# Time to completion of web-based physics problems with tutoring

Rasil Warnakulasooriya, David J. Palazzo, and David E. Pritchard
Department of Physics & Research Laboratory of Electronics
Massachusetts Institute of Technology, Cambridge, MA 02139.

Rasil Warnakulasooriya, Research Laboratory of Electronics, Massachusetts Institute of Technology; David J. Palazzo, Research Laboratory of Electronics, Massachusetts Institute of Technology; David E. Pritchard, Research Laboratory of Electronics, Massachusetts Institute of Technology and founder Effective Educational Technologies.

Correspondence concerning this article should be addressed to Rasil Warnakulasooriya, Research Laboratory of Electronics, 26-251, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. E-mail: rasil@mit.edu

Abstract

We study students performing a complex learning task- solving multipart physics problems with interactive tutoring. We extract the rate of completion and fraction completed as a function of time on task by retrospectively analyzing the log of student-tutor interactions. We find a spontaneous grouping into three groups of students, the central (and largest) group being those who solve the problem in real time after multiple interactions (primarily submitting wrong answers and requesting hints) with the tutorial program. This group displays a sigmoidal fraction completed curve as a function of *logarithmic* time. This sigmoidal shape is qualitatively flatter for problems that do not contain hints and wrong answer responses. We conclude that the group of students who respond quickly is obtaining the answer from some outside source. The third group represents those who interrupt their solution; presumably to obtain outside help.

The time necessary to complete a task is an important property of that task. This is well recognized in education: years required to obtain a given degree, days of school per year, hours of class time per semester, and minutes on a given test are common metrics. Time to completion is arguably of equal importance with regard to homework tasks since homework often dominates the total time spent on a course especially when the grade level increases (Cooper, 1989). Unfortunately, the unpredictable location and time at which students actually do their homework combined with its unsupervised nature make such studies impractical. Online homework administration and tutoring programs now provide opportunities to collect time data on homework tasks. This study represents an initial investigation of such data and provides novel insight into student behavior vis a vis homework.

It is well known that "time on task" correlates strongly with the amount learned (Long, 1983; Pintrich, 1988; Wellman & Marcinkiewicz, 2004). However, since students can spend only a limited amount of time on homework, it is important to study the allocation of this important resource in part to optimize learning per unit time. Studies of the time taken to complete homework tasks are necessary for this, and are valuable for other reasons also: identifying tasks that in the instructor's judgment are taking too long, predicting the time students will take on a test or on the entire homework assignment, and identifying students who may be receiving outside assistance with the tasks or who allocate their time inefficiently.

While detailed studies of time spent on homework tasks is sparse, the time to complete a task, and the rate at which this time is reduced by instruction (often repetitive practice) – called the "learning curve" - are common variables in psychology involving both humans and animals (Coombs, 1970; Luce, 1986). Such findings show that performing simple tasks quickly is an indication of skill in the task domain; hence we might expect time to completion of homework problems to be an indication of skill in the associated knowledge domain (Pritchard & Warnakulasooriya, 2005).

Here we investigate the time students take to complete tutorial problems in a physics on-line homework environment. "Tutorial problems" here means multi-part problems that contain both spontaneous responses to incorrect answers and requestable hints that help the students along an expert route to solution. Some of the hints are sub-problems representing useful checks of procedures and results toward the solution. Some tutorial problems include tutorial text preceding the problem.

Surprisingly, we have discovered that students do not complete these problems at an even rate, but that the rate of students completing such problems typically has two minima that divide the students into three distinct major groups: the large majority of students are in the central group which we show are "real time solvers." This group is preceded by "quick responders" most of whom we argue are simply copying the answer from elsewhere. The third group solves the problem after a delay during which we suspect they receive help from some outside source.

We show that the overall fraction completed within the real time solvers group follows a sigmoidal curve similar to earlier studies on simpler tasks in both humans and animals (Bizo & White, 1995; Hull, 1943; Hunt, 1975; Machado & Guilhardi, 2000; Machado & Keen, 1999; Schnipke & Scrams, 1997; Tulving, Mandler, & Baumal, 1964). Unlike their data, we find that the curves fit standard sigmoidal model functions better when we use logarithmic time rather than time itself as the independent variable. The real time solvers group further divides into a group that makes only wrong responses on the way to solution and a somewhat slower group that also requests hints.

### *The Study*

The study we report here involved about 400 students taking "Introductory Newtonian Mechanics" course at the Massachusetts Institute of Technology (MIT) in fall 2003. The students were assigned homework problems (we argue later that they are not exercises for most students) using Mastering Physics, a web-based homework tutor (www.masteringphysics.com).

Students can work the assigned problems at any time over a several day period prior to the due date. There is no requirement that the problems be completed in one sitting. The problems of interest were given in pairs and two equally skilled groups of students must sequentially solve each of the problems in 6 pairs of related problems. (The sequential requirement enables another study on knowledge transfer, but has no significant effect on the present study – the students may work all other problems in the assignment in any order.) The study consists of 6 problem pairs or 12 problems and since each problem is solved by two equally skilled groups, we have 24 instances in total. The two instances for a given problem will be labeled instance 1 (I1) and instance 2 (I2).

Most of the assigned problems are tutorial in nature and were interactive. They contain supplementary information that enables a Socratic dialogue where students are provided with hints and simpler sub-problems upon request, and are given criticism (feedback) when specific incorrect answers are proposed. This pedagogy approximates mastery learning (Bloom, 1981) in that less skilful students travel a longer path, but eventually over 90% of all students obtain the correct solution. Guessing behavior is discouraged by the use of free response answer types for a majority of all requested responses. Also, a penalty is applied for incorrect submissions. Furthermore, hints requests were also penalized to promote thinking before requesting help. End-of-chapter (EOC) problems based on Young & Freedman (2004) that did not contain hints or feedback were also administered through Mastering Physics for comparisons with the tutorial problems.

Student interactions with Mastering Physics are then analyzed for time to completion, the number of hints requested and the number of wrong answers submitted. Time to completion in this paper is defined as the time duration between the first opening of a problem and the submission of the completed problem. No allowance for interactions in between (even for logins and log-offs) was made. Completion is defined as finishing all the main parts of a given problem correctly. (It is not required that students solve requested sub-problems, although there is a stiff penalty for requesting the solution to them.)

