The Problem State: As a Bottleneck in Multitasking

Ganesh Pitchiah
IIT Kanpur
Department of EE
ganeshp@iitk.ac.in

Mentor: Prof. Amitabha Mukerjee
IIT Kanpur
Department of CSE
amit@cse.iitk.ac.in

ABSTRACT

In the area of multitasking, the main challenge is to predict when and how tasks interfere. Of the many theories, Threaded Cognition theory [Salvuci & Taatgen, '08] is a recently proposed integrated theory of multitasking. Based on this, here we study the role of problem state (PS) resource in causing interference. Problem state is the directly accessible intermediate information involved in mental transformation tasks. The prediction made is that interference is caused whenever two tasks require the problem state. To test the prediction, an experiment is carried out where subjects have to carry out a text entry and a subtraction task concurrently. Both the tasks have two versions: one that requires a PS and one that doesn't. There is an over-additive interaction effect, indicative of interference when both the tasks involve a PS. To compare the observations, a cognitive computation model built by [Borst & Taatgen, '10] was used. The model too shows an over-additive interaction effect which confirms PS as a bottleneck.

Keywords

problem state (PS), threaded cognition, rANOVA

1. INTRODUCTION

For a brief etymology, the use of the term multitasking arose with the rise of the computer industry in 1960's. It indicated the computer's new ability to process multiple tasks simultaneously. However, psychologists have been investigating the human ability to multitask atleast since the 1930's. Based on this work, various detailed cognitive models have been developed: concurrent multitasking, task switching, sequential multitasking etc. It is interesting to note that the exact same words and their meanings work out for computers as well. So, when it comes to building computation models for multitasking, components of the proposed theories might already have an unintended implementation in computers to which we can relate to.

There have been many studies and common results in the area of limits of human multitasking. A select observation in this matter with a neurological basis is: "the most anterior part [of the brain] allows [a person] to leave something when it's incomplete and return to the same place and continue from there, while Brodmann's Area 10, a part of the brain's frontal lobes, is important for establishing and attaining long term goals. When we multitask dissimilar tasks, the brain is forced" [Wallis, '06] to carry out both the activities in its anterior: planning as well as saving entire states of incomplete tasks which places limits on our multitasking ability. The role of Brodmann's Area 10 will be related to the goal state in Section 1.1.

1.1 Related Work

Problem State Resource: In their terminology [Borst & Taatgen, '10] "it is used for storing intermediate information that is necessary for performing a task". For example, if asked to solve $5x - 7 = 10$ mentally, the intermediate solution $5x = 17$ is stored in the PS resource. This concept has a neurological basis: "blood oxygen level dependent activity [was found] in the posterior parietal cortex that correlates with the transformation in mental representations" [Borst & Taatgen, '10].

The concept of PS is not entirely new and similar concepts have existed in previous models of multitasking. For example: the episodic buffer [Baddeley '00] as a component of working memory, which can store temporary information deduced from available information (from the world as well as long-term memory). It is distinguished by its conscious awareness during retrieval.

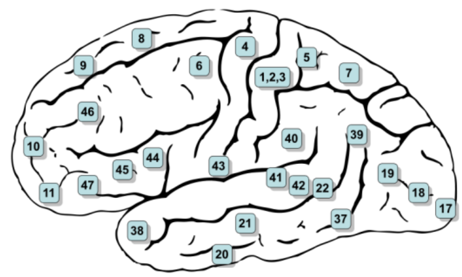


Figure 1. Lateral surface of the brain with Brodmann's areas numbered. [File:Gray726-Brodman.png - Wikimedia Commons, '07]

Threaded Cognition Theory: [Salvuci & Taatgen, '08] According to this theory, every task can be represented by a cognitive thread. More specifically, a thread can be identified by the associated goal of a task. For example: while driving a car, though multiple modalities (vision, touch, hearing etc.) are utilized, the goal of reaching the destination can be represented by a thread which can access the multiple resources in parallel. The idea is that while the multiple modalities and resources (declarative and procedural knowledge) can operate in parallel without a central supervisory control, each can be accessed by only thread at once. For example: talking on the phone while driving, at any given instant the hearing modality provides input to only one of the two active threads. While this serial nature places a limit on multitasking ability, some resources act as a bottleneck by increasing the execution time from more than the normal. It means, the execution time of a single thread increases not only due to the competition among the threads for the procedural processor resource but also due to change in the response time of the particular resource based on the tasks.

