The Effect of Prerequisite Skill Response Time on Current Skill Correctness

Douglas Selent  Supreeth Gowda  Neil T. Heffernan III  Yutao Wang

Worcester Polytechnic Institute
100 Institute Rd Worcester, MA 01609

dselent@wpi.edu  smgowda@wpi.edu  nth@wpi.edu  yutaowang@wpi.edu

ABSTRACT
In this paper the effects of prerequisite skill average response time was examined on both the average correctness of the prerequisite skills and the average correctness of the current skill. The intent is to find a way to help students who are struggling to learn a current skill by having the student practice a skill they know to quicken response times. This will hopefully make it easier for the student to learn the current skill. Our hypothesis is that students who answer prerequisite skills faster will have an easier time learning the current skill. We used a subset of data from the ASSISTments online tutoring system collected from August 2010 to December 2012. Based on the ASSISTments skill graph we chose 5 random skills with enough data on both the current and prerequisite skills. The results of our analysis show that prerequisite response time affects the current skill’s correctness for “easy” skills. Faster prerequisite response times outperformed slower prerequisite response times by up to 30% for prerequisite and up to 10% for current skills. For “hard” skills, fast prerequisite response times had a higher percentage correct for the prerequisite skill, but showed no difference in the percent correct for the current skill. The R^2 values for the relationship between prerequisite response time of “easy” skills and median/mean correctness of the current skill were .74/.95, .5/.93, and .88/.6. From these results we can conclude that faster response times on prerequisite skills lead to higher percentages correct on “easy” current skills but have no effect on “hard” current skills.

Keywords
prerequisite skill, response time, web tutoring, skill graph

1. INTRODUCTION
1.1 Inspiration
The idea for this paper came from the first author’s personal experience at middle school as a PIMSE fellow. There was a student who was unable to learn how to solve a certain problem. No matter what the author tried, the student was still “not getting it”. The connection just could not be made. The author and the teacher both talked about teaching techniques, but both were unable to find a good way to teach this student the skill. As suggested in [1], students who answer questions on prerequisite skills faster might have an easier time learning the current skill. This paper looks to see if there is a relationship between the times it takes students to answer problems on prerequisite skills and how well the students do on the current skill. The intent is that if quicker response times on prerequisite skills result in higher percentages correct on current skills then students who have trouble learning a current skill can practice a prerequisite skill they already know. This will make learning the current skill easier for them since they can practice on a skill that they know (perhaps not too well) to prepare them to learn a skill that they do not know. Response times have been shown to be a factor in predicting student performance shown in [4] and on some data sets in [5]. Response time for hints has also been analyzed in [7].

1.2 ASSISTments System
Our data is collected from the online intelligent tutoring system called ASSISTments founded by Neil Heffernan. ASSISTments is used by mostly math students between grades 4-12 in the United States with a large number of students located in or near Massachusetts. Students use ASSISTments on classwork and homework, which may be done with or without the use of a paper copy. Instant feedback is typically provided to the student upon answering the problems, so they know immediately whether or not they have answered correctly. Students are not allowed to skip any problems and must answer correctly to continue. A response is correct only if it was both correct and the first response. If the student answers correctly after answering incorrectly, that student is recorded as incorrect for that problem. The number of tried it takes the student to answer correctly is also recorded, as well as if the student has received any hints on the problem.

2. Data
Our initial data consisted of over 60 GB of all ASSISTments problem level data on mathematics from August 2010 to December 2012 for students between grades 4-12. From this data we decided only to use a small subset that we thought would be reliable. We programatically removed all rows where the problem type was a non-skill builder. The reason for this is because non-skill builder problem sets are often given to students both on paper and on ASSISTments. The problem with this is that several students tend to do everything on paper first and then enter their answers on ASSISTments. This would lead to very fast response times. Since our goal is to see the effect of prerequisite skill response time on current skill correctness, we need reliable time data. The Assistment skill builder problem type is where students are given random problems from a pool of 50-200 related problems. The student completes the skill builder
when he or she has answered three problems correctly in a row. The student is also only allowed to answer 10 problems in a single day. We believe this problem set type gives more reliable time data. Due to the number of possible problems randomly chosen it is highly unlikely the student will do these problems on paper first. It would be hard to search through 50-200 problems in random order on a piece of paper. Teachers are allowed to create their own skill builders, however must use the pre-created ones because it saves them time. From Doug’s experience visiting middle school once a week, we can also back up this statement since he has never seen students do skill builder problem sets on paper.

