learning and problem-solving transfer [11] takes place from one related problem to another in the sense that the prepared group have a time advantage of 14.6±2.2% on average over the

unprepared group on the same problem. This effect is seen across all the fourteen problems thus providing robust statistical evidence. We arrive at this result by fitting the differentials of the above sigmoidal functions to the rate of completion data and identifying the time corresponding to the peak rate for the real-time solvers (see Fig. 5 for an example). This method also reveals that the median time on a typical problem is about 15 minutes, which is quite realistic (there are only a few students at the margins of the time interval 2.5 minutes to 2.2 hours). We also find that averaging over six pairs of problems, the fraction of prepared students requesting hints is 7±2% lower than the unprepared students. In other studies [12] we have identified quicker time to solution and use of fewer hints with better performance on the final exam.

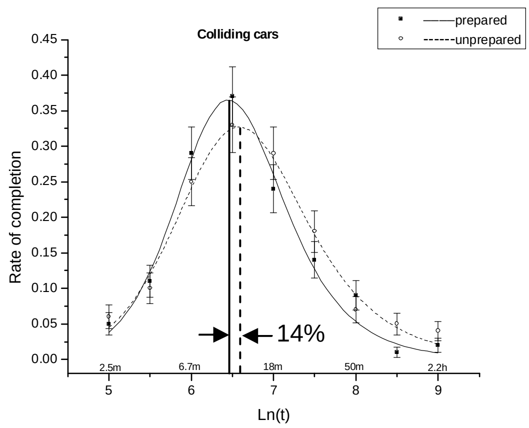

FIGURE 5. The prepared group is able to solve a given tutorial problem in less-time compared to the unprepared group. The average advantage in time seen across fourteen problems is 14.6%.

DISCUSSION

We have demonstrated a methodology of extracting useful information about time on task in an uncontrolled homework environment – namely, measuring the fraction complete against the variable logarithmic time for problem completion. Identifying the real-time solvers we have demonstrated that their time to completion curve follows the sigmoidal shape against logarithmic time rather than linear time as observed in other psychological tasks. Furthermore, we have shown evidence that the problems with tutorial help can be solved in a reasonable time by twice as many students as end-of-chapter problems. We also find robust evidence of problem-solving transfer as evidenced by the reduction in time to solution having solved a related prior problem across fourteen problems covering seven concept domains. The effectiveness of the tutorial problems can be attributed to the hints and wrong answer responses leading to learning in much the same way Simon and Reed [13] have observed in the context of the "cannibals and missionaries" problem.

We note that our results are not biased since the effects we see are reciprocal; that is the prepared group for a given problem in a pair is the unprepared group for its related problem and vice versa. The various comparisons we have made in this study are general in nature in the sense that the problems are related only by content and the relevant conceptual domain. Also, the tutorial problems and the EOC problems are not isomorphs of each other. Since our larger goal is to investigate and establish best pedagogical practices in web-based tutoring, a study designed to look at the isomorphic problems with and without hints and wrong answer responses is being carried out. The findings from these studies will be reported in the future.

[We thank the National Science Foundation for providing assistance through its grant 0231268.]

REFERENCES

1. B. S. Bloom, *All our children learning*, McGraw-Hill, New York, 1981.
2. H. D. Young and R. A. Freedman, *University physics*, Pearson, San Francisco, 2004.
3. K. P. Burnham and D. R. Anderson, *Model selection and multimodel inference: a practical information-theoretic approach*, Springer, New York, 2002.
4. C. L. Hull, *Principles of behavior: An introduction to behavior theory*, Appleton-Century-Crofts, New York, 1943.
5. D. L. Schnipke and D. J. Scrams, *Journal of Educational Measurement* **34**, 2132 (1997).
6. E. Tulving, G. Mandler, and R. Baumal, *Canadian Journal of Psychology* **18**, 62 (1964).
7. A. Machado and P. Guilhardi, *Journal of the Experimental Analysis of Behavior* **74**, 25 (2000).
8. H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Second international symposium on information theory*, edited by B.N. Petrox, F. Caski, Budapest, 1973, pp. 267-281.
9. N. Sugiura, *Communications in Statistics, Theory and Methods* **A7,** 13 (1978).
10. L. F. Hornke, *Psicológica* **21,** 175 (2000).
11. R. E. Mayer, *Learning and instruction*, Pearson, New Jersey, 2003.
12. D. E. Pritchard and R. Warnakulasooriya, *EdMedia* (2005) – To appear.
13. H. A. Simon and S. K. Reed, *Cognitive Psychology* **8**, 86 (1976).