In threaded cognition theory [Salvuci et al., '09] and ACT-R [Anderson '05] mental representations have two separate components: a goal state and a problem state. Goal state stores the state of current goal: complete, incomplete etc. while problem state stores "temporary intermediate information" [Anderson '05]. The division can now be justified with a neurological basis of two physically separated regions: Brodmann's Area 10 in the anterior for goal state and posterior parietal cortex (Brodmann's Area 7) [FitzGerald et al., '11] for problem state.

Timeline: From the primary paper [Borst & Taatgen, '10], PS resource appears to succeed Threaded Congition Theory. However, this is not entirely true and both developed hand-in-hand with a common author Taatgen in all the four papers. As mentioned earlier, PS is not an entirely new concept and its implementation already existed in ACT-R [Anderson '05] before its formal identification as such.

Table 1. Timeline expressing relation between PS and Threaded Cognition Theory

Paper	Focus
[Borst & Taatgen, '07]	Identifies the PS resource.
[Salvuci & Taatgen, '08]	Propose the Threaded Cognition Theory for the first time.
[Salvuci et al., '09]	Revised Threaded Cognition Theory to include a PS module.
[Borst & Taatgen, '10]	To investigate the possible role of PS as a bottleneck in multitasking within the Threaded Cognition Theory.

2. EXPERIMENT

All content in quotes in **Section 2** is borrowed from **[Borst & Taatgen, '10]**. The experiment was carried in strict adherence to those followed in the primary paper with the only variation of 10 male subjects instead of "15 subjects (10 female)".

2.1 Participants

The participants were 12 students from the Indian Institute of Technology, Kanpur. All participants were male in the age group of 21-24 years and had "normal or corrected-to-normal" vision.

2.2 Design [Borst & Taatgen, '10]

Participants have to perform two tasks concurrently: a subtraction task and a text entry task. The subtraction task is displayed to the right and the text entry to the left as shown in **Figure 2.**

The subtraction task involves carrying out a 10 digit subtraction with a positive 10 digit result. The digits are entered using the number keys on the keyboard. In the subtraction task there are two versions: easy and hard. In the easy version, all the digits of the upper row are greater than or equal to the corresponding digits in the lower row i.e. no carry. In the hard version, there are any six columns where the upper digit is lower than the lower digit i.e. requires a carry. A proposed further modification [Borst & Taatgen, '10] was made such that on entering a digit, the entire column is hidden and participants have to remember whether a carry operation took place, to perform the next column of subtraction. With lack of information in the world, the assumption is that PS would be required to store the information.

The text entry task involves entering letters displayed above the virtual keypad by clicking with a mouse. This task has no visual feedback for the entered letters, the reason for which would soon be clear. This task too has two versions: easy and hard. In the easy version, letters from a meaningful 10 letter word are displayed in random order one after the other which makes predicting the next letter difficult. In the hard version, a meaningful 10 letter word is displayed at the beginning of a trial when participants have to remember the word. Once the first letter is entered, the word disappears; subjects have to enter the letters from memory. Again, here the assumption is that PS would be required to remember the word and the previous letter entered. The cursor could be used an indicator of the previous letter entered by not moving it after clicking a letter. If so, only the word would be stored by the PS.

Based on the difficulty level of text and subtraction task, participants could face any of the 4 possible versions (Subtraction/Text: easy-easy, easy-hard, hard-easy or hard-hard). Apart from this, 2 versions (Subtraction/Text: easy-hard or hard-easy) of single tasking (i.e. complete one entire text or subtraction task before starting the other) were also presented. This was to record the normal response times.

To enforce concurrency, after entering the first digit or letter, only one interface is visible and active at any given instant of time and after each digit or letter is entered, the corresponding interface is hidden and the other interface (i.e. text or subtraction) one is made visible. At the beginning of each trial, the subject is given 200 points which is displayed at the top of the screen. The points reduce at a rate 2 per second. At the end of each trial, 10 points per every correct letter or digits is added to the total score along with remaining of 200 points and display on a score page. This encourages the subjects to be quick as well as accurate.

2.3 Stimuli and Apparatus[1]

The numbers for subtraction task are generated randomly at the beginning of each trial with the already mentioned constraints. The primary paper [Borst & Taatgen, '10] used 10 letter Dutch words from the CELEX database while we are using English words from the same database.

