Andrew Sullivan

9/26/13

CMSI 370

Assignment 0926

For our project we were testing three pieces of technology that accomplish the domain of processing arithmetic. The three tools that we tested were three different calculators. The first was using the Google search bar to calculate arithmetic. The second was the calculator that is built into MacBooks. The third and final tool we tested was a standard TI-83 calculator. On each device we had each test subject perform three tasks, while recorded data.

The three tasks that we asked the test subjects to perform were to solve three arithmetic problems, three trigonometric problems, and finally to solve three problems which consisted of several arithmetic calculations.

Arithmetic:

1. 5 + 3
2. 187,652 × 27
3. 5,783 ÷ 12

Trigonometric:

1. Sin(45**°)**
2. 2cos(52**°)**
3. Tan(134**°)**

Several:

1. 3(7+2))⁄15
2. (3 ⁄ 2)(8 − 2)
3. (7 (64 + 128)) ⁄ 2.7

These were our three measurement activities. The three usability metrics that we chose to evaluate this experiment were efficiency, errors, and finally satisfaction. For efficiency we timed how long it took the test subjects to complete each of the three tasks. For errors we simply recorded how many, if any, errors the subjects had for each of the tasks. For satisfaction we had the test subjects rate on a scale from 1 to 10 how much they enjoyed performing the tasks on each device.

The results can be seen at Figure 1 below. Even though we only tested six subjects, the results were quite clear which devices performed the best in the hands of our users. Figure 2 shows the average time of completing a problem on each device, the average rating of satisfaction, and the total amount of errors for each device. Both the Google search bar and the MacBook built-in calculator scored similarly on their average time. It took the test subjects an average of 41.89 seconds to complete a problem. It took the subjects a slightly longer average time of 49.94 seconds to complete a problem, while it only took an average of 16.33 seconds to complete a problem using the TI-83 calculator. These results clearly indicate that the TI-83 calculator has a much higher efficiency than the other two devices, and that the Google calculator has a slightly higher efficiency than the MacBook calculator.

The other two results were even more revealing. For the total amount of errors that occurred during each device across all users Google and the TI-83 scored 2 and 1, respectively. The MacBook calculator on the other hand collected a total of 19 errors, a significant increase from the other two devices.

Finally we take a look at the satisfaction rating that each test subject gave us after using each device. The TI-83 calculator scored the highest with an average rating of 9.5 out of 10. The Google calculator scored an average of 7.58, and the MacBook calculator scored a mere 3.3 overall. This test clearly showed that the satisfaction rating of the TI-83 was very high and that the MacBook calculator was extremely low.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **GOOGLE BUILT IN CALCULATOR** | | | |  |  |  |  |
|  | Efficiency |  |  | Errors |  |  | Satisfaction |
| Subject | Arithmetic | Trigonometry | Several | Arithmetic | Trigonometric | Several | Google |
| 1 | 57 | 32 | 61 | 0 | 0 | 0 | 9.5 |
| 2 | 44 | 31 | 71 | 0 | 0 | 0 | 5 |
| 3 | 24 | 35 | 55 | 0 | 0 | 0 | 8 |
| 4 | 32 | 31 | 50 | 0 | 1 | 0 | 8 |
| 5 | 30 | 42 | 44 | 0 | 1 | 0 | 8 |
| 6 | 31 | 37 | 47 | 0 | 0 | 0 | 7 |
| **MACBOOK BUILT IN CALCULATOR** | | | |  |  |  |  |
|  | Efficiency |  |  | Errors |  |  | Satisfaction |
| Subject | Arithmetic | Trigonometry | Several | Arithmetic | Trigonometric | Several | Macbook |
| 1 | 46 | 55 | 115 | 0 | 2 | 2 | 4 |
| 2 | 43 | 35 | 58 | 0 | 3 | 3 | 1 |
| 3 | 38 | 37 | 75 | 0 | 1 | 0 | 3 |
| 4 | 38 | 35 | 72 | 0 | 0 | 0 | 5 |
| 5 | 30 | 24 | 67 | 1 | 1 | 3 | 3 |
| 6 | 37 | 31 | 63 | 0 | 1 | 2 | 4 |
| **TI GRAPHING CALCULATOR** | | |  |  |  |  |  |
|  | Efficiency |  |  | Errors |  |  | Satisfaction |
| Subject | Arithmetic | Trigonometry | Several | Arithmetic | Trigonometric | Several | Ti Graphing |
| 1 | 15 | 13 | 31 | 0 | 0 | 0 | 10 |
| 2 | 18 | 16 | 16 | 0 | 0 | 0 | 10 |
| 3 | 15 | 17 | 17 | 0 | 0 | 0 | 9 |
| 4 | 15 | 20 | 20 | 0 | 0 | 0 | 10 |
| 5 | 13 | 12 | 12 | 0 | 1 | 0 | 9 |
| 6 | 14 | 15 | 15 | 0 | 0 | 0 | 9 |

