EXAMINING THE EFFECTS OF MENTAL WORKLOAD AND TIME PRESSURE ON COGNITIVE TASK PERFORMANCE

An Undergraduate Scholars Thesis

by

MAHMOUD MOHSEN EL-SHERIF

Submitted to Honors and Undergraduate Research
Texas A&M University
in partial fulfillment of the requirements for the designation as

UNDERGRADUATE RESEARCH SCHOLAR

Approved by:
Research Advisor: Thomas Ferris

May 2014

Major: Industrial and Systems Engineering

TABLE OF CONTENTS

ABSTRACT

Examining the combined effects of mental workload and time pressure on cognitive task performance. (May 2014)

Mahmoud Mohsen El-sherif
Department of Industrial and Systems Engineering
Texas A&M University

Research Advisor: Thomas Ferris
Department of Industrial and Systems Engineering

For the proposed research, I will investigate ways to reliably measure the effects of time pressure on mental workload using physiological measurement techniques and subjective workload assessments, as well as the resultant effects on performance in a cognitively-demanding task. This task will involve subjects solving mental arithmetic problems of varying difficulties to induce different workloads. Time pressure has not been systematically researched to the extent of mental workload, but has been observed to affect performance in two contrasting ways: 1) limiting performance in a manner similar to cognitive overload; and by 2) focusing cognitive activities to improve performance. After this experiment we expect to have a better understanding of the interaction between mental workload and time pressure, and the resultant effects on performance. Reliably being able to measure and quantify mental workload and time pressure will allow evaluation of the strain placed on human operators' mental resources when performing all kinds of different tasks with different machines and tools, and can in turn inform the design of tasks and machines/tools to maximize the performance and safety of operations.

ACKNOWLEDGEMENTS

I would like to thank Dr. Ferris for his help, guidance, and mentorship throughout this project. He has also given me the opportunity and freedom to study and work on the projects which I find most interesting and intellectually stimulating.

CHAPTER I

INTRODUCTION

In this current age, many work systems involve human workers performing tasks that are increasingly demanding; requiring interaction with machinery that is growing in complexity and with shrinking margins of error. These tasks range from driving a car and piloting a plane, to welding metals at very high temperatures and operating dangerous machinery in factories and warehouses. Any errors committed during these tasks can result in dire consequences including serious injuries and in some cases even loss of life. Human factors research attempts to understand the task factors that affect humans in order to be able to predict operator performance and provide support in situations where it may be needed.

As technology continues to provide us with devices that are increasingly sophisticated and complex, we must also evolve our understanding of workload and all of the factors that affect it. One of these factors which is relatively underexplored is time pressure, i.e., the stress induced by having limited amounts of time to make a decision or perform a control action. Time pressure tends to interact in interesting ways with both high and low workload levels (Wickens, Hollands, Banbury, & Parasuraman, 2012)[1] in ways that can affect human operators' performance. Some studies have explored time pressure effects on performance and decision making (Ordonez & Benson, 1997)[2] but not to the extent of mental workload. We do know that is interacts with pressure in two different ways: it can sometimes lead to a lower level of performance similar to the effect of increasing mental workload, while at other times it can serve to increase mental focus which can actually improve performance.

Examples of the performance effects of time pressure can be seen in our everyday life. One of the biggest examples can be seen in sports competitions, where across all different sports certain athletes are known to be “closers” or “clutch time players”. These players are at their best towards the end of games, where timing starts to become a major factor. Some athletes make a living playing in these situations where you only have a limited amount of time to score a certain amount of points or prevent a small lead from slipping away.

There are numerous studies that have investigated different mental workload measurement techniques (Mehler, 2012)[3]. Fewer studies have described the effects of time pressure on performance (Andrews & Farris, 1972)[4]. Finally, very few if any research efforts have investigated ways to reliably measure the effects of time pressure. In this study, I will be exploring time pressure effects as well as attempting to find ways to reliably measure the imposed time pressure using established workload measurement techniques in a method similar to (Murray & Janelle 2003)[5] and but with time pressure instead of anxiety levels. I will also explore this interaction of mental workload and time pressure.

Measuring mental workload can be done in three different ways (Tsang & Vidulich, 2006)[6]: performance-based measurement, subjective ratings of workload/effort, and physiological measures that correlate with workload. Each of these methods has been proven to have a correlation with mental workload, but they all also have their own limitations.

The first mentioned method of measuring mental workload is the performance based measurement. This is a very direct method which assumes an inverse relation between workload and performance and uses this relation to predict the exerted mental workload. This method is not always pursuable as sometimes performance can be difficult to manipulate even with an increase in workload, or in our case time pressure. Also the performance of some tasks cannot always be measured. (Tsang & Vidulich, 2006)[6]

Another mental workload measuring technique is the subjective rating of the participants of the task. In this method, participants who perform the task are asked to rate the amount of workload they felt during different times of the task. This can be very useful for researchers because it gives them actual feedback from participants. This may reveal some information that could not have been found using quantitative analysis. However, collecting subjective ratings measures can be disruptive of task performance, and the ratings are subject to inter-subject variability and so can result in data that might be somewhat problematic for statistical analysis.