The experimental conditions of the primary paper [Borst & Taatgen, '10] have been recreated. On a 19" monitor, the interfaces each measured: 9cm in width and 4.8cm in height. The distance between the interfaces measured 10cm. The screen was 75cm away from the participants.

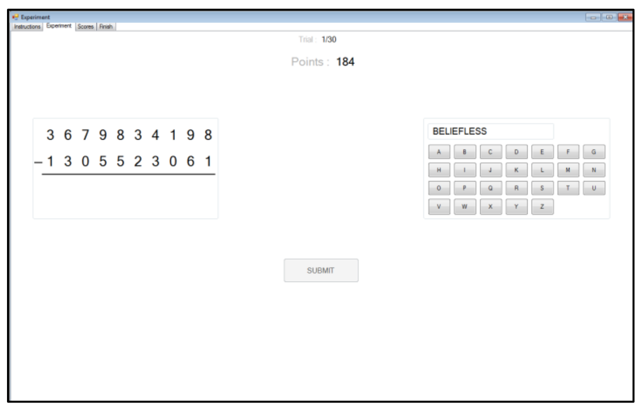


Figure 2. Screenshot of Experiment.

2.4 Procedure [Borst & Taatgen, '10]

The participants are first shown an instructions page with an outline of the experiment and the available user inputs. After this, participants go for 10 sample runs: 6 of single tasking and 4 of multitasking. Here the participant familiarizes to distinguish the difficulty of the individual tasks (for example: vanishing column indicates hard subtraction).

[1] The Windows form is available for download at: http://home.iitk.ac.in/~ganeshp/se367/project/code.zip

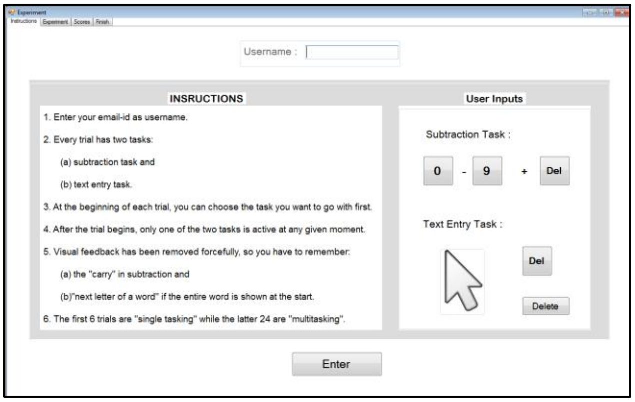


Figure 3. Screenshot of Instructions

Once the experiment begins, the first 6 trials of the experiment are single tasking. The next 24 trials comprise of dual tasking. The experiment is carried out in 6 blocks, each block comprises of the 4 versions in random order, with the constraint that the beginning trial of a block and the last trial of the previous block are not the same. Thus, participants face 24 trials of multitasking in a semi-random order. The score after each round is displayed on the score page for a 5-s period. Apart from this, a 30-s break was provided after 15 trials. The entire experiment takes 50 minutes (10 min. for sample runs and 40 min. for the experiment).

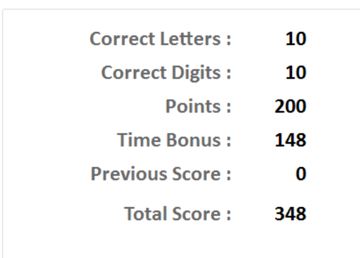


Figure 4. Screenshot of Score

2.5 Model [Borst & Taatgen, '10]

A computational cognitive model has been built using threaded cognition in ACT-R with the problem state module [Borst '10]. As mentioned earlier ACT-R has a problem state module, however, it can store only one chunk of information at a time. One chunk here could be understood as the intermediate information stored for one task (ex: the word and letter entered in hard text entry). In ACT-R [Anderson '05] information retrieval from PS module takes no time however, it takes 200ms to change contents in it. This is because the value of PS is pushed to declarative memory and retrieved when required.

This threaded cognition model in ACT-R [Anderson '05] predicts interference when two tasks require a PS due to two reasons. First, the execution time increases theoretically by 200ms when tasks are carried out concurrently. Second, ACT-R introduces error whenever retrieving information from older memory in accordance with human behavior.

The focus of the primary paper [Borst & Taatgen, '10] is to check the role of PS as a bottleneck. If the model explains the results, PS acts a bottleneck possibly by its limited capacity and the time it takes to integrate available information. (An example might be retrieving information from declarative memory but it is not the only possible one).

