

Scientific Calculators: Workload and Layouts

Barmartini

Paul George

Sunil Katkar

Vu Hoang Le

Barbara Martinez

Akhil Ravindran

April 28, 2019

1.0 SYSTEM INTRODUCTION

Scientific calculators are utilized by the majority of undergraduate and graduate students in order to perform a variety of problems that range in complexity. The purpose of our project is to compare two existing scientific calculators using workload theory and performance measures to determine which calculator best supports ease of use and functionality. Furthermore, we will analyze our findings to identify potential problems in calculator designs and present solutions that would improve the user experience.

In contrast to basic four-function calculators found in phones, scientific calculators provide a variety of complex mathematical functions. Given the vast number of functions available and the limited amount of space, the necessary steps required to complete certain functions can be too extensive or complicated to remember. Calculators can attempt to alleviate this potential issue by making the steps easy to find or intuitive for novice users, which can also decrease the number of needed button presses. This can be done by configuring and designing the buttons and their labels in such a way that would aid users in finding the necessary buttons with minimal search. The display configurations also provide users with more or less feedback, which in turn impacts their performance on a given task.

Additionally, users who are familiar one calculator may not find that another is consistent in the layout, display, or required steps to complete the same function. By comparing the Casio FX-991ES scientific calculator and the TI-36X Pro, we will be able to collect a variety of data from users who range in their experience with both calculators to understand their respective strengths and weaknesses.



Figure 1. Casio FX-991 ES Plus

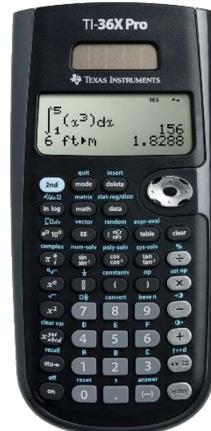


Figure 2. TI 36X-Pro

2.0 DESIGN AND EVALUATION METHODS

Preliminary Survey. Prior to beginning the task, participants will complete a short survey indicating their familiarity and/or experience with the two calculators we will be utilizing in our evaluation. These two scientific calculators will be a Casio FX-991ES plus scientific calculator and the TI-36X. This will allow us to analyze our results accordingly (i.e. experienced Casio users should perform better than novice Casio users).

Task Design. In order to evaluate both systems for relevant comparison, we have designed the following experiment. The experiment included six participants.

Participants were asked to solve three problems per calculator. Two problems had a form of time constraint, while the matrix problem did not. We allowed participants to take as long as necessary to solve the matrix problem. The two time-constrained problems were visually displayed as projections on laptop screens in which participants needed to look up and away from their workspace to receive the necessary information. One problem (calculus) was displayed on a single screen, displaying one element of the problem at a time. The other timed problem (running operations) similarly displayed one element at a time, but instead alternated between screens, one on the left and right of the participant. Both these display procedures forced the participant to look up consistently in order to solve the problem completely and accurately. The untimed matrix problem was displayed constantly on one screen until the participant completed the task. After the completion of all problems, participants were asked to complete a demand vector for the two time-constrained. These steps will be repeated for the second calculator. Below are the mathematic problems presented to each participant.

Calculus

$$A: \int_0^{\pi} (10 * \sin(17x + 6)) dx = 16.366$$

$$B: \int_0^{2\pi} \left(15 * \cos \left(\frac{13x}{5} \right) \right) dx = 92.976$$

Order of Operations

$$A: \frac{3 * (4 + 13)}{(11 - 13)^2} = \frac{2601}{4}$$

$$B: \frac{5 * (2 - 20)}{(15 + 25)^3} = -\frac{729}{64}$$

Matrix

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} * \begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 10 & 8 \end{bmatrix}$$

Workload Method. We used workload theory to evaluate the results from the time-constrained problems, as they are divided into two concurrent tasks. Task A required them to read the problem from the laptop displays based on the problem type, while Task B involves the use of the calculator itself to solve the problem. We utilized conflict matrix to analyze how each system promotes or hinders dual-task timesharing (i.e. demand of task A, demand of task B, conflict between task A and B) based on the demand vectors that were produced by the participants. We chose this method because the completion of the problem required both reading the problem from a distant display and utilizing the calculator in order to solve. The workload method allowed us to compare which calculator had fewer cognitive conflicts in completing both concurrent tasks under the give time constraint.

Time and Accuracy Measurements. In addition to the workload method, we also measured user time and accuracy. Participants were timed on their execution of the problems. Those who finished at or before a predetermined time had the predetermined time recorded as their completion time. For the matrix problems, we recorded the total time for the task completion from beginning to end.

To measure accuracy, we recorded the last display shown on the calculator screen. We noted whether the participant completed the task or had steps to go and whether their display matched the answer given in the earlier section.

Post-Task Survey. At the conclusion of the calculations, participants were asked to provide any additional notes or observations, such as preferences, issues with certain tasks or calculators, suggestions, etc.