***Figure 1***

|  |  |  |  |
| --- | --- | --- | --- |
|  | Average Efficiency | Total Errors | Average Satisfaction |
| Google | 41.89 | 2 | 7.58 |
| MacBook | 49.94 | 19 | 3.3 |
| Ti Graphing | 16.33 | 1 | 9.5 |

***Figure 2***

Taking all of the results into consideration the MacBook calculator was designed the worst of the three. The TI-83 has the best design to it, and the Google calculator is somewhere in the middle. Every usability metric that we tested supported these conclusions. There are many possible reasons as to why these devices performed the way that they did, and why they all seemed to perform the same way for each user. First of all, the TI-83 calculator is very common amongst college students, and therefore many people that we tested have had extensive use with the device. The combination of the TI-83s memorability and the user’s familiarity with the device most likely had an influence on the results. Most of the test subjects were either not aware of the existence of the built in MacBook calculator or rarely use it. Most of them have used Google before but not extensively. These match up with our results, implying that familiarity with and memorability of a certain device can have drastic influences over how efficient a device can be to a particular user.

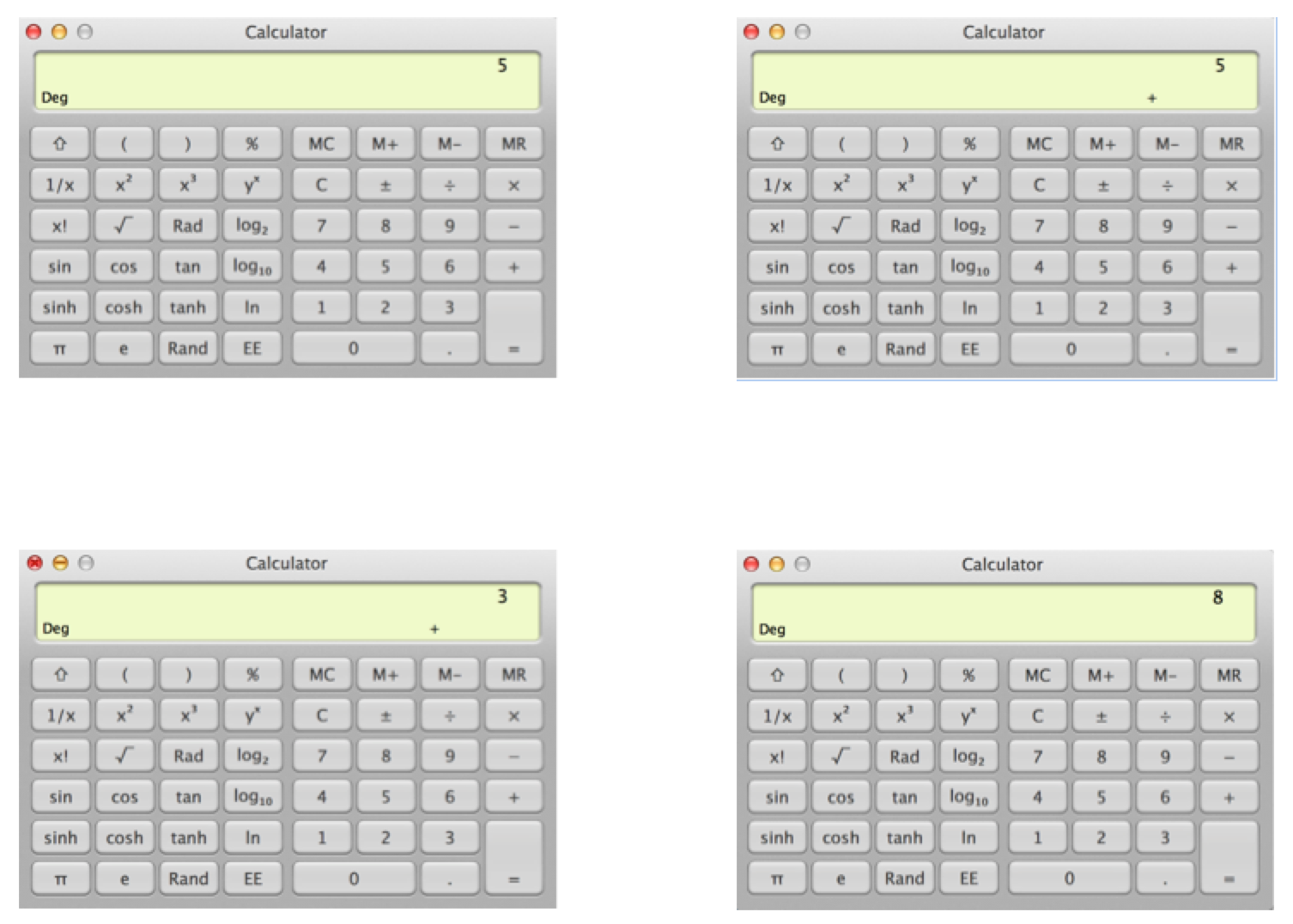
Another aspect of each device that may have affected its performance was the means of entering the data into the device. Both the Google and MacBook calculators were controlled by a mouse and keyboard, whereas the TI-83 was controlled by the users fingers. Fitt’s Law (figure 3) applies directly to this distinction. The shorter distance between the user and the object, the easier it is to interact with the interface. When using the two calculators on the computer the user had to drag the mouse around the screen to hit each button that they needed. When using the TI-83 the user’s fingers had to move a considerably smaller distance to hit each button. Because of this, the time it took to perform each task was decreased. This may have contributed to the high efficiency scores that the TI-83 scored (over twice as fast as the other two devices).

 = a + b \log_2 \Bigg(1+\frac{2D}{W}\Bigg)

***Figure 3***

Another aspect that majorly affected how well the user’s performed the tasks was the way that the calculators were designed to function. Both the TI-83 and the Google calculators functioned like a standard calculator. For the trigonometric problems the MacBook calculator, however, performed differently. For example when entering the first trigonometric problem into either the TI-83 or Google, the commands were “Sin”, “45”, “Enter/Shift”. On the MacBook calculator however the commands had to be “45”, “Sin”. Because of this, when the trigonometric problems came up, every user except for one experienced errors. Once the users learned that this was how this specific calculator was designed to function, they could easily correct themselves and reenter the data correctly, however this lost them time and gave them errors at first. If we had studied learnability in this experiment instead and had allowed the users to come back again and retake the test, the results may have been much different in all the metrics.