Problems in ASSISTments can be set in two modes, “tutor mode” and “test mode”. Tutor mode is when the student is informed whether or not they have answered correctly after each problem. Test mode is when the student is not given any feedback on their answer. To have more consistent time data we chose to only use problems that were in tutor mode. We removed all data instances where the problem was in test mode.

We also know that the ASSISTments system has a large number of "fake" classes. This is because all teachers are encouraged to create "test" or "sandbox" classes to test their material. There are also several test classes that developers use to do testing. Our approach to dealing with this was to first programmatically remove all data where the class that the student belonged to had less than five students. Since there are so many classes it was not reasonable to look through each one. After removing those classes we ended up with roughly 2600 classes. From this list we manually scanned through all the class names and removed any classes we thought were not real classes, such as “trash”, or “junk”. The result was 2473 classes that we believed were reliable. All data that did not belong to those classes was removed from our data set.

Since our analysis deals with skills and prerequisite skills, all problem data that was not associated with a skill was of no use to us. We removed all data where there was no skill associated with the problem. We also removed all data where the problem was associated with more than one skill. The reason for this is because we did not want to deal with how much each skill affects the problem correctness. This would introduce the credit/blame problem of why a student answered incorrectly, which has been studied in [8] and [3]. Having a one-to-one relationship between problems and skills makes our analysis more clear.

Another issue we encountered was that over 500,000 rows of data had bad times for how long it took a student to complete a problem. We considered bad times as negative times, null times, or times less than one second. We found that a bug in the system accounted for some of the null times. We tried looking into the issue but only found one problem that did not account for all the bad times. One reason this occurs is a result of when a teacher deletes a student’s progress on a problem set. However this occurrence did not account for all the bad times. We deemed the times unreliable and removed those data instances. We also removed data instances with times greater than one hour. We believe those times are unreliable because it is unlikely a student is really working on a single problem for over an hour. It is more likely the student left the computer and came back later.

The result of our data cleaning process left us with 32,196 rows of data, of which 23,695 had data on both prerequisite and current skill data per student. This included 19 different skills and 11,481 distinct students. Our final dataset is online and can be found at [6]. The column that represents the average student prerequisite response time is called “Average Prereq User+Skill”. We note that there are values of -1 in our dataset, which mean that the value could not be calculated. If a student did not have any data on a prerequisite skill then the average student prerequisite response time will be shown as -1. These values were not included in our analysis or calculations.

3. Skill Graph

The terms prerequisite skill and current skill refer to the idea of a skill graph. A skill graph is basically a graph of all the skills where a prerequisite skill is a skill that must be learned before learning the current skill. An example is that “addition of whole numbers” is a prerequisite skill to “subtraction of whole numbers”. A slightly out dated version of the skill graph used in the ASSISTments online tutoring system can be found at [2]. Since the ASSISTments skill graph is constantly changing there is currently no published up to date version. Only the database itself contains the exact skill graph we used, but it is very similar to the one online. A piece of the skill graph is shown in figure 1. The grade number at the top shows what grade the skill is typically leaned at. The colors represent the type of skill, of which only one was included in the figure. Most importantly are the directed edges on the graph. These edges represent the links between prerequisite skills and current skills. Only direct prerequisite links are included. This means that although a student would need to learn addition before calculus, there is no direct link between the two skills. It is possible for a single skill to have multiple prerequisite skills. It is also possible for a prerequisite skill to have multiple current skills. In our analysis all of our skills have only a single prerequisite skill.