### ***Three Peaked Rate of Completion***

Time to completion ranges from several seconds to several days for the problems studied. Thus, we use a logarithmic time (Ln(t) where t is measured in seconds) as the independent variable, which in this study runs from 2 to 13. This corresponds to a time scale from about 7 seconds to 5 days. We chose to bin the data in time intervals of 0.5 in logarithmic time scale (a factor of 1.6 in real time). We studied both the "rate of completion" – i.e. the fraction of students who have completed a given problem during that interval - and the time integral of this curve, the fraction of students who have completed the problem at the end of that time interval after opening it. (The fraction is relative to all the students who have opened the problem.)

The major discovery reported here is that the rate of completion graphs reveals three distinct groups of students; i.e. they are characterized by a three-peak pattern that is common to all the 12 problems that contained help (requestable hints, feedback, and advice within the problem text). Figure 1 shows the three-peak pattern for three representative problems for the entire student pool. The first peak consists of "quick responders" who predominantly answer correctly on the first attempt with a median time to completion of $1.4 \pm 0.3$ min, with the local minimum between this group and the real time solvers being typically at 2.5 min. Finally, a group of "delayed solvers" is separated from the real time solvers by a broader local minimum typically centered on 2.2 hours. These "delayed solvers" display long periods of inactivity, suggestive of absence from their computer terminal. Their median time for problem completion is $21.5 \pm 5.4$ hours. We shall argue that most members of both of these groups receive help outside the web-based tutor. Ten of twelve problems suggest that the time duration from 2.5 minutes to 2.2 hours is a reasonable duration in identifying the "real time solvers" marked by the prominent central peak. (For two problems involving gravitation the central peak or the real time solver duration is shifted to the right and is within 7 minutes to 6 hours.) The median time to completion, hints per user, incorrect responses per user, and the total interactions per user for the three groups of students are given in Table 1. The data indicate that the quick responders have a significantly different behavior from that of the real time and delayed solvers.

We now argue that the "quick responder" group is largely made up of students who obtain the answer without either inspecting or working the problem themselves, for example from a student who has previously worked the problem. This belief is buttressed by two observations. First, skillful expert problem solvers in this domain take longer than 2.5 minutes to solve the problems given and display a higher error rate than the quick responder students when attempting to work problems this rapidly: one of the authors (DEP) took 7.6 ± 2.5 min on average for five representative problems whereas the quick responder students took 1.6 ± 0.2 min for the same five problems. (The real time solvers averaged 20.4 ± 5.3 min for the same five problems.) Second, there is a strong *negative* correlation between solving a higher percentage of tutorial problems as a quick responder and doing well on the final exam. Figure 2 shows that students who respond quickly to more than 50% of the problems average 1.3 standard deviations below the class average in the paper-based final examination, hardly the behavior expected for students who are truly able to solve the majority of problems quickly and with far fewer mistakes than average students in the course.

Turning to the delayed solvers we find that the fraction of delayed solvers who has given either a correct or an incorrect first response to a main part of a problem within 2.5 min is 13.8 ± 3.9%. By 2.2 hours this fraction is only 39.6 ± 5.6%. This suggests that about 60% of the delayed solvers give their first response 2.2 hours or more after opening the problem. Also, by 2.2 hours only 33.3 ± 4.1% of all their interactions are registered. Thus, over 65% of the interactions with the tutor for the delayed solvers take place after 2.2 hours. These interactions include hint requests, submitting solutions (correct/incorrect) to main/sub-parts, and requesting solutions. This suggests that many delayed solvers decide to postpone their attack on the problem (or to seek outside help) upon initial inspection of the problem, or possibly open it only to print it for later attempts. Indeed, within the delayed solvers group, there are a subset of students who make no interactions with the problem before closing and opening another problem. For ten studied problems, this group comprises of about 35% of the delayed solvers. When this group returns to solve the bypassed problem about 20% make no interactions with the tutor other than to submit the correct answer. The other delayed solvers make multiple interactions. The mean duration of problem completion for these students is 2.3 ± 0.3 hours, just about the top of the transition from a real time solver to a delayed solver.

***Real Time Solvers***

The students who fall within the dominant central peak give evidence of solving the problems in real time and are denoted as *real time solvers*. They interact with the tutoring features of the online environment. Unlike the quick responders, who rarely make mistakes and almost never ask for hints, 85% of the real time solvers submit at least one incorrect response, and these divide into two subgroups of students – those who requested hints and those who did not (Fig. 3). As seems logical, those who request hints take about 4% longer to finish the problem.

The results shown in Figure 3 support our use of the term “problem” to describe these tasks at least for the real time and delayed solvers. In the problem solving literature, a “problem” is distinguished from an “exercise” in which the route to solution is apparent to the student upon initial inspection (Kahney, 1993; Mayer, 2003; Smith, 1991). The observation that the vast majority of students submit at least one wrong answer, that typically one half feel sufficiently confused to sacrifice some credit in order to receive a hint, and that they take substantially longer than expert solvers to solve the problem all argue that they are not able to treat the tasks as an exercise.

The large sample size (n ~ 400) allows us to study the detailed shape of the fraction complete curves. This leads to the novel observation that the fraction completion curves for the real time solvers (which are the integral curves of the rate of completion curves) have a sigmoidal shape if plotted against *logarithmic* time. This consistently occurs in all 24 instances with tutorial help (Fig. 4). Our data have sufficiently high signal-to-noise ratio that we can explore the detailed shape of the sigmoidal curves.

We explored two aspects of the sigmoidal curves in detail: first whether the apparent preference for Ln(t) as the independent variable is statistically significant, and second, which particular functional form of the sigmoidal curve best represents the data. We considered three functional forms: Logistic (Eq. 1), Boltzmann (Eq. 2), and Gompertz (Eq. 3).