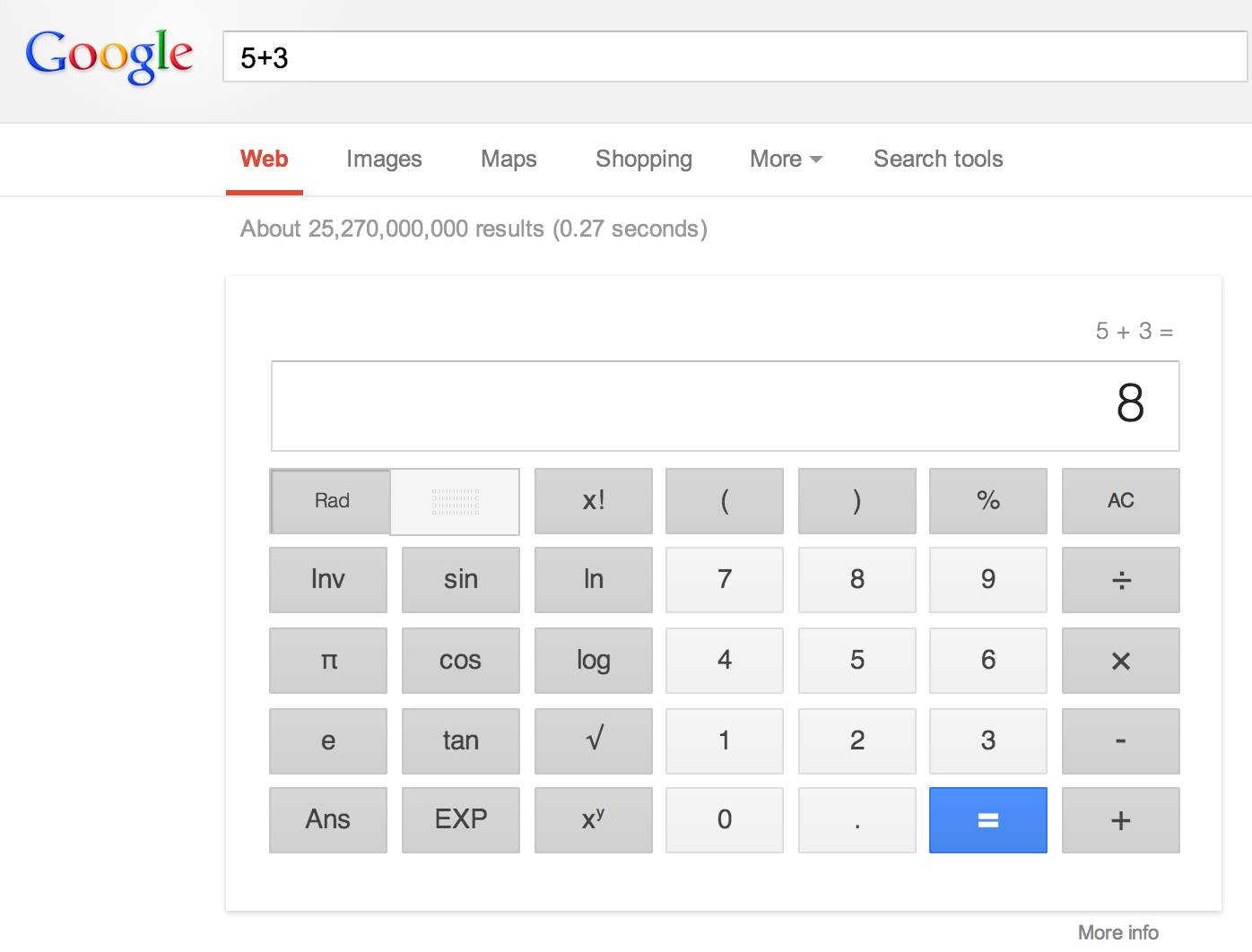
Another aspect that may have contributed to the poorer performance of the MacBook calculator was the style of how the problem is displayed. On the TI-83, for example, after pressing enter, the work you did before is still displayed above it, which can make it much easier to keep track of where you are in a longer problem such as the last set of problems. On the MacBook calculator, however, after you press enter once, the work from that problem is replaced by the answer. Another example of this is shown in the figures below.