3.0 RESULTS

Workload Observations. The results of data collection can be found in Appendix A and the conflict matrices the workload results are based from can be found in Appendix B. For each individual, all the workloads tended to hover within a three-quarters point range. We performed a paired t-Test for both the calculus workloads and the operations workloads, with data sets from each calculator. For the calculus problems, we found a t-Stat of -0.0304, and for the operations problem, we found a t-Stat of -0.2417. Compared to a two-tailed t-Critical of 2.2281, we determined that there is no significant

difference between the workload exerted by the use of either calculator. Some limitations to this methodology will be discussed further in this paper. We performed other measures and did find other interesting results.

Experience and Workload Correlation. Referencing Tables 1 and 2, User 2 had previously utilized both calculators and rated their skills at “Proficient”, therefore it stands that their mental workload was the same for both calculators as well as both problem types. This user also had the same success rate and time, which shows that a proficient user will have similar results regardless of the calculator used based on their expertise and automaticity. This user did note that they preferred the TI calculator over the Casio, even if their time, successful, and mental workload values were equal across. Similarly, User 1 was a new user to both calculators and had the same mental workload for both calculators, although they found the order of operations problem to demand a slightly less mental workload. This user noted that the calculus problems required more mental workload in order to find the less commonly used functions. They also tried to remember values or elements that they did not have time to find so they could later input them. Comparing these results to User 3, who was also a new user to both calculators, they rated the Casio equal for mental workload across both problem types but found calculus to be less mentally demanding and order of operations to be more mentally demanding. When asked, this user noted that the operation symbols were not very salient and difficult to find initially, but once accustomed to the layout, they were able to perform the calculus problem with less mental resources. Additionally, they observed the less commonly used functions easier to find on the TI than the Casio, hence the lower mental workload value. This coincides with our observations that the TI was more searchable in terms of functions, but Casio was better in terms of numeric values and basic operations.

Users 4, 5, and 6 can be classified as experienced Casio users but new TI users, given that User 4 had worked with a different Casio model in the past. An initial hypothesis would be that these users would find the TI significantly more demanding than the Casio. While User 4 and User 5 both rated the mental demand to match this hypothesis for both calculus and order of operations, the difference was no more than 0.375, which is not as significant as originally expected. Furthermore, User 6 found calculus to be less mentally demanding on the TI than the Casio and rated order of operations equal on both calculators. This matches our previous claim that the TI was more searchable in terms of functions, which were necessary in the calculus problems. Furthermore, these workload values indicate that the mental workload for an experienced Casio user is similar or only slightly lower than the mental workload for a new TI user. A user who is experienced in both calculators would either have the same mental

workload, as did User 2, or a slightly less demanding workload for the TI, as User 2 had noted in observations.

Time and Workload Correlation. Due to the nature of how we designed the calculus and running operation problems, many users felt the need to solve the problem immediately when the equal sign (=) was flashed on the display and therefore fell within 1 minute, plus or minus 10 seconds. Only 3 participants went over this time frame, though none of those participants went over 2 minutes total time. As a result of the overall uniform time of most users, with only 60 seconds or less additional time for users that went over, there was little variance across times. Additionally, there was very little difference in workload between the two calculators for either type of problem. Considering these two factors, we were unable to find any correlation or relationship between workload and time. (See appendix for Workload vs Time graphs)

Experience and Accuracy Correlation. From the 6 participants, half had previous experience with the Casio calculator and the other half had not. Of the 3 who had no experience, 1 had previously used a different Casio model in the past and therefore was able to complete all 3 problems utilizing the Casio calculator. The 2 users who had no experience and rated their skills as “Low” were either unable to solve any of the problems or only able to solve one. Of the 3 users who had experience, 1 rated their skills as “Proficient” and was able to complete all the problems correctly. The other 2 experienced users were able to correctly complete 2 of the 3 problems.

Regarding the TI calculator, only 1 user had previous experience with the calculator and rated their skills as “Proficient”. This user completed all 3 problems correctly and later noted that they preferred the TI over the Casio as it yielded a faster computation speed. It should be noted this was the same user that also rated their skills as “Proficient” for the Casio. Of the 5 inexperienced users, one had previous experience with a TI graphing calculator and was able to correctly complete all 3 problems. The remaining 4 inexperienced users varied in accuracy rates: 1 user answered all 3 correctly, another user answered 2 of the 3 correctly, and the last 2 users only were able to answer 1 of the 3 correctly.

Comparing the two calculators for overall success, inexperienced users were able to correctly complete at least 1 or more problems with the TI calculator and yielded better performance accuracy results for these new users. New Casio users were only able to correctly complete 1 problem, if any problems at all (disregarding the user who had worked with a different Casio model previously). Experienced users of the Casio did not all perform with 100% accuracy as would be expected. Additionally, the “Proficient”

user who was equally successful in both calculators with 100% accuracy noted that they preferred the TI for easier and faster computation. This can be attributed to the variety of problems identified previously, such as button searchability and intuitiveness, also accounting for the stress of time pressure and additional cognitive workload if the participant had to commit values to working memory.

When using the Casio to perform the calculus problem, half of the users were able to complete the problem correctly but only 2 of those users had previous experience with the Casio. This means that one of the experienced users was unable to correctly complete the calculus problem. The experienced user who was successful was the user who had used another Casio scientific calculator. Alternatively, the users who were unable to solve included 2 inexperienced users and 1 experienced user. The order of operations problem had more success, with 4 successful solutions. Similar to the results of the calculus problem, 1 experienced user was unsuccessful and therefore 2 inexperienced users were successful (again, 1 was the user who worked with a different Casio). Finally, the matrix problem had the same success rate as the order of operations but both unsuccessful completions were attributed to new users.