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

Boltzmann $$f_B(x) = \frac{a-b}{1+e^{(x-x_c)/d}} + b \qquad (2)$$

Gompertz $$f_G(x) = ae^{-e^{-k(x-b)}} \qquad (3)$$

where $x$ = Ln(t) or t for independent variable being logarithmic or linear time, respectively. To remove bias to the sigmoid curve in the time range of 2.5 min-2.2 hours that arises from the presence of quick responders and delayed solvers, and so that each point (i.e. time interval) affects the result independently of the others, we fit the rate of completion data with the *differentials* of the above models assuming a Poisson error variation for each point (i.e. error proportional to square root of the number of students in that time bin).

We used two statistical tools to explore the fitting: the chi-square goodness-of-fit coupled with the second order Akaike Information Criterion or $AIC_c$ (Akaike, 1973; Sugiura, 1978) which is a relation between the Kullback-Leibler distance (Kullback & Leibler, 1951) and the maximized log-likelihood function. According to Burnham and Anderson (2002), “AIC provides a simple, effective, and objective means for the selection of an estimated ‘best approximating model’ for data analysis and inference.” AIC penalizes a model with more parameters which guards against overfitting according to Pitt and Myung (2002). The results are shown in Table 2.

Sigmoidal (S-shaped or psychometric) curves have been observed in various psychological tasks (Bizo & White, 1994; Hull, 1943; Machado & Guilhardi, 2000; Machado & Keen, 1999; Schnipke & Scrams, 1997; Tulving, Mandler & Baumal, 1964), but as a function of linear rather than logarithmic time (i.e. taking $x$ = t in Eqs. 1-3). Therefore, we fitted our data with the sigmoidal functions in Eqs. 1-3, but considered as differential functions of linear (rather than logrithmic) time. For the 24 cases, the median superiority of the logarithmic time in (unreduced) chi-squared was 3.7 and the AIC evidence ratio was 10 – deemed a considerable difference (Burnham & Anderson, 2002). Although there were 4 cases for which linear time was slightly better, the median chi-square difference was 1 and the AIC evidence ratio was 2 - not enough to rule-out the logarithmic time (or accept the linear time) for these minority cases. In summary, the differential model in logarithmic time is statistically superior than in linear time in accord with the fact that the prominent symmetrical ‘S’ shape is present to the eye only with respect to logarithmic time (Fig. 4 and Fig. 5).

Turning now to the question of which function of Ln(t) fits our data, the Logistic and Boltzmann models are significantly preferred for our data relative to the Gompertz in 23 of 24 cases. The median reduced chi-squared ($\chi_v^2$) is of the order of 1, suggesting that the Logistic and Boltzmann functions account for essentially all information in the data. This judgment is confirmed by noting that in cases where the reduced chi-square is ~2 or greater the data lie above the curve at one end or the other of the range, a likely indication that some students belonging to the quick responders or the delayed solvers “spill over” into the real time solver group (Fig. 6).

In summary, the sigmoidal curves observed here differ significantly from those observed earlier in that they are sigmoidal functions of Ln(t) rather than linear t. The problems studied require multiple responses and typically take several minutes to be solved compared with the type of simple recognition tasks or pigeon learning tasks, which occur on a scale of several seconds or less (Bizo & White, 1994; Tulving, Mandler & Baumal, 1964). This suggests that the shape of these curves depends significantly on the type of mental task which in turn suggests the possibility that explicit mathematical models of learning and problem solving (Bizo & White, 1995; Coombs, 1970; Hull, 1943; Machado & Guilhardi, 2000; Atkinson, Bower & Crothers, 1965; Friedman et al, 1995; Luce, 1986) will provide connections between the types of mental processes and the shape of the time to completion curves.

A further indication of the significance of the shape of the fraction complete curves comes from a comparison of the above curves with those obtained on problems that lack the tutoring described above. These are typical end of chapter (EOC) problems that do not contain hints or spontaneous specific responses to students’ wrong answers (i.e. the student just receives the message “try again”). Several such curves are shown in Fig. 7. It is obvious that, these curves lack the steep central part observed for tutorial problems (Fig. 8). In fact the fraction of students who are able to complete tutorial problems within the real time solver interval is 65 ± 4% for 10 problems compared to 33 ± 5% for 5 EOC problems within the same interval. This may possibly indicate that some students are “getting stuck” and subsequently waste time or go away from their computers to seek help. The key point is that the shape of the curve reveals a strong qualitative difference that most probably arises due to the hints and wrong answer responses in the tutorial problems.

### *Discussion*

This study shows that the time to completion provides valuable insight into how students are approaching problems in their online homework assignment. We have shown that it allows one to identify students who are solving a particular problem in real time or after a delay preceded by several hours and perhaps a short attempt at a solution, and to separately identify students who are obtaining the solution from outside sources without making an honest attempt to solve it (quick responders) or after an unsuccessful struggle (delayed solvers).

Identifying the quick responders provides a way of determining how many students (and which ones) are engaging in intellectual dishonesty. Moreover we have shown that such behavior is strongly anti-correlated with success on the final examination (at least at MIT). While this may reflect the fact that these students are denying themselves the opportunity to learn the material from the homework tutor, it is also possible that they find this material too difficult and time-consuming and therefore resort to cheating in order to keep the total time spent on their homework within tolerable bounds. The opportunity to study cheating directly reveals the power of analyzing web-based homework and will allow us to go beyond typical research on academic dishonesty that is based on self-reporting (Jensen et al, 2002; Murdock, Hale & Weber, 2001).

We have found that the real time solvers display a sigmoidal shaped fraction complete curve if Ln(t) is taken as the independent variable. This is a novel finding because previous psychological studies generally reveal a sigmoidal shape against linear time. This difference may well indicate that solving complex tasks like physics problems is different from solving simple tasks such as memory recall. Our results show additional discrimination in that the Boltzmann and Logistic models are significantly better fits than is the Gompertz; however we have no explanation for this finding. Clearly more research is needed on the connection between mathematical shape of the fraction complete and the nature of the mental task, for example a mathematical model of the cognitive process of physics problem-solving (Bocaneala & Bao, 2004; Newell, 1990; Pirolli & Wilson, 1998).