2.6 Results & Analysis

2.6.1 Response Times

During the experiment, two parameters were measured: the response time and accuracy. The response time for a single letter entry can be defined as the time elapsed between the appearance of the interface and the user clicking a letter. Similarly, for subtracting a digit, it would be the time elapsed between the appearance of the interface and the user entering a digit. In the experiment of 24 trials, every version has 6 occurrences each having 10 response times for text entry, which amounts to 60 samples from a single user for text entry in a single version. Thus, we have 600 response times for text entry in each version from all the users, whose average for each version has been shown below. Following the primary paper, outliers with RT greater than 9000ms and less than 250ms have been removed followed by values outside 3 standard deviations from the mean per version per person.

The x-axis shows the task version and all error bars denote standard error. R^2 and RMSD displayed on the graph are for the ACT-R model. Graphs from [Borst & Taatgen, '10] have been borrowed to compare results. Black bars represent experimental measurements while grey bars represent those obtained from ACT-R model. As a general rule, all results of this paper are presented to the left and that of the primary paper to the right.

Experimental **[Borst & Taatgen, '10]**

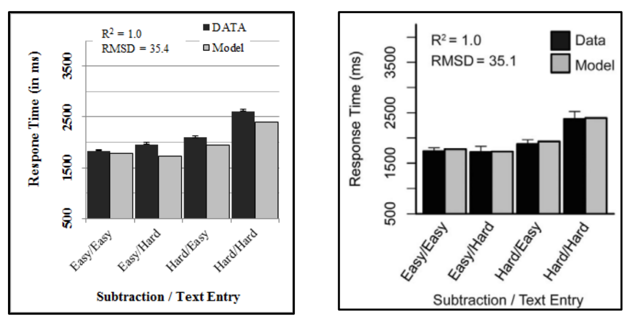


Figure 5. Text Entry Response Time

Comparison with [Borst & Taatgen, '10] :

Except for the easy-easy version all other response times are greater than that observed from the primary paper. One reason could be the increased difficulty level in hard subtraction implemented (negihbouring columns disappear in hard version of the task) , which was suggested but not implemented in the primary paper.

Comparison with ACT-R Model :

The ACT-R model doesn't implement the improved hard subtraction and hence is more in accordance with the primary paper than our results.

A simple consideration of values shows that the response time of increases with task difficulty and there is an overadditive effect in the hard-hard condition. The overadditive effect has been shown diagramatically; only increasing text difficulty adds 262.42ms to the response time and only increasing subtraction difficulty adds 129.39ms, the combined effect is an increase of 775.88ms. This is more than 391.82ms had there been an additive effect, hence, this is an overadditive effect.

Table 2. Average Response Time for text entry (in ms)

Text / Subtraction	Easy	Hard	Difference
Easy	1831.367	2093.79	262.4229
Hard	1960.764	2607.245	
Difference	129.397		775.8776

To test the validity of this overadditive effect we carry out a repeated measure analysis of variances (rANOVA) which gives us the F and p values. rANOVA has been explained in the Appendix. For the text entry response times, interaction effect between text entry difficulty and subtraction difficulty followed by the simple effects are tabulated below along with the values borrowed from [Borst & Taatgen, '10].

Table 3. Effects in Text Entry Response Time – rANOVA

	Experimental			[Borst & Taatgen, '10]		
	F	p	Significant	F	p	Significant
Interaction	13.76	<0.01	Yes	22.15	<0.01	Yes
Hard Subtraction	36.513	<0.01	Yes	10.78	<0.01	Yes
Hard Text Entry	53.69	<0.01	Yes	47.16	<0.01	Yes
Easy Subtraction	4.93	0.03	**Yes**	1.88	0.2	**No**
Easy Text Entry	24.41	<0.01	**Yes**	3.35	0.09	**No**

A p value of 0.05 has been choosen, i.e. only effects with 95% confidence are considered significant. An interaction effect between text entry and subtraction difficulty is found (F = 13.76 & p <0.01) which cofirms the overadditive interaction effect. Simple effect analysis has also been performed, "Hard Subtraction" in the table denotes the effect of text difficulty when subtraction is hard. Tasks with hard subtraction or hard text entry show signifance as expected due to the interaction.