Comparatively, the failure calculations when using the TI for either calculus (3 failures) or order of operations (2 failures) were only done by inexperienced users. It should be noted that one of the new TI users was successful with all 3 problems, similar to the case described previously with the Casio model. The distinguishing factor between these two cases is that the Casio user had previously worked with a different Casio scientific calculator; the TI user had worked with a different TI graphing calculator. The similarities between Casio scientific calculators would be closer than the similarities between a TI scientific and TI graphing calculator. All users, new and experienced, were able to solve the matrix.

Comparing the calculators for individual problem types, there were an equal number of successes for the calculus and order of operations problems. The notable difference is that no experienced user was unsuccessful with the TI, or rather new users had more success with the TI. Another notable difference is that all users were able to solve the matrix, while only experienced Casio users (current or other models) were able to solve the matrix problem. As mentioned, this is due to the more intuitive design for matrix calculations using the TI calculator over the Casio model.

Time and Accuracy Correlation. As previously defined, the calculus and running operations problems were time sensitive in that each element of the problem was flashed for a given number of seconds, including the equal symbol. Therefore, participants would generally press enter at the end of the projection and fall within the one-minute time frame, plus or minus 10 seconds (see Tables 1 and 2).

Some users went over this time to go back and correct any elements or search for missed elements that they were originally unable to find before the next projection. One user went over the designated time frame for both problems when using the Casio, although this was the user who had previously worked with a different Casio scientific calculator. Going over the designated time frame, in this case, led to correct responses for both questions. An experienced user went over time for the order of operations problem on the Casio, approaching 2 minutes, but was still unsuccessful. In contrast, two users went over time for the calculus problem on the TI and both were successful. It should be noted that one of these users had previous experience with a TI graphing calculator. This user, as well as the Casio user previously mentioned, probably needed additional time to search for buttons that they had expected in a different location based on their previous experience with a different but somewhat similar calculator. Although all these users went over the time frame to search for the necessary buttons, they still needed to have these elements committed to memory. Therefore, while time was no longer a stress factor, memory was a critical aspect in success.

Users who fell within the time frame had success dependent on their searching capabilities and/or prior experience. More new users were successful within the time frame when using the TI than the Casio, which highlights the ease of searchability of the TI over the Casio as previously noted. Furthermore, not all experienced Casio users were successful within the time frame, also illustrating searchability issues. It should be noted that participants were asked to complete the problems with the Casio first and therefore could potentially pre-search for these elements in the few seconds before presented with the problems using the TI. This would level the prior experience of seeing the buttons between experienced users and new users. Therefore, it still stands the experienced Casio users were not always successful with this prior knowledge or experience in button location, but new TI users may have benefitted from even a brief visual scan of the buttons.

The matrix problem was untimed and constantly displayed, therefore there was no time constraint to act as a stressor while participants searched for buttons and possible strategies to solve. Additionally, there was less stress placed on working memory since the problem and all its elements remained available the entire time. Experienced Casio users, as well as the user who had worked with a different Casio model, were able to successfully solve in 1:33 or less. The new users had times of 1:33 and 4:36. The first gave up and knew their answer was unsuccessful while the second was more determined to solve, although also unsuccessful. This indicates that even with an extended amount of time, the strategy to solve was not easily and intuitively found by new users. In contrast, all users were successful

in solving the matrix problem with the TI. The times ranged between 1:05 and 3:40 for new users, and 0:28 for the proficient user. In comparing the successful matrix times between calculators, experienced Casio users had the same time new TI users. Furthermore, the proficient user had a time of 1:33 for the Casio and 0:28 with the TI. These findings highlight the ease and intuitiveness of matrix calculations with the TI compared to the difficulty of discovering these steps in a Casio, as previously described.

4.0 LIMITATIONS OF EXPERIMENTAL DESIGN

Our results were not as decisive as we would have liked. We made some observations that might have affected our results and additional things we could have done that would have made our findings more significant. For example, due to how we ran our experiment, the participants could reasonably be expected to perform a quick search on the TI calculator immediately upon receipt. This could correspond to the higher correct answer rate demonstrated by those using the TI calculator.

If we were to use a larger range for the workload demand vectors, we might have been given clearer results in whether the workloads between different calculators were significantly different. Another method that we could have used was the NASA TLX. Because this test tends to provide larger and more variant numbers, we might have at least observed means that were further spaced apart for individual observations.

5.0 DESIGN AND AESTHETIC PROBLEMS

Display Screen. Casio has larger display font for elements on screen, which means more elements are cut off when looking at longer problems. TI has a smaller font on the display but it can be so small it is difficult to read while concurrently looking at a projection of problems. This means that more of the whole problem is displayed and requires less toggling back and forth to edit any issues or mistakes. More information is transmitted within the given screen space and therefore the user utilizes less resources, cognitively and manually, in order to determine whether the entered values are correct or incorrect. In contrast, the Casio only transmits a portion of the information if it exceeds screen space, forcing the user to utilize more resources, cognitively in committing the values to working memory as they disappear from the screen and manually in pressing more buttons to access all the necessary information. If too many numbers must be committed to working memory, the user may have the toggle back and forth more than one time. For longer and more complex problems, the Casio forces

users to view that problem in multiple parts or chunks, while the TI allows the user to see the problem as either a whole if it fits within the display screen or a smaller quantity of chunks when compared to the Casio.