The fact that the real time solver group is twice as big on the tutorial problems as for the EOC problems (which don't have tutorial help) is a strong suggestion that the tutoring dramatically reduces the time it takes students to work problems. While the EOC and tutorial problems were not variants of the same problem, they otherwise seem to us to be of comparable difficulty. This suggests that the hints and wrong answer responses generate significant reduction in time to solution, and a several-fold reduction in the fraction of delayed solvers (i.e. those who have to go elsewhere for help after getting stuck). More research is needed here.

Finally, we note that the presence of delayed solvers for a tutorial problem may indicate that the students are unable to solve it with the aid of only the available tutorial hints and wrong answer responses (however, some students may elect to seek outside help rather than accept the hint penalty). Together with the fact that the shape of the time curves of EOC problems differs systematically from that of the tutorial problems, the presence of delayed solvers may allow problem authors to identify (and rectify) problems for which the current hints and wrong answer responses are inadequate.

A significant aspect of this study is that its results are obtained during the regular course of instruction in the area (homework) that consumes the largest fraction of students' time and hopefully imparts the most learning. The ability to isolate the real time solvers and track their actions offers an opportunity to make educational studies with very large samples and correspondingly high signal-to-noise ratio. Importantly, improvements (e.g. in learning per unit time) found in further studies will be directly applicable to future students' homework without concern for extensive further development, class testing, and issues of scalability.

[This work was funded by the National Science Foundation under grant 0231268. ]

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrox & F. Caski (Eds.) *Second International Symposium on Information Theory* (pp. 267-281). Budapest: Akademiai Kiado.

Atkinson, R. C., Bower, G. H., & Crothers, E. J. (1965). *An Introduction to Mathematical Learning Theory*, New York: Wiley.

Bizo, L. A., & White, K. G. (1994). The behavioral theory of timing: Reinforcer rate determines pacemaker rate. *Journal of the Experimental Analysis of Behavior*, **60**, 19-33.

Bizo, L. A., & White, K. G. (1995). Biasing the pacemaker in the behavioral theory of timing. *Journal of the Experimental Analysis of Behavior*, **64**, 225-235.

Bloom, B. S. (1981). *All Our Children Learning*, New York: McGraw-Hill.

Bocaneala, F., & Bao, L. (2004). A computational model for physics learning, *2003 Physics Education Research Conference*, AIP Conference Proceedings **720**, (pp. 117-120).

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*, New York: Springer-Verlag.

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970) *Mathematical Psychology: An Elementary Introduction*, New Jersey: Prentice-Hall.

Cooper, H. (1989). *Homework*, New York: White Plains.

Friedman, D., Massaro, D. W., Kitzis, S. N., & Cohen, M. M. (1995). A comparison of learning models. *Journal of Mathematical Psychology*, **39**, 164-178.

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*, New York: Appleton-Century-Crofts.

Hunt, E., Lunneborg, C., & Lewis, J. (1975). What does it mean to be high verbal? *Cognitive Psychology*, **7**, 194-227.

Jensen, L. A., Arnett, J. J., Feldman, S. S., & Cauffman, E. (2002). It's wrong, but everbody does it: Academic dishonesty among high school and college students. *Contemporary Educational Psychology*, **27**, 209-228.

Kahney, H. (1993). *Problem solving: Current issues*, Buckingham: Open University.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, **22**, 79-86.

Long, D. R. (1983). Time-on-task in beginning students of Spanish at the Univeristy level: A case study, Paper presented at the Annual Meeting of the American Council on the Teaching of Foreign Languages, ERIC Document Reproduction Service No. ED 239 514. San Francisco, CA.

Luce, R. D. (1986). *Response times: their role in inferring elementary mental organization*, New York: Oxford University.

Machado, A., & Keen, R. (1999). Learning to time (LET) or scalar expectancy theory (SET)? A critical test of two models of timing. *Psychological Science*, **10**, 285-290.

Machado, A., & Guilhardi, P. (2000). Shifts in the psychometric function and their implications for models of timing. *Journal of the Experimental Analysis of Behavior*, **74**, 25-54.

Mayer, R. E. (2003). *Learning and Instruction*, New Jersey: Pearson.

Murdock, T. B., Hale, N. M., & Weber, M. J. (2001). Predictors of cheating among early adolescents: Academic and social motivations. *Contemporary Educational Psychology*, **26**, 96-115.

Newell, A. (1990). *Unified theories of cognition*, Cambridge, MA.: Harvard University.

Pintrich, P. R. (1988). A process-oriented view of student motivation and cognition. In J. Stark, L. Mets (Eds.) *Improving teaching and learning through research* (pp. 65-70). San Francisco, CA: Josey-Bass.

Pirolli, P., & Wilson, M. (1998). A theory of the measurement of knowledge content, access, and learning. *Psychological Review*, **105**, 58-82.

Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, **6**, 421-425.

Pritchard, D. E., & Warnakulasooriya, R. (2005). Data from a web-based homework tutor can predict student’s final exam score, *EdMedia – World conference on educational multimedia, hypermedia and telecommunications*, Proceedings (pp. 2523-2529).

Schnipke, D. L., & Scrams, D. J. (1997). Modeling item response times with a two-state mixture model: A new method of measuring speededness. *Journal of Educational Measurement*, **34**, 213-232.

Smith, M. U. (1991). A view from biology. In M. U. Smith (Ed.) *Toward a unified theory of problem solving* New Jersey: Hillsdale.

Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics, Theory and Methods*, **A7**, 13-26.

Tulving, E., Mandler, G., & Baumal, R. (1964). Interaction of two sources of information in tachistoscopic word recognition. *Canadian Journal of Psychology*, **18**, 62-71.

Wellman, G. S., & Marcinkiewicz, H. (2004). Online learning and time-on-task: Impact of proctored vs. un-proctored testing. *Journal of Asynchronous Learning Networks*, **8**, 93-104.

Young, H. D., & Freedman, R. A. (2004). *University Physics*, San Francisco: McGraw-Hill.