<u>Comparison with [Borst & Taatgen, '10]</u> :

- Given the easy subtraction task, a significance means the mean response time for easy text entry and difficult text entry are different. It is comparable though there is confidence (p = 0.03, 97%) as our significance F = 4.93 is close to the F-critical = 3.83.
- Given the easy text entry task, a significance means the mean text response time of the participants is affected by the subtraction difficulty. (Easy Subtraction:1831ms vs Hard Subtraction: 1960ms). Participants are showing some response to context as well.

Thus, the text response time increases with task difficulty and hard-hard condition shows an overadditive interaction effect.

Similarly, the average response times for subtraction task have been calculated and shown below. It too shows an over-additive effect in the hard-hard condition that is shown below.

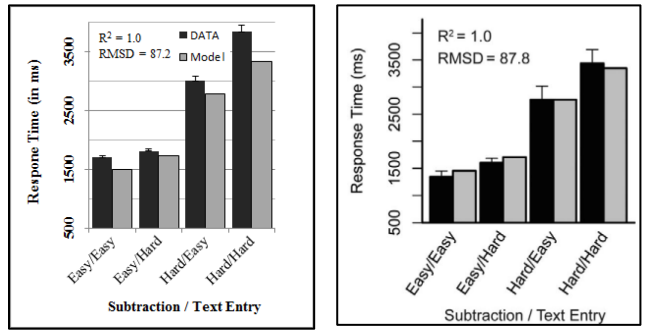
Table 4. Average Response Time for subtraction (in ms)

Text / Subtraction	Easy	Hard	Difference
Easy	1656.66	1741.63	84.97
Hard	2857.64	3422.18	
Difference	1200.976		1765.52

Figure 6. Subtraction Response Time

As mentioned earlier, one reason for the response times being

Experimental **[Borst & Taatgen, '10]**



greater is the increased difficulty of hard subtraction. Another reason being; results from single tasking reveals the average response time for easy subtraction is atlest 1400ms, so, participants in our experiment couldn't have responded below this in the easy-easy version. Again to test the validity of the overadditive effect rANOVA was carried out whose F and p values are shown below.

Table 5. Effects in Subtraction Response Time – rANOVA

	Experimental			[Borst & Taatgen, '10]		
	F	p	Significant	F	p	Significant
Interaction	8.39	<0.01	Yes	6.24	0.03	Yes
Hard Subtraction	12.80	<0.01	Yes	11.81	<0.01	Yes
Hard Text Entry	175.13	<0.01	Yes	111.6	<0.01	Yes
Easy Subtraction	1.85	0.17	**No**	11.65	<0.01	**Yes**
Easy Text Entry	107.63	<0.01	Yes	69.04	0.09	Yes

The significance of the interaction effect (F = 8.39 & p <0.01) confirms the overadditive effect observed.

<u>Comparison with [Borst & Taatgen, '10]</u> :

- Apart from the interaction effect, all simple effects show significance except the case when subtraction is easy. This means the average subtraction response time isn't affected by the text level difficulty. This is a result in our favor because we expect the text level difficulty to be not affecting the subtraction response time when subtraction is easy.
- Moreover as we mentioned earlier, the average response time in single tasking itself is 1400ms whereas results in the primary paper show a subtraction response time of 1400ms in the easy-easy version.
- Like our particpants text response time-given easy text entry, the participants in the primary paper are respoding to the context, for subtraction response time given easy subtraction.

Thus, the subtraction response time increases with task difficulty and hard-hard condition shows an overadditive interaction effect

2.6.2 Accuracy

Accuracy for text entry has been defined as the ratio of number of correct letters entered to 10. The graphs below show accuracy expressed in percentage for text entry vs the task version/difficulty.

Experimental **[Borst & Taatgen, '10]**

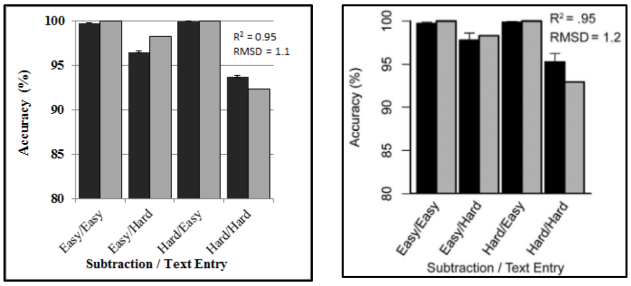


Figure 7. Text Entry Accuracy

As accuracy is derived from discrete qunatites (right or wrong), to calculate the interaction effects the accuracies are transformed using the arcsine transformation [Borst & Taatgen, '10] the results of which are shown below.