Saliency of Buttons and Searchability. TI operation buttons, while a different color than the digit buttons and function buttons, has no contrast between the button and the font coloring for the operation. Therefore, it was more demanding to look for the correct operation button when compared to the Casio, given that the button (background) is indistinguishable from the operations (figure). Additionally, the color of the basic operation buttons was metallic and therefore had a glare when hit by light. Therefore, in addition to having difficult distinguish the operation as a result of the lack of contrast, users had to take more time to adjust the positioning of the calculator to remove the glare. While the Casio was easier to quickly search and locate the basic operation buttons, the labeling of function buttons was not as easy. Each button had at least 2 functions, the one identified by the label on the button itself and another labelled above the button in orange. This is similar to the design of the TI; the only difference is that the TI utilizes blue. The main difference is that some of the buttons on the Casio have more than 2 functions per button and so the labelling above the button becomes cluttered with up to 3 different colors. Additionally, the functions and labels have low processing proximity but are still displayed in close spatial proximity. This increases the chances of entering the wrong function or element as well as the probability of performing in the wrong mode. Some of the TI buttons have 2 labels on the button itself, but these functions are similar and have high processing proximity, such as ln and log functions or sin and sin-1. Therefore, this design utilizes the law of similarity to take advantage of limited space efficiently in a way that would minimize slips during calculations.

6.0 FUNCTION AND USABILITY PROBLEMS

Intuitiveness. TI provides feedback and cues that aides user in forming matrices, even for unexperienced users. Casio, while it has a matrix button, requires other buttons to be pressed that an inexperienced user would not think to look for or use in order to solve. First time users or novices were unable to determine the necessary steps to create a matrix with the Casio, with some giving up completely due to frustration. In contrast, the TI provides clear menu options after selecting the matrix button, allowing first time users or novices to at least problem solve. Additionally, the TI provides more detailed feedback errors that are understandable to a human and better equip them with how to fix their mistake. The TI is better in supporting the top-down mental models with relevant and

understandable bottom-up visual cues, without an overload of insignificant data and feedback. In contrast, the Casio relies too heavily on top-down processing, which would benefit an expert users who does not need bottom-up feedback or cues to perform tasks, but hinders a novice user who is trying to problem solve and form a strategy.

Mode Confusion. When completing the matrix problem, the Casio provided no bottom-up feedback that notified the user they were in fact operating within the matrix mode. As a result, users are functioning with primarily top-down influences from past experiences and expectations to develop a strategy when in fact that strategy is being operated in a different mode completely. Given that the calculator was not actually in matrix mode, as that required additional steps beyond pressing the “Matrix” button, the top down mental model developed by the new users are for a different and incorrect mode. This, in turn, means that the top down mental models do not match the task either. As shown in Figure 1, many of the buttons have more than one function or mode, and therefore the lack of bottom-up feedback to indicate the mode or function that is being used can result in mode confusion and furthermore incorrect values.

Inconsistency Between Calculators. As illustrated in Figure 1, the design and layouts of both calculators vary. This variation also applies to the functions and the steps needed to complete complex calculations. Ideally, both calculators should follow the principle of consistency so that the design and functions match what users expect from a calculator to accurately support a wide range of top down mental models. This is particularly important for new users that have no mental model of the given calculator but may have mental models of other types of calculators. The inconsistency between calculators, particularly when analyzing the data and observations from the matrix problem, shows that the TI was closer in matching new users’ mental models (i.e. intuitive cues to guide users to the next step, sufficient feedback if errors arose). Users found that it matched the steps normally taken when completing a matrix calculation by hand or those taken with a different calculator. In contrast, the Casio only support top down mental models of users who had used the Casio before, limiting the scope of users who could find similarities and consistencies between their past calculator experiences and Casio.

Ease of Error Correction. When users forget an element of a calculation and toggle to add that element into the equation, the element replaces another one instead of being inserted. If a user is unaware of this the first time, they will not think to commit the element being replaced to working memory and may need to restart the entire task from the beginning. Alternatively, they may try to recall the number that was replaced by bringing the entire problem to working memory. Given that the user was unaware

of this potential problem though, their efforts to initially commit these values to working memory was not as focused. As a result, they may insert a different number from the given task, which would result in an incorrect answer. Users who are familiar with this feature or who may anticipate it based on previous experiences with similar calculators, will utilize more cognitive resources to commit values to memory. First, they will make a stronger effort to commit the values of the problem to memory in order to avoid the issue previously mentioned. Additionally, they will commit the values surrounding the one to be inserted to working memory so that they are not lost when replaced. In actuality, the TI has an insert button to avoid this extra demand, but not all users may be aware of this function or may find it just as demanding or wish to avoid the extra steps it requires.

7.0 REDESIGN SOLUTIONS

Practical Solution. The proposed practical solution involves a new design that incorporates the number pad layout of the Casio with the function pad layout of the TI. The salience of the number keys (due to the Casio layout) along with improved accessibility of the most frequently used variables and constants (due to TI function layout) would help in improving user experience and overall performance of the human operator.