Table 1

| | Quick responders | Real time solvers | Delayed solvers |
|---|---|---|---|
| Median time | 1.4 ± 0.3 min | 16.9 ± 8.6 min | 21.2 ± 5.8 hours |
| Hints/user | 0.4 ± 0.1 | 2.7 ± 0.6 | 3.8 ± 0.9 |
| Incorrect/user | 0.7 ± 0.1 | 4.1 ± 0.5 | 5.2 ± 0.8 |
| Interactions/user | 6 ± 1 | 13 ± 2 | 15 ± 2 |

Table 2

| Problem | Group | Boltzmann | | | Logistic | | | Gompertz | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ΔAIC | ER | $\chi_\nu^2$ | ΔAIC | ER | $\chi_\nu^2$ | ΔAIC | ER | $\chi_\nu^2$ |
| colliding cars | I1 | 2.1 | 3 | 1.072 | 0 | | 0.851 | 8.6 | 74 | 2.213 |
| | I2 | ~0 | 1 | 1.488 | 0 | | 1.488 | 7.4 | 39 | 3.361 |
| collision at an angle | I1 | 0 | | 0.142 | 7.1 | 35 | 0.313 | 23.2 | >1000 | 1.866 |
| | I2 | 1.4 | 2 | 0.522 | 0 | | 0.449 | 9.8 | 137 | 1.342 |
| *mass-spring system* | I1 | 0 | | 0.858 | 3.0 | 5 | 1.199 | 13.5 | 860 | 3.853 |
| | I2 | 3.0 | 5 | 1.416 | 0 | | 1.014 | 5.2 | 14 | 1.812 |
| *shooting a block* | I1 | 0.3 | 1 | 0.889 | 0 | | 0.858 | 8.2 | 61 | 2.142 |
| | I2 | 9.8 | 136 | 3.884 | 6.0 | 20 | 2.549 | 0 | | 1.304 |
| person on a ladder | I1 | 0 | | 1.184 | 3.0 | 4 | 1.652 | 10.1 | 156 | 3.637 |
| | I2 | 0 | | 0.459 | 3.7 | 6 | 0.689 | 11.1 | 262 | 1.582 |
| calculating torques | I1 | 0 | | 0.383 | 0.8 | 2 | 0.418 | 20.6 | >1000 | 3.788 |
| | I2 | 0 | | 1.020 | 5.2 | 13 | 1.807 | 15.9 | >1000 | 5.946 |
| *asteroid impact* | I1 | 0.7 | 1 | 1.370 | 0 | | 1.264 | 5.3 | 14 | 2.264 |
| | I2 | ~0 | 1 | 2.755 | 0 | | 2.752 | 3.6 | 6 | 4.109 |
| *post-collision orbit* | I1 | 2.1 | 3 | 2.539 | 0 | | 2.015 | 0.6 | 1 | 2.154 |
| | I2 | 0 | | 1.501 | 0.85 | 2 | 1.648 | 7.8 | 50 | 3.586 |
| flywheel kinematics | I1 | 1.8 | 2 | 0.837 | 0 | | 0.688 | 10.9 | 232 | 2.307 |
| | I2 | 1.4 | 2 | 1.335 | 0 | | 1.148 | 6.3 | 24 | 2.315 |
| angular motion | I1 | 1.3 | 2 | 1.084 | 0 | | 0.937 | 7.7 | 46 | 2.199 |
| | I2 | 0.8 | 2 | 0.631 | 0 | | 0.580 | 6.9 | 33 | 1.258 |
| *parallel-axis theorem-I* | I1 | 0 | | 1.939 | 1.7 | 2 | 2.335 | 8.5 | 69 | 4.969 |
| | I2 | 0 | | 4.404 | 2.3 | 3 | 5.686 | 7.7 | 48 | 10.404 |
| *parallel-axis theorem-II* | I1 | 1.5 | 2 | 1.398 | 0 | | 1.181 | 3.9 | 7 | 1.831 |
| | I2 | 2.0 | 3 | 2.230 | 0 | | 1.778 | ~0 | 1 | 1.780 |
| Median | | 0.5 | | 1.259 | 0 | | 1.190 | 7.8 | | 2.238 |

Differential model fits with independent variable Ln(t). The ΔAIC values and evidence ratios (ER) are given relative to the best fit model. The median ΔAIC and reduced chi-squared ($\chi_\nu^2$) values are shown at the bottom. Related problem pairs are shown by normal letter pairs and *italicized* pairs alternatively.

# Figures

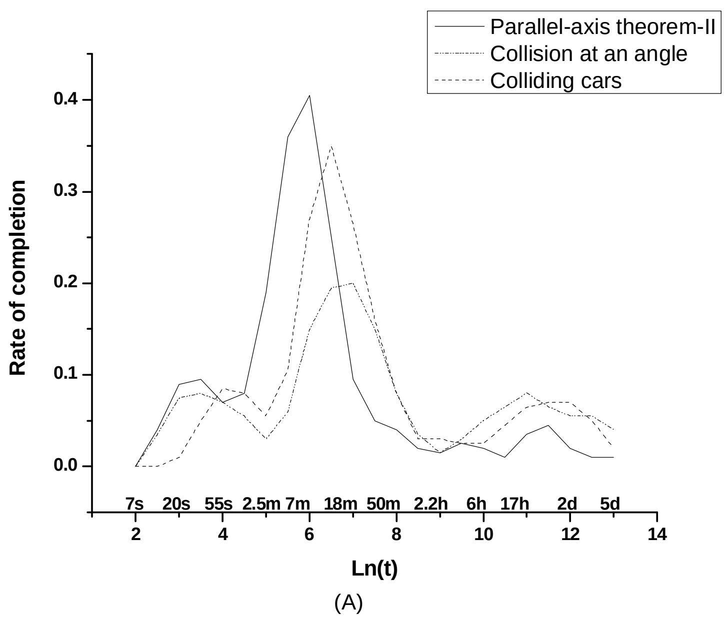



Fig. 1. The Rate of completion for three representative problems. The plots reveal three distinct groups. The three group structure is preserved although the problems have different median times as measured by the central peak.