Figure 1. Skill Graph Piece

A piece of the ASSISTments skill graph is shown above. The complete skill graph can be found at http://teacherwiki.assistment.org/wiki/images/2/29/Assistment_Skill_Diagram_Poster_PDF.pdf

4. Analysis

For our analysis we converted our problem level data into skill level data. We calculated a skill difficulty for each skill. We took the number of problems answered correctly associated with a skill and divided that number by the total number of problems for that skill. The average prerequisite time per student per skill was also calculated. For each student and each skill, the time taken on all the problems was added up and then divided by the number of problems. Similarly the percent correct per student per skill was calculated for each student and each skill.
We did not consider the ordering of prerequisite skills and current skills. Therefore our data may contain some instances where the student has done the current skill before the prerequisite skill. Something like this would be common in middle, where we know that students in later grades such as 8th grade continue to practice the same skills that they did in 7th grade. ASSISTment skill builders are often done as review.

We also did not include correctness in reference to average response time. The trends we noticed were that the average and median response time for both correct and incorrect answers increased as the average response time increased. This makes sense since the average response time is a combination of the average response times for both correct and incorrect answers. The time it took a student to complete a problem with incorrect answers was longer than the time it took a student to complete a problem with correct answers. We feel that just the average response time was a good enough time metric since factoring in the correctness did not seem to add any useful information.

4.1 Binning Time

We decided to bin response time into four different bins for our analysis. Binning of student response time was done in [1]. However we encountered problems when we tried binning time by increments of time. We ended up with 95% of the time data in one bin and the other bins nearly empty. This is because most of the time data is grouped together with the exception of huge outliers. Instead we decided to bin time into equal quarters. For each skill we sorted our file by average prerequisite skill time per student. We broke the files for each skill up into quarters, where the first quarter contained the fastest times and the last quarter contained the slowest times. Binning time made our analysis less susceptible to outliers due to the large range of times. By breaking the files up into equal fourths it also allowed us to compare times across different skills despite the fact that different skills require different amount of time to complete. We could compare times based on what bin they were in. We chose four bins because it allowed us to look at as many different time levels as we could while still having enough data instances for each bin.

Table 2. Skill Information

<table>
<thead>
<tr>
<th>Skill ID</th>
<th>Skill Name</th>
<th>Grade taught</th>
<th>Number of Data Instances</th>
<th>Difficulty (% correct)</th>
<th>Prerequisite Skill Name</th>
<th>Prerequisite Difficulty (% correct)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Probability of Two Distinct Events</td>
<td>7.9</td>
<td>1151</td>
<td>48</td>
<td>Probability of a Single Event</td>
<td>71</td>
</tr>
<tr>
<td>21</td>
<td>Interior Angles Figures with More than 3 Sides</td>
<td>7.3</td>
<td>1139</td>
<td>53</td>
<td>Interior Angles Triangle</td>
<td>73</td>
</tr>
<tr>
<td>75</td>
<td>Square Root</td>
<td>8.4</td>
<td>2552</td>
<td>82</td>
<td>Order of Operations +,-,/,,* ()</td>
<td>81</td>
</tr>
<tr>
<td>95</td>
<td>Substitution</td>
<td>7.8</td>
<td>1600</td>
<td>72</td>
<td>Exponents</td>
<td>67</td>
</tr>
<tr>
<td>298</td>
<td>Area Triangle</td>
<td>5</td>
<td>2330</td>
<td>70</td>
<td>Area Rectangle</td>
<td>79</td>
</tr>
</tbody>
</table>

Figure 2. Example Problems for Each Skill

**Skill 17 Example**

Luis is going to toss two coins. What is the probability that he will toss one head and one tail?

**Skill 21 Example**

What is the measure of angle A?

**Skill 75 Example**

The square root of 48 is between which two whole numbers?

**Skill 95 Example**

Evaluate this expression if m = 5, n = -8, and p = -3. p - m + n

**Skill 298 Example**

What is the area of the triangle shown below?
4.2 Skill Information
We looked at five different skills to see how the average prerequisite response time affected both the correctness of the prerequisite skill and the correctness of the current skill. Brief descriptions of the skills are shown in Table 2. An example problem of each skill type is shown in Figure 2.

Table 2 shows each of the five skills we used in our analysis. The “grade taught” column in Table 2 is what school grade the skill is usually taught at. For example skill 95, “substitution” is usually taught 8/10 of the way through 7th grade. Therefore the grade taught column for skill 95 is 7.8. The number of data instances represents how many data rows we had on each skill. Since there were four time bins each bin will have the total number of data instances divided by the number of bins (4).