Two additional features were incorporated into this design:

- Change of the exponential symbol from [x^{\square}] to the caret symbol [\wedge]
 - This ensures that the symbol on the calculator is consistent with math terminology
- Incorporating a raised dot on the number 5
 - This would help the user know the relative position of the other keys without looking

No-Budget Solution. For this solution, the function pad from the previous design was replaced with a context sensitive touchscreen. This would be flanked by a few menu buttons at the top along with the physical number pad at the bottom. The options displayed on the touchscreen would depend on the mode of operation that the calculator is in. This would help in reducing the number of clicks needed to access a particular sub-function. Provisions have also been made to incorporate voice inputs of the user in order to perform hands free computation.

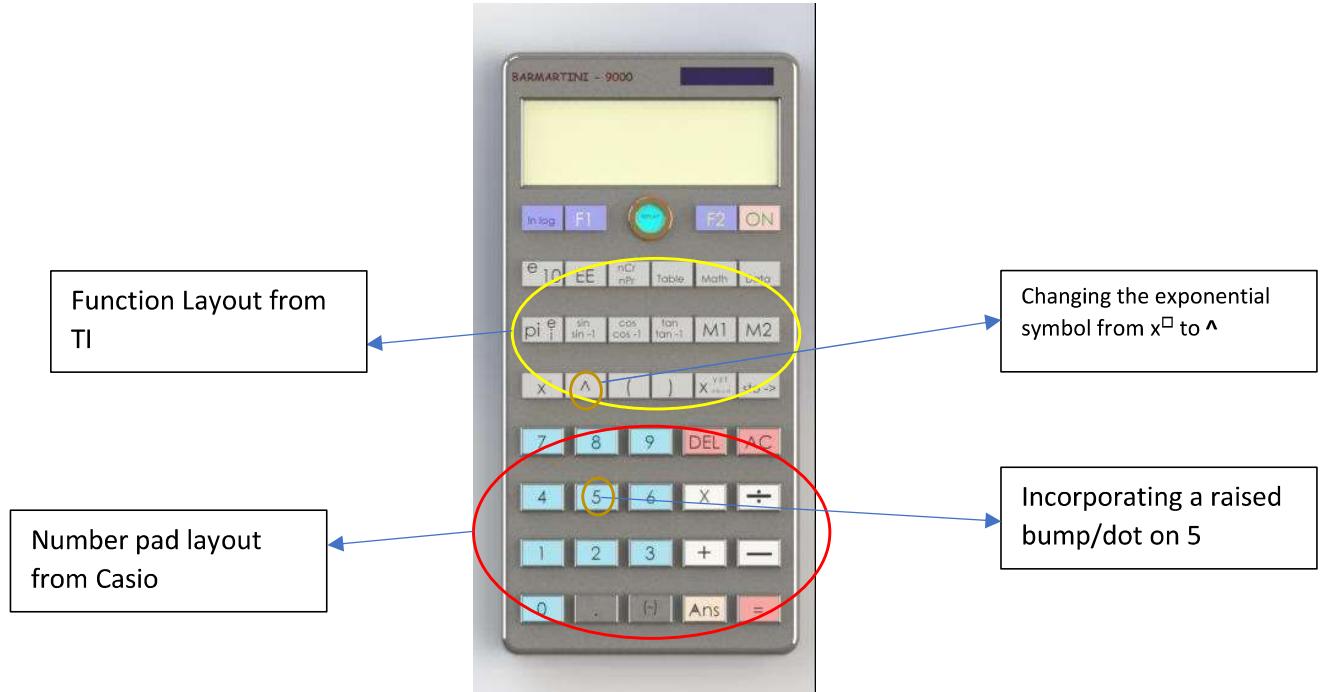


Figure 3. Calculator Redesign, Practical Solution

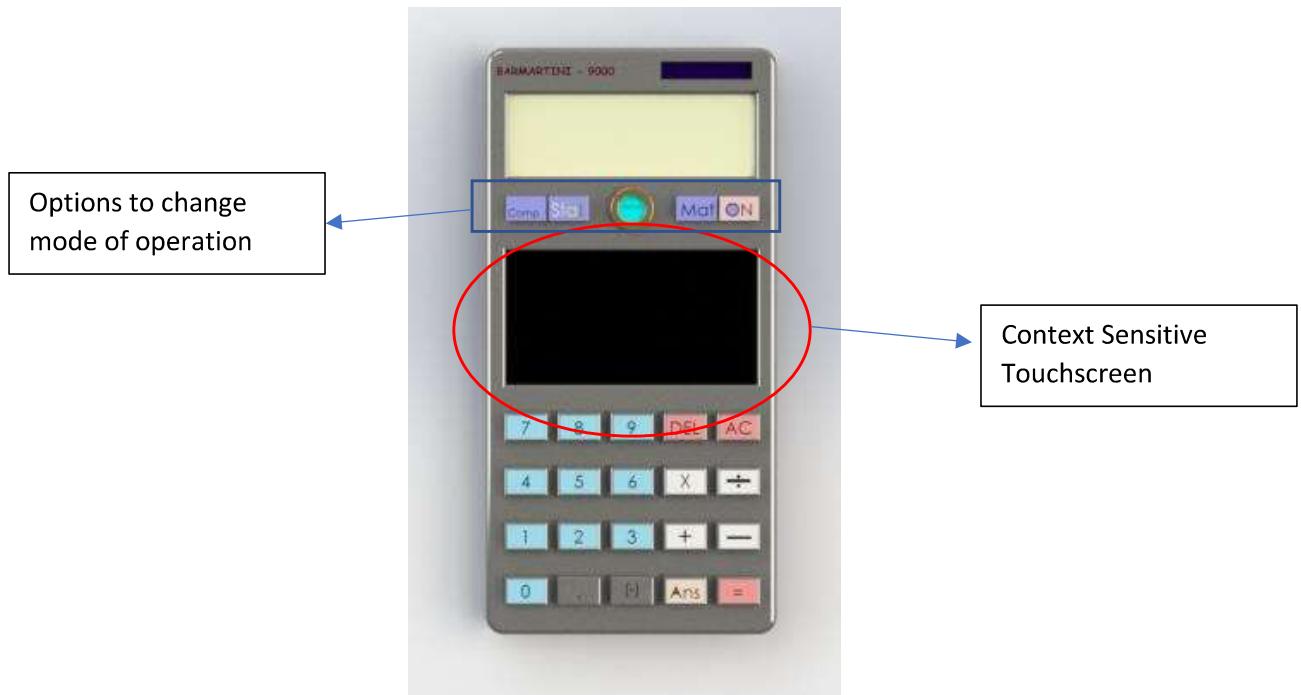


Figure 4. Calculator Redesign, No-budget solution

APPENDIX A. DATA COLLECTION

Table A1. Casio Calculator Observation

	User 1	User 2	User 3	User 4	User 5	User 6	Accuracy