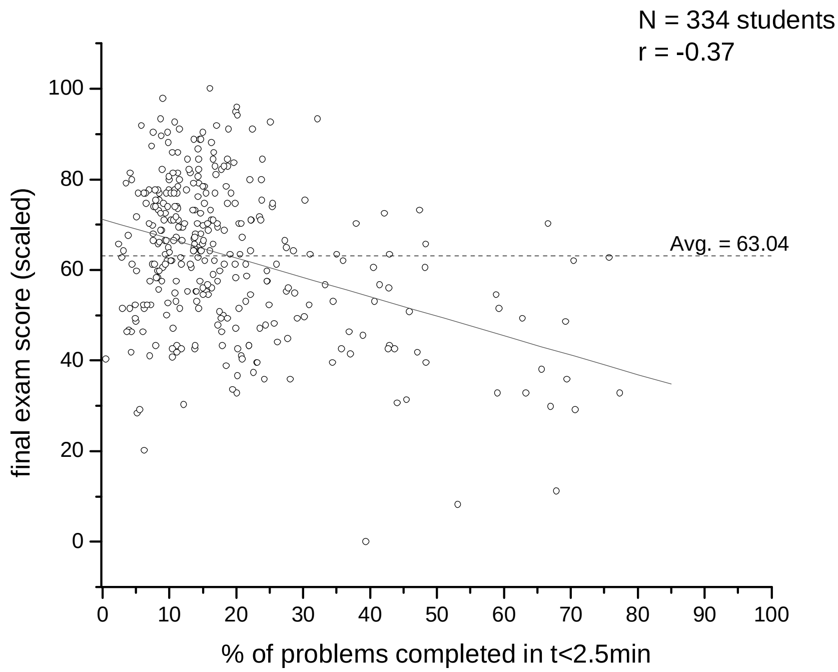
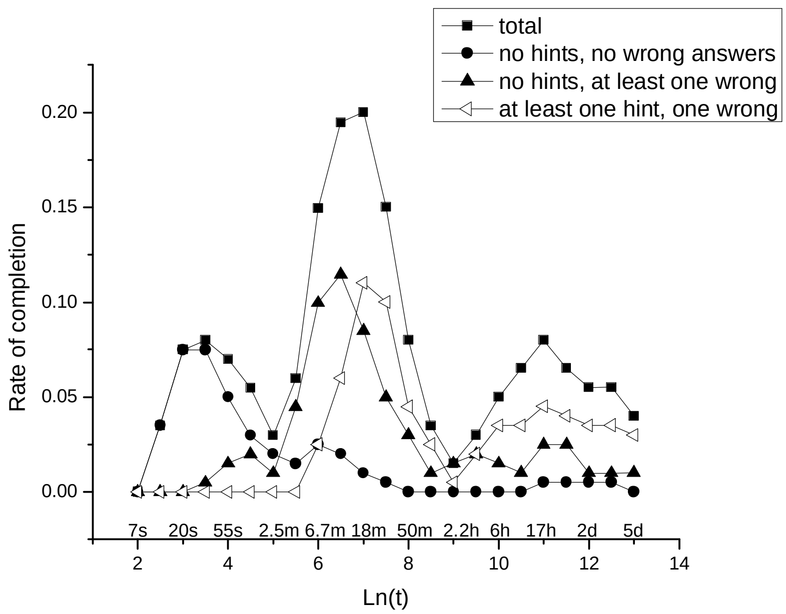
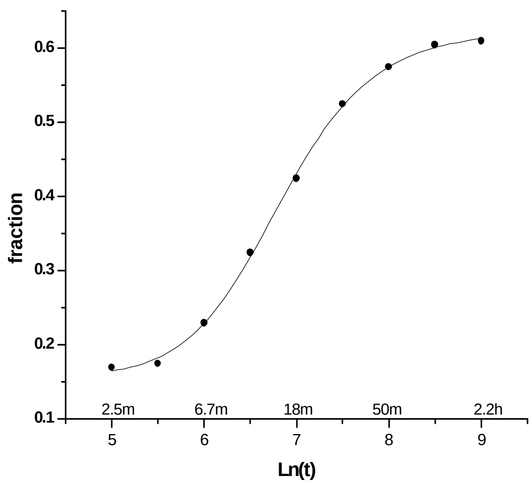



Fig. 2. The more problems that a student completes in less than 2.5 minutes, the lowers is his/her final exam score, suggesting academic dishonesty.




(B)

Fig. 3. A breakdown of the rate of completion for a representative problem (B). The “quick responders” (typically< 2.5 minutes) are dominated by students who do not ask for hints or give incorrect answers. The “real-time solvers” (typically, 2.5 minutes to 2.2 hours) and “delayed solvers” (typically>2.2 hours) are dominated by those who make mistakes and ask for help (hints).



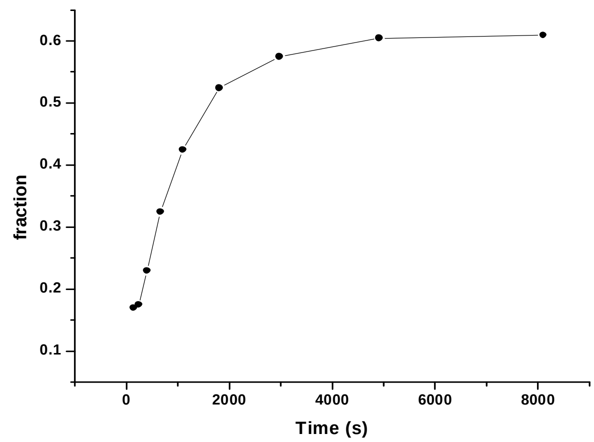




Fig. 4. Time to completion in log time (left) and linear time (right) for the real-time solvers. The prominent ‘S’ or sigmoidal shape occurs in logarithmic time.