Both the current skill difficulty and the prerequisite skill difficulty were rounded to the nearest percent. A skill with higher difficulty such as skill 17 has a lower percent correct (48%). We use the terms “easy” and “hard” skills comparatively among the five skills we looked at. Skill 17 is hard because it has a much lower percent correct than other skills.

5. Results
For all five skills we looked at the mean and median percent correct for both prerequisite and current skills at each bin level. The results are shown in Table 3. We chose to look at both the mean and the median since both have weaknesses and strengths. Means can get skewed by outliers and medians can get skewed by several values that are the same. It is common for over half the students to get 100% correct on a given skill. It is also common for several students to get extremely low percentages correct on a given skill. Using both means and median trends give a better reflection of the data. All time values are rounded to the nearest second and all percent values are rounded to the nearest whole percent. We also provide graphs figures 3-6 showing all skills by bin level for both mean current skill correctness and median current skill correctness.

Table 3. Prerequisite Time vs. Prerequisite and Current Percent correct

<table>
<thead>
<tr>
<th>Skill ID</th>
<th>Time (Seconds)</th>
<th>Bin #</th>
<th>Median % Correct Pre</th>
<th>Mean % Correct Pre</th>
<th>Median % Correct Current</th>
<th>Mean % Correct Current</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>1 - 26</td>
<td>1</td>
<td>89</td>
<td>87</td>
<td>58</td>
<td>59</td>
</tr>
<tr>
<td>17</td>
<td>26 - 39</td>
<td>2</td>
<td>77</td>
<td>78</td>
<td>57</td>
<td>58</td>
</tr>
<tr>
<td>17</td>
<td>39 - 64</td>
<td>3</td>
<td>76</td>
<td>76</td>
<td>56</td>
<td>57</td>
</tr>
<tr>
<td>17</td>
<td>64 - 1665</td>
<td>4</td>
<td>66</td>
<td>76</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>21</td>
<td>1 - 25</td>
<td>1</td>
<td>100</td>
<td>88</td>
<td>70</td>
<td>69</td>
</tr>
<tr>
<td>21</td>
<td>25 - 36</td>
<td>2</td>
<td>100</td>
<td>85</td>
<td>71</td>
<td>69</td>
</tr>
<tr>
<td>21</td>
<td>36 - 56</td>
<td>3</td>
<td>100</td>
<td>85</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>21</td>
<td>56 - 700</td>
<td>4</td>
<td>83</td>
<td>73</td>
<td>70</td>
<td>68</td>
</tr>
<tr>
<td>75</td>
<td>1 - 14</td>
<td>1</td>
<td>100</td>
<td>96</td>
<td>100</td>
<td>89</td>
</tr>
<tr>
<td>75</td>
<td>14 - 22</td>
<td>2</td>
<td>100</td>
<td>90</td>
<td>100</td>
<td>86</td>
</tr>
<tr>
<td>75</td>
<td>22 - 38</td>
<td>3</td>
<td>88</td>
<td>85</td>
<td>100</td>
<td>86</td>
</tr>
<tr>
<td>75</td>
<td>38 - 156</td>
<td>4</td>
<td>78</td>
<td>77</td>
<td>93</td>
<td>84</td>
</tr>
<tr>
<td>95</td>
<td>1 - 19</td>
<td>1</td>
<td>100</td>
<td>93</td>
<td>90</td>
<td>85</td>
</tr>
<tr>
<td>95</td>
<td>19 - 34</td>
<td>2</td>
<td>83</td>
<td>82</td>
<td>100</td>
<td>84</td>
</tr>
<tr>
<td>95</td>
<td>34 - 98</td>
<td>3</td>
<td>76</td>
<td>75</td>
<td>86</td>
<td>80</td>
</tr>
<tr>
<td>95</td>
<td>98 - 643</td>
<td>4</td>
<td>71</td>
<td>67</td>
<td>78</td>
<td>75</td>
</tr>
<tr>
<td>298</td>
<td>1 - 14</td>
<td>1</td>
<td>100</td>
<td>94</td>
<td>89</td>
<td>84</td>
</tr>
<tr>
<td>298</td>
<td>14 - 22</td>
<td>2</td>
<td>100</td>
<td>89</td>
<td>83</td>
<td>80</td>
</tr>
<tr>
<td>298</td>
<td>22 - 57</td>
<td>3</td>
<td>80</td>
<td>77</td>
<td>78</td>
<td>75</td>
</tr>
<tr>
<td>298</td>
<td>57 - 955</td>
<td>4</td>
<td>83</td>
<td>80</td>
<td>80</td>
<td>74</td>
</tr>
</tbody>
</table>