New user	Proficient	New user	Other Casio	Experience	Experience		
Calculus Time	1:02 (N)	0:53 (Y)	0:56 (N)	1:42 (Y)	0:58 (N)	1:08 (Y)	50%
Workload	14.25	13.125	13.5	13.125	16.025	13.625	
Operations Time	0:57 (N)	0:53 (Y)	1:05 (Y)	1:15 (Y)	0:57 (Y)	1:53 (N)	66%
Workload	13.75	13.125	13.5	13.125	15.777	13.375	
Matrix	1:33 (N)	1:33 (Y)	4:36 (N)	0:58 (Y)	1:25 (Y)	1:06 (Y)	66%
Accuracy	0%	100%	33%	100%	66%	66%	

Table A2. TI Calculator Observation

	User 1	User 2	User 3	User 4	User 5	User 6	Accuracy
TI Graphing	Proficient	New user	New user	New user	New user		
Calculus Time	1:43 (Y)	0:53 (Y)	1:10 (N)	0:53 (N)	0:59 (N)	1:58 (Y)	50%
Workload	14.25	13.125	13.125	13.5	16.4	13.375	
Operations Time	0:55 (Y)	0:53 (Y)	0:58 (N)	1:08 (Y)	0:55 (N)	0:57 (Y)	66%
Workload	13.75	13.125	14	13.25	16.025	13.375	
Matrix	1:53 (Y)	0:28 (Y)	3:40 (Y)	1:58 (Y)	1:23 (Y)	1:05 (Y)	100%
Accuracy	100%	100%	33%	66%	33%	100%	

Tables 1 and 2 show the user time for each problem, Table 1 for Casio and Table 2 for TI. Additionally, they indicate the user's prior experience and accuracy (Y for correct completion, N for incomplete or unsuccessful). There is an additional row below Calculus and Running Operations to indicate each user's mental workload demand per each problem type for both calculators.