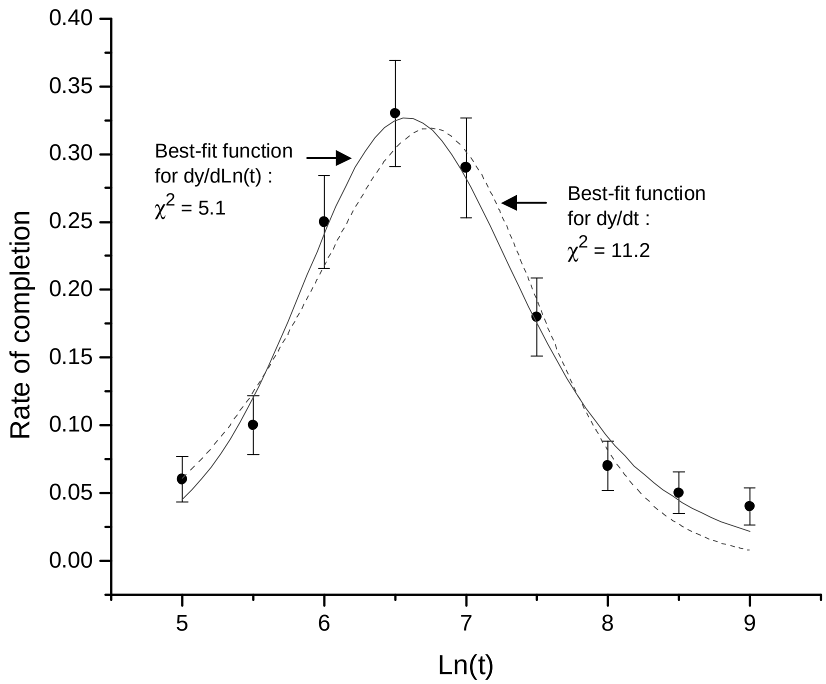



Fig. 5. An example of a case where the function that best-fit dy/dt in t (where y = fraction complete) does not fit well for dy/dLn(t) in Ln(t). This shows that y vs Ln(t) is better described by a function different from that of y vs t.

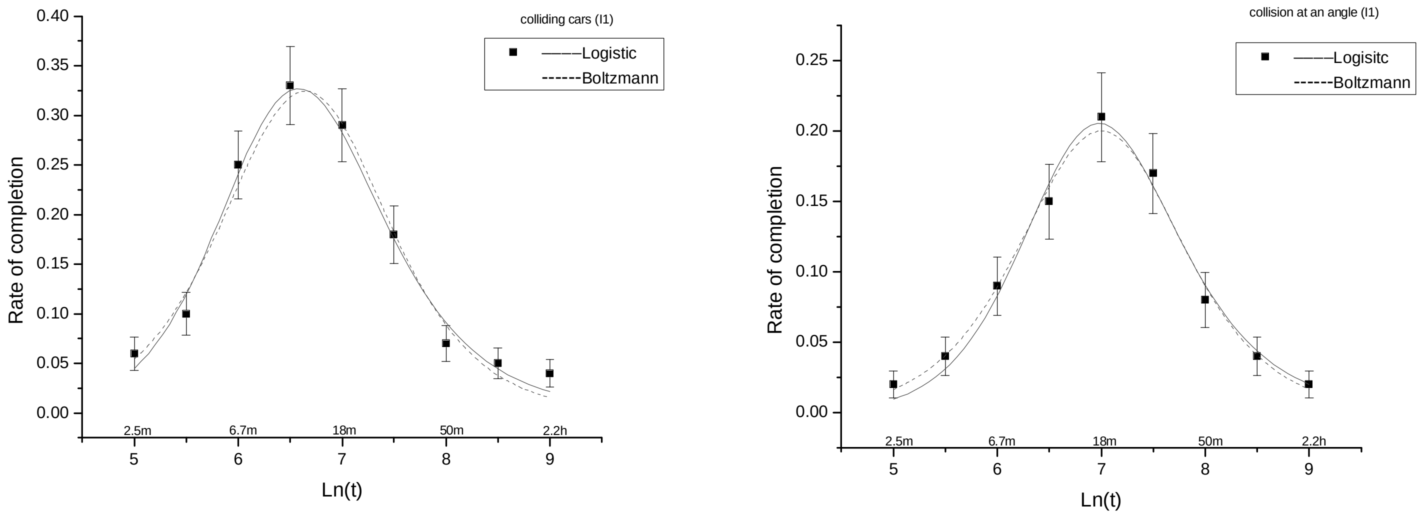



Fig. 6. Two cases where the Logistic (left) and the Boltzmann (right) fit well in real time solver region. In the left figure the Boltzmann does not account well for the right-most points while in the right figure the Logistic does not account well for the left-most points. The points in between are more or less well accounted for by both models in both figures. Thus, the preference of one model over the other seems to be influenced by the end-points which are sensitive to the existence of students from the quick responder and the delayed solver regions.