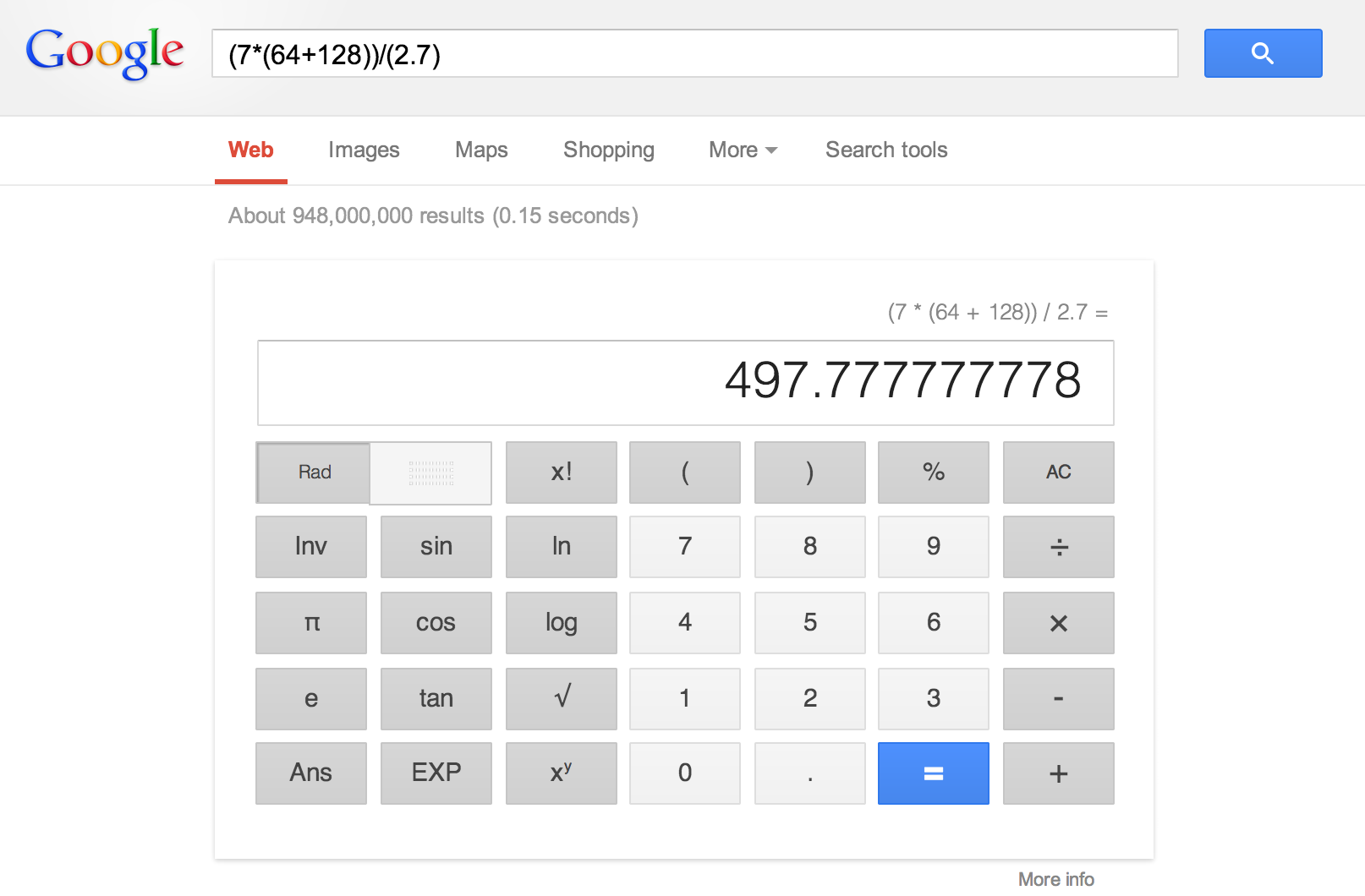
***Figure 4:***

***5 + 3 = 8 on MacBook***

Now compare this system to how the Google calculator functions. Figure 5 shows how the entire problem remains on the screen, even after pressing enter. In figure 6, you can see how this is very helpful for longer problems.



***Figure 5***



***Figure 6***

If I were to do this experiment again, or to do a follow up test, I would increase the amount of test subjects in order to get a more accurate result of the designs of each device. I would also like to test other usabillity metrics such as memorability. I think that this would have interesting results on these devices. It would be interesting to see that if memorability was tested if the difference between the devices would increase or if it would help the MacBook calculator raise its test results.