APPENDIX B. WORKLOAD CONFLICT MATRICES

Table B1. User 1 Conflict Matrices

Table B2. User 2 Conflict Matrices

Table B3. User 3 Conflict Matrices

Table B4. User 4 Conflict Matrices

Table B5. User 5 Conflict Matrices

Table B6. User 6 Conflict Matrices

APPENDIX C. TIME VS WORKLOAD REGRESSION

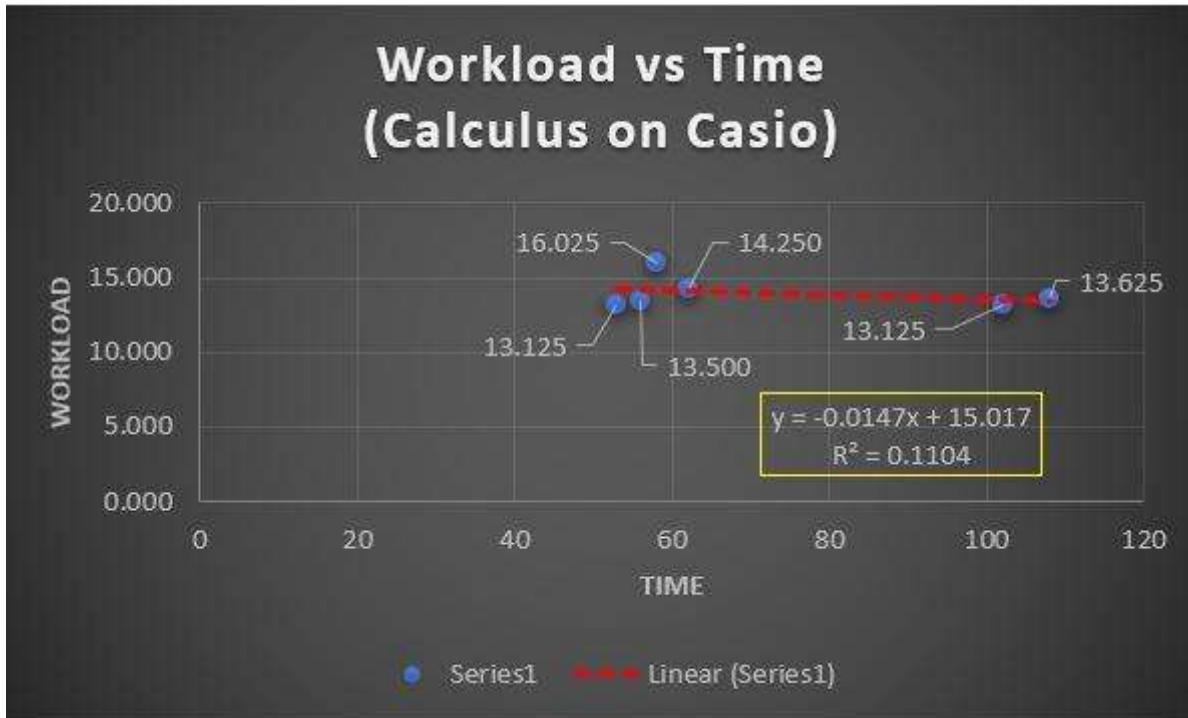