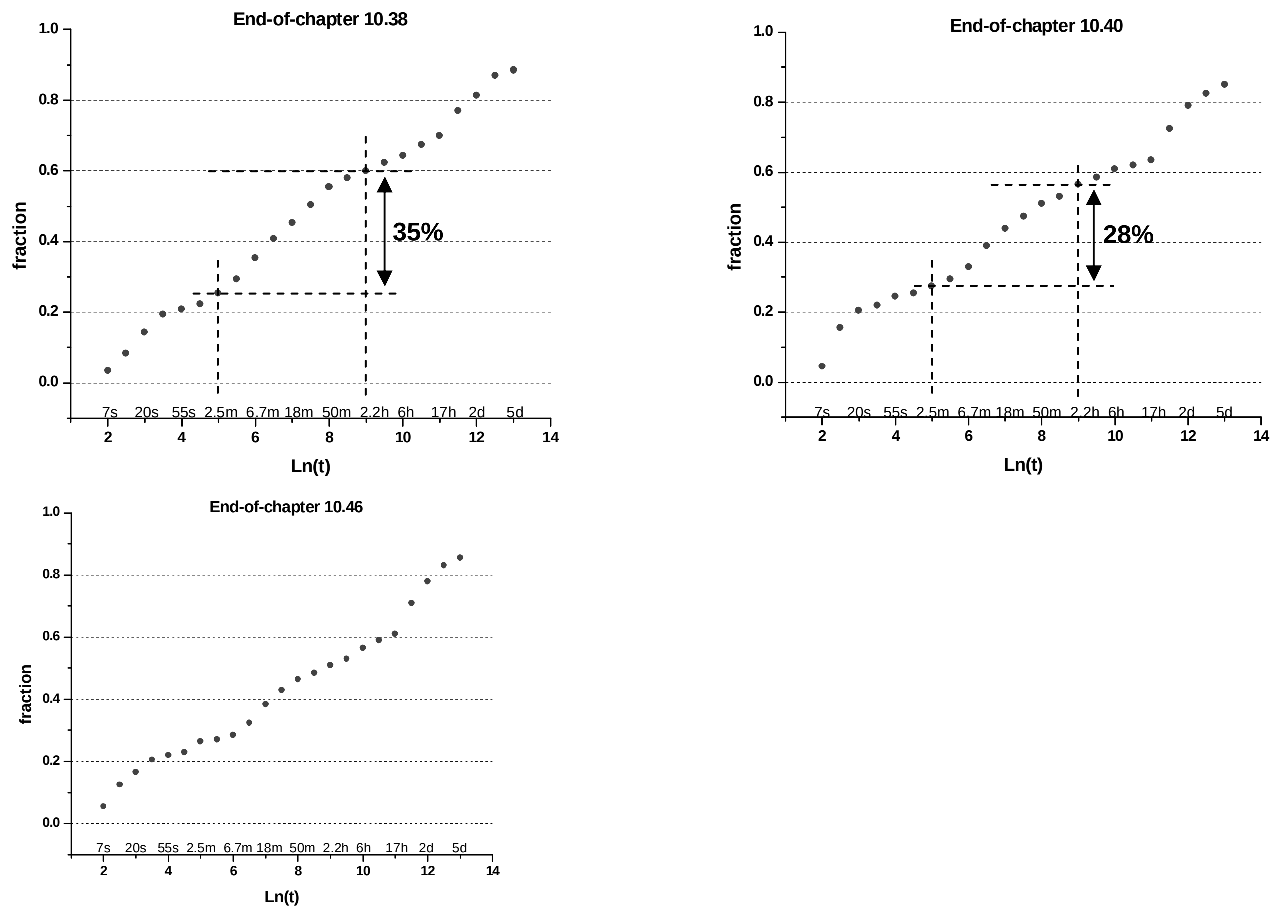



Fig. 7. Three representative End-of-Chapter (EOC) problems. The prominent increase in fraction of students able to solve the problem within 2.5 minutes – 2.2 hours is absent in contrast to tutorial problems (see Fig. 8).

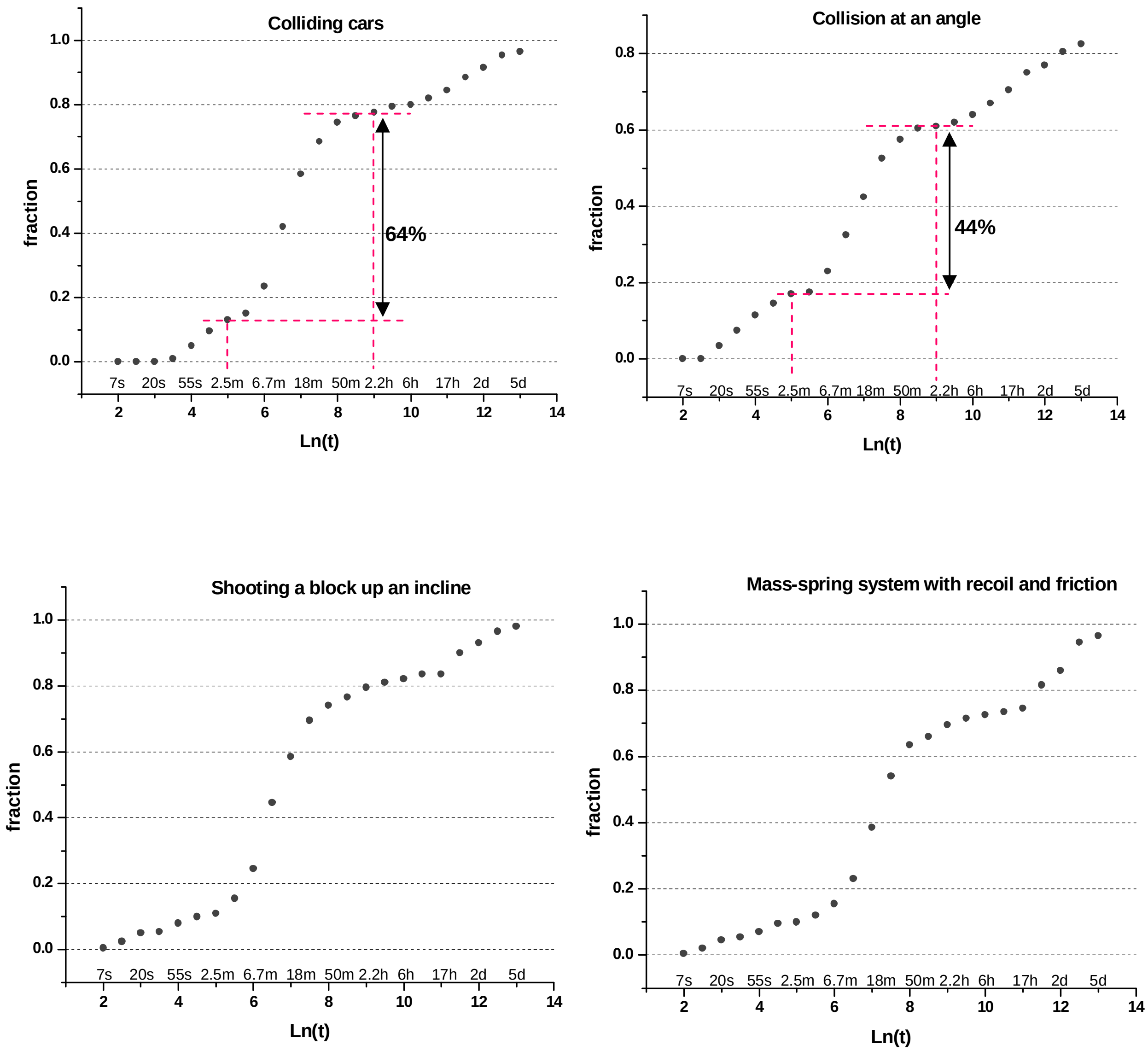



Fig. 8. Four representative tutorial problems. The prominent increase in fraction of students able to solve the problem within 2.5 minutes – 2.2 hours can be seen in contrast to EOC problems (see Fig. 7).