Table 6. Effects in Text Entry Accuracy – rANOVA

	Experimental			[Borst & Taatgen, '10]		
	F	**p**	**Significant**	**F**	**p**	**Significant**
Interaction	2.20	0.14	No	4.65	0.052	No
Subtraction	1.33	0.25	No	7.31	0.02	No
Text Entry	16.6	<0.01	Yes	21.57	<0.01	Yes

F = 2.20 & p =0.14 mean no interaction effect between text entry difficulty and subtraction difficulty. With no interaction, the individual effects can be studied separately.

Comparison with [Borst & Taatgen, '10] :

- The results agree with those in the primary paper showing significance only for the text entry task. This means text entry accuracy is affected only by the difficulty of text entry task.
- The primary paper says that the interaction effect between text entry difficulty and subtraction difficulty shows a trend towards significance (p =0.052). This however, is not the case with our results.

Thus, text entry accuracy decreased with text entry difficulty and shows a stronger decrease for the hard-hard condition.

Similarly, the accuracies for subtraction tasks were defined and calculted, whose results are shown below.

Experimental **[Borst & Taatgen, '10]**

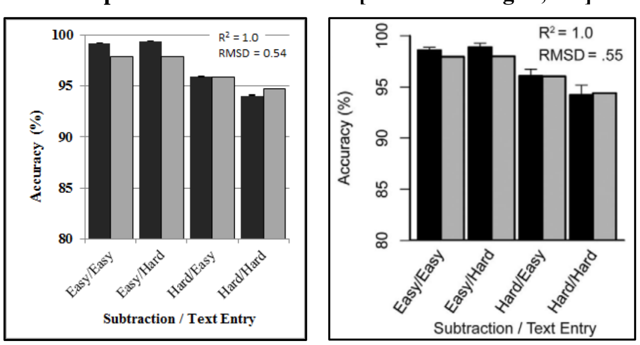


Figure 8. Subtraction Accuracy

To study the interaction effects, again rANOVA is carried out over the arcsine transformed accuracies whose F and p values are shown below.

Table 7. Effects in Subtraction Accuracy – rANOVA

	Experimental		
	F	**p**	**Significant**
Interaction	2.33	0.13	No
Subtraction	35.75	<0.01	Yes
Text Entry	1.26	0.26	No

There was no interaction effect between text entry difficulty and subtraction difficulty in our results. So the individual effects were studied separately. A singificance of only subtraction difficulty means the subtraction accuracy is affected only by subtraction difficulty.

Comparison with [Borst & Taatgen, '10] :

- The primary paper had interaction effects between text entry difficulty and subtraction difficulty, hence, all simple effects had to be considered.

Table 8. Effects in Subtraction Accuracy – [Borst & Taatgen, '10]

	[Borst & Taatgen, '10]		
	F	**p**	**Significant**
Interaction	10.5	<0.01	**Yes**
Hard Subtraction	6.68	0.02	**Yes**
Hard Text Entry	87.7	<0.001	**Yes**
Easy Subtraction	3.64	<0.01	**Yes**
Easy Text Entry	7.17	0.02	**Yes**

- All the simple effects show significance which means the average accuracies are significanlty diffirent among each other across the versions.
- A reason why subtraction accuracy in general was not affected by the text entry difficulty in our results could be that participants tended to place more weightage for accuracy than time and had several re-trials failing the time constraint, sometimes barely finishing in time.

Thus, subtraction accuracy decreased with subtraction difficulty and shows a stronger decrease for the hard-hard condition.

3. DISCUSSION

The interaction effects are in semi-agreement with the model predictions: an over additive effect for response times but not for the error rates. The reason behind this has been mentioned earlier: participant's tendency to be accurate than quick.

The overall increase in response time due to the increased difficulty of the subtraction task couldn't be captured due to the use of the primary paper's original model.

We have also looked at the effect of one condition on other by carrying out the rANOVA analysis. For instance, text response time is affected by subtraction difficulty when text entry is easy.

4. CONCLUSIONS

The over-additive interaction observed in the response times and the results obtained from the computational model indicate well the role of PS as a cognitive bottleneck in multitasking.
The primary paper carries out further experiments to establish this bottleneck is indeed due to PS alone and not cognitive

loading. In a future work, this can be carried out to strongly establish the PS as a bottleneck within threaded cognition.