Figure C1. Workload vs Time, Calculus on Casio

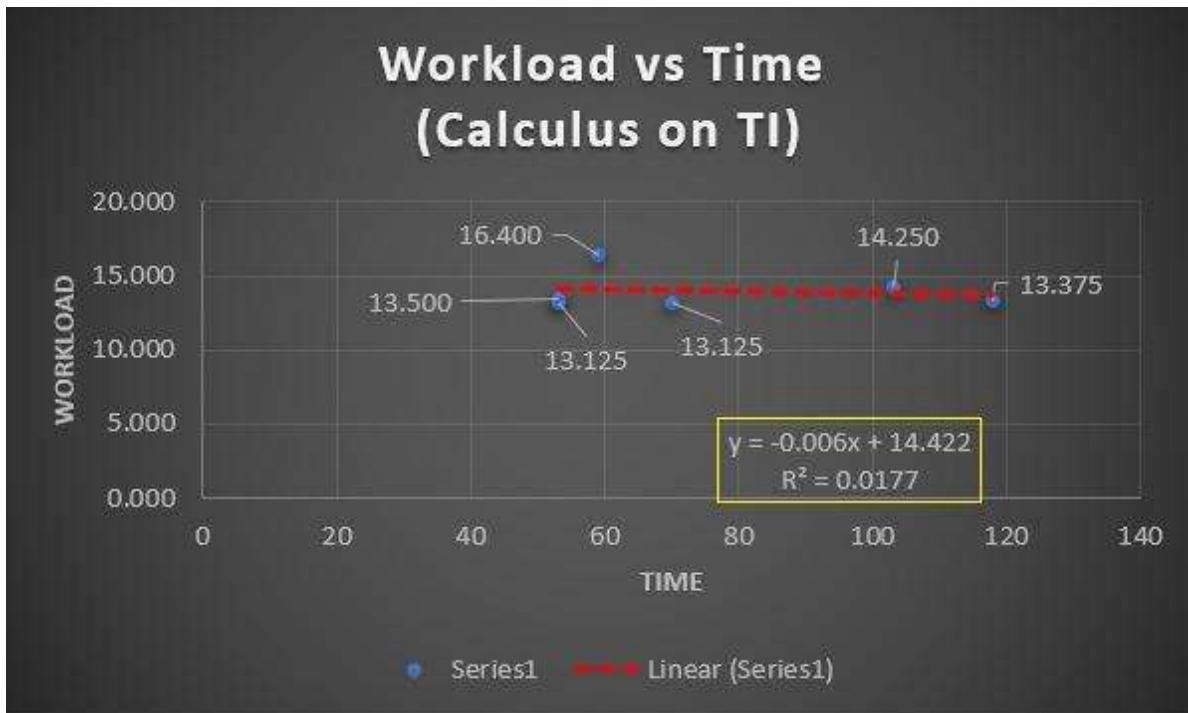


Figure C2. Workload vs Time, Calculus on TI

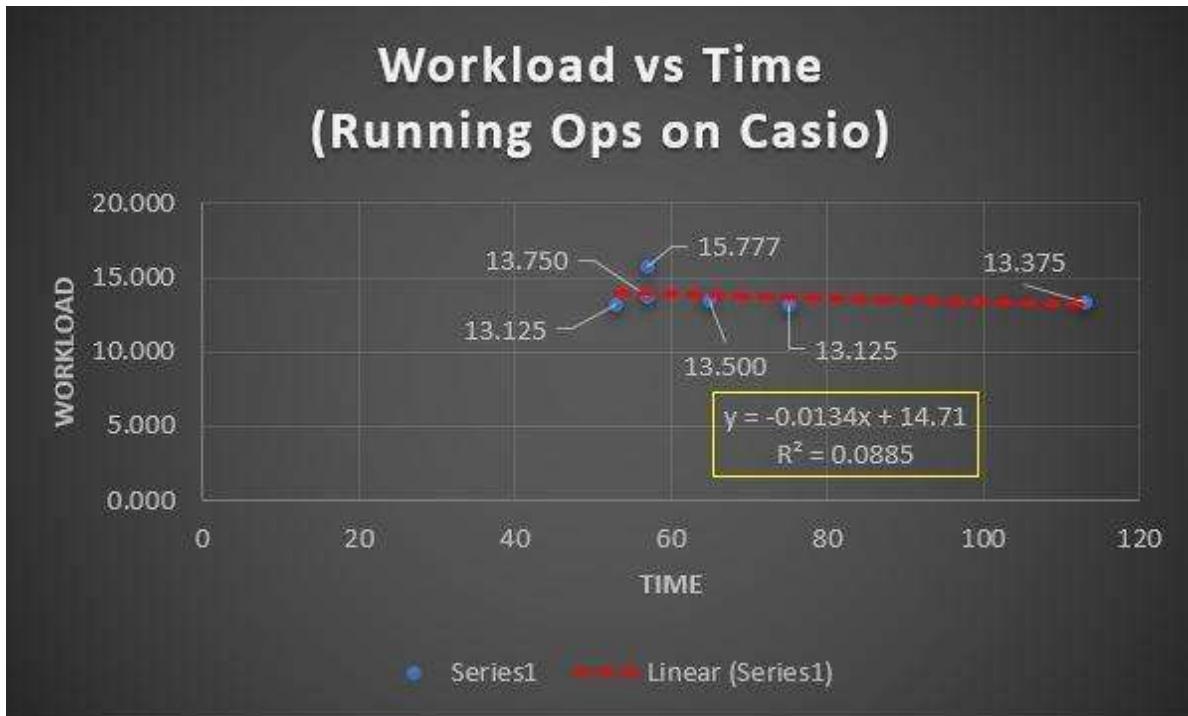


Figure C3. Workload vs Time, Operations on Casio

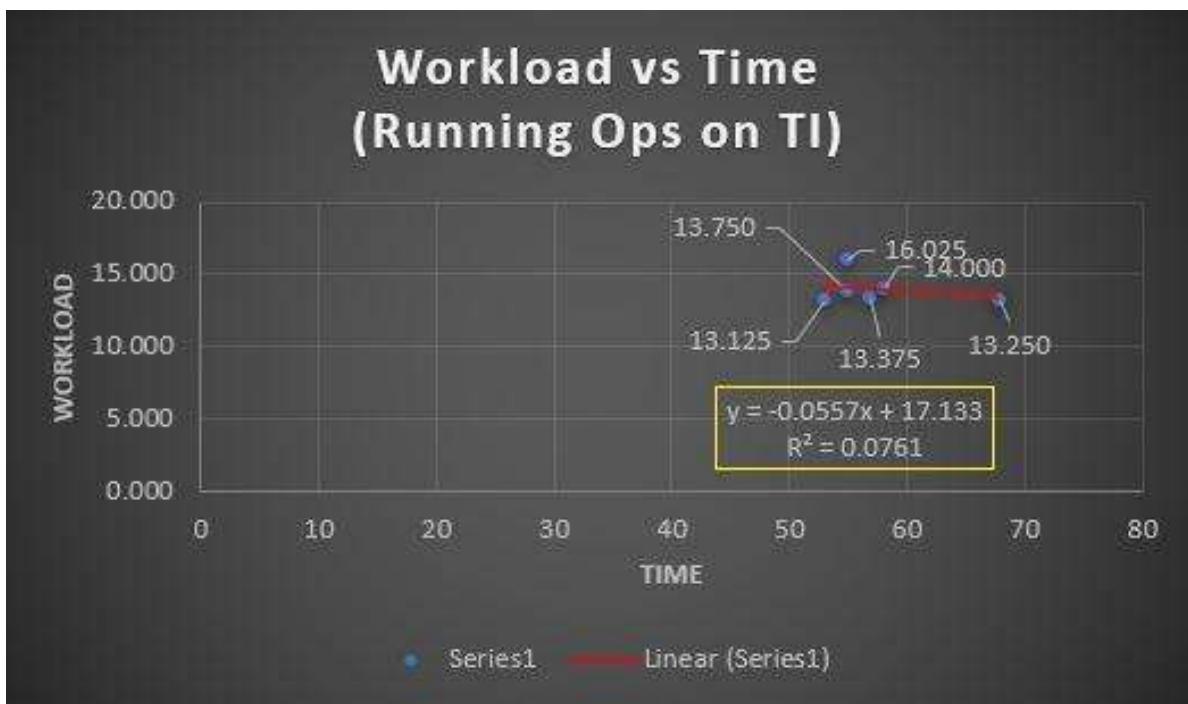


Figure C4. Workload vs. Time, Operations on TI