The time range and bin number are shown for each skill. Bin 1 has the fastest times and bin 4 has the slowest times. All times are rounded to the nearest second and all percentages are rounded to the nearest whole percent.
Skills 17 and 21 show a large difference in prerequisite skill correctness between the fastest and slowest prerequisite time bin. Skill 17 drops from a median prerequisite percent correct of 89% to 66% and from a mean prerequisite percent correct of 87% to 76%. Skill 21 drops from a median prerequisite percent correct of 100% to 83% and from a mean prerequisite percent correct of 88% to 73%. This trend does not carry over to the current skill correctness. Skills 17 and 21 show almost no difference in median and mean current skill percent correct between the fastest and the slowest prerequisite time bin. Skills 75, 95, and 298 show a similar trend with prerequisite percent correctness. All three skills show that as prerequisite response time decreases so does prerequisite percent correct. Unlike skills 17 and 21, these trends carry over to the current skill correctness. Skills 75, 95, and 298 show a drop of 7%, 12%, and 9% respectively between the fastest and slowest prerequisite time bins for median percent correctness on the current skill. The skills show similar drops of 5%, 10%, and 10% respectively for mean current percent correct. Since the drop in percentage says nothing about the trend we include $R^2$ values for our trend lines. The $R^2$ values for the relationship between prerequisite response time median/mean correctness of the current skill for skills 17, 21, 75, 95, 298 were .14/.04, .09/.4, 88/.6, .74/.95, .5/.93, and 88/.6 respectively.

For the skills 75, 95, and 298 there are strong correlations between current percent correct and prerequisite skill response time.

One odd observation is that the median percent correct for the current skill in figure 6 goes up from 90% to 100% in bin 2 for skill 95. The mean current percent correct for skill 95 in figure 5 still shows percent correct going down as bin number increases. This can easily be seen on the graphs. After looking at the data more closely we saw that in bin 1 there were less students with 100% correct, but more students within the 80-90% range. Bin 2 had more students with 100% correct, but much fewer students in the range from 80-90%.

We believe the reason why some skills were affected by prerequisite response time and other skills were not is due to the skill difficulty. For the “hard” skills (17 and 21) the probability that a student will get problems with those skills correct is 48% and 53%. For the “easy” skills (75, 95, and 298), the probability that a student will get problems with those skills correct is 82%, 72%, and 70%.

Figure 3. Time Bin # vs. Mean Prerequisite Skill Percent Correct

Figure 4. Time Bin # vs. Median Prerequisite Skill Percent Correct
6. Conclusions + Future Work

Our results show that prerequisite skill response times are strongly correlated with a large increase on prerequisite skill performance. More importantly our results also show a strong correlation with a large (but not as large) increase on current skill performance for easy skills. However, prerequisite skill response time does not seem to make a difference for hard skills. The prerequisite skill response time effect does not seem to carry over to the current skill for hard current skills. One possible explanation is that if the prerequisite skill itself is hard then students tend to suffer in both prerequisite skill as well as current skill.

For students who are struggling to learn an easy skill, it might be helpful to have them practice the prerequisite skill to quicken the response time. This may solidify the student’s knowledge on the prerequisite skill making it easier for the student to learn the current skill. This will make it easier to teach the student and less painful for a student to learn a new skill.

One factor not considered was the time between the prerequisite skill and the current skill. We think that the time between learning the prerequisite skill and the current skill might also be an important factor on how well a student is able to learn the current skill. Possible future work can be to examine the time in between learning the two skills and see if that makes a difference on how well the student does on the current skill.

7. Acknowledgements

We would like to acknowledge funding from the U.S. Department of Education’s GAANN grant and by the National Science foundation grant “Graduates in K-12 Education” (GK-12) Fellowship.

8. REFERENCES