5. ACKNOWLEDGEMENTS

I thank Prof. Amitabh Mukherjee for providing an avenue through the course SE367 which helped in materializing this project. I want to specially mention Jelmer Borst for making the computational model for the experiment available. I highly approve of Avinash Verma for generously providing me with the monitor and all the subjects for giving me time and bearing with me.

6. REFERENCES

[Borst & Taatgen, '10]
Borst, Jelmer P., Niels A. Taatgen, and Hedderik van Rijn. "The problem state: A cognitive bottleneck in multitasking." Journal of Experimental Psychology: Learning, memory, and cognition 36.2 (2010): 363.

[Borst & Taatgen, '07]
Borst, J. P., and N. A. Taatgen. "The costs of multitasking in threaded cognition." Proceedings of the Eighth International Conference on Cognitive Modeling. 2007.

[Salvuci & Taatgen, '08]
Salvucci, Dario D., and Niels A. Taatgen. "Threaded cognition: an integrated theory of concurrent multitasking." Psychological Review 115.1 (2008): 101

[Salvuci et al., '09]
Salvucci, D. D., Taatgen, N. A., & Borst, J. P. (2009, April). Toward a unified theory of the multitasking continuum: From concurrent performance to task switching, interruption, and resumption. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1819-1828). ACM.

[Anderson, '05]
Anderson, J. R. (2005). Human symbol manipulation within an integrated cognitive architecture. Cognitive science, 29(3), 313-341.

[Anderson, '07]
Mellon, J. R. A. R. K. (2007). How can the human mind occur in the physical universe?. Oxford University Press, USA.

[Baddeley, '00]
Baddeley, A. (2000). The episodic buffer: a new component of working memory?. Trends in cognitive sciences, 4(11), 417-423.

[FitzGerald, '11]
FitzGerald, M. J. T., Gruener, G., & Mtui, E. (2011). Clinical Neuroanatomy and Neuroscience E-Book. Saunders.

[File:Gray726-Brodman.png - Wikimedia Commons, '07]
Gray726-Brodman. (n.d.). In Wikipedia. Retrieved April 19, 2013, from http://en.wikipedia.org/wiki/File:Gray726-Brodman.svg

[Wallis, '06]
Wallis, C. (2006). The multitasking generation. Time Magazine, 167(13), 48-55.

[Research Basics, '11]
Research Basics - What is the Scientific Method? In Explorable, Retrieved April 19, 2013, from http://explorable.com/anova

[Donald, '07]
Handbook of Biological Statistics: Two-way anova (n.d.). In UDEl, Retrieved April 19, 2013, from http://udel.edu/~mcdonald/stattwoway.html

7. APPENDIX

7.1 ANOVA: Analysis of Variances

ANOVA can be used to compare the means of two or more groups. It starts with the null-hypothesis that the means of the all the groups are equal. The variables being controlled are called factors and their possible values are called levels.

For example, in our experiment, the factors are text entry difficulty and subtraction difficulty and the levels are: easy and hard for each factor.

The general aim is to study the effect of a single/set of factors on a measureable (here it is the response time). To establish an effect the means of the two samples should be different: i.e. the rejection of the null hypothesis. To this effect, performing an ANOVA on two groups gives us F and p values for each possible combination of all the levels. (here it is subtraction/text: easy/easy, easy/hard, hard/easy, hard/hard). F is a test statistic for measuring the difference in the means of the two groups, while p is the probability of occurrence of F given the null-hypothesis. So, lower the p value higher the confidence for rejecting the null-hypothesis.

7.2 Two-way ANOVA with replication

Repeated measures design is one where a single participant is subjected to all possible combination of levels in the experiment. This reduces the number of subjects required. As with all other ANOVA's, it assumes "observations within each sample are normal distributed and have equal variances" [Donald, '07]. Two-way ANOVA with replication would have multiple observations for every combination. In our experiments, we have 2 factors, 2 levels and 10 subjects. For 3 subjects, the table would look like:

Table 9. Sample Response Time for 3 subjects

	Text Easy	Text Hard
Subtraction Easy	2069.917	2057.377
	1611.51	1851.554
	1597.752	2027.703
Subtraction Hard	2386.085	2723.486
	1981.418	2240.807
	1946.344	2789.091

The samples from subjects in each block are the replication here. Various statistical tools exist for carrying out ANOVA. Microsoft™ Excel also has this facility for N-way ANOVA with replication under the Data Analysis tab.