Finally, physiological measures give an objective quantification of workload which can be easier to analyze in real-time while a task is being conducted (rather than after the fact, as in in performance and subjective ratings methods). Examples of these measurements include heart rate and heart rate variability (Eggert, Lara & Labrador. 2013) [7], and skin conductance (O'Donnell & Eggemeier, 1986)[8]. We used devices to capture some these measurements and use the combination of the resulting data to have a clear picture of the physiological state of the person who is performing the task at all times. However physiological measures are also limited because they are additionally affected by non-workload-related factors such as emotional state and body movement.

No mental workload measuring techniques are faultless and each has its limitations and strengths. This research will use a combination of multiple physiological, subjective and performance techniques to improve the robustness of the overall measurement

.

CHAPTER II

METHODS

Overview

Four study participants were recruited from the student body of Texas A&M University (TAMU) via email. The first 5-10 minutes included an introduction to the task, consenting to participate, and then the participants practiced completing examples of the arithmetic problems at each of the three difficulty levels. This practice session helped that participants get better acquainted with the interface and assured that they understood the format. The participants then started the actual study which consisted of 6 two-minute blocks. The total participation time required for this study was approximately thirty minutes.

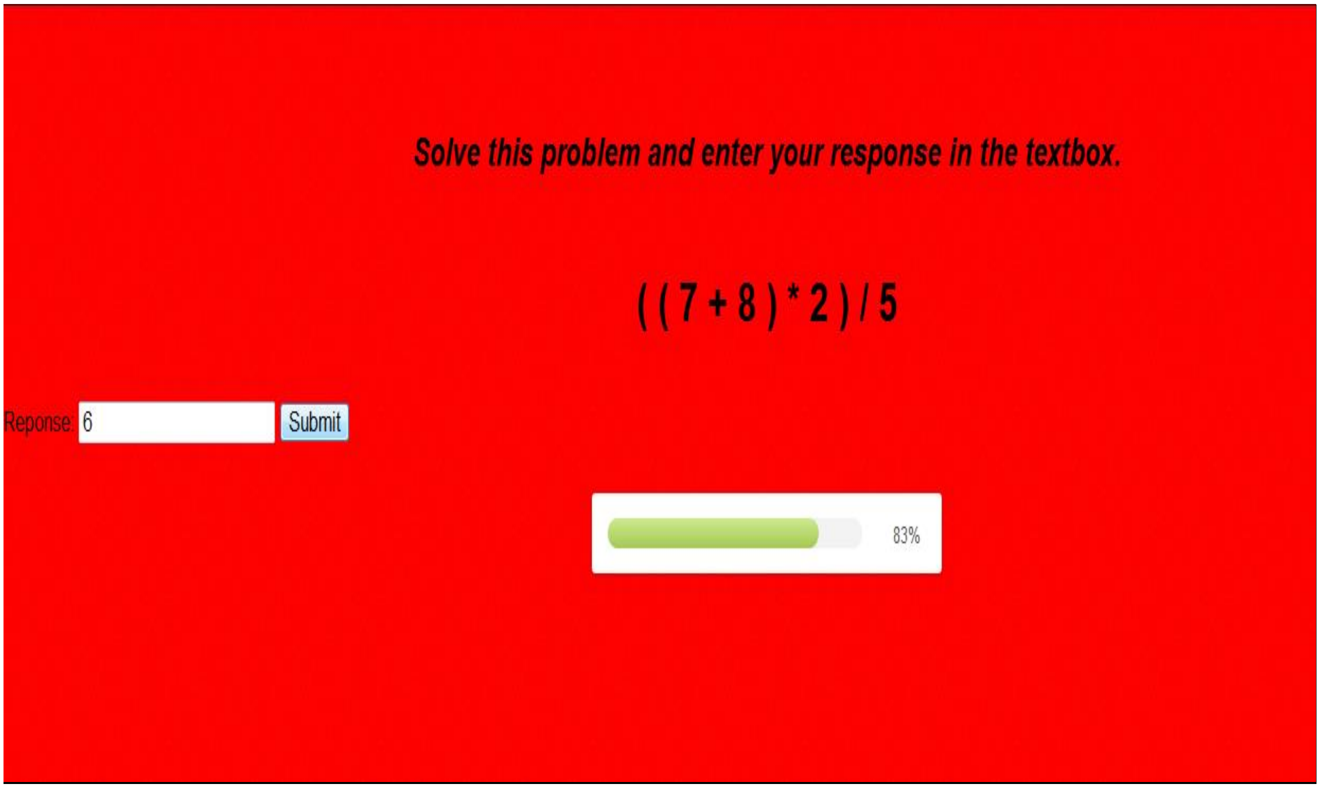
There were 3 different difficulties (D1, D2, and D3) and there were 2 experimental blocks for each difficulty. The order in which these blocks were presented was balanced between participants. Arithmetic problems consisted of one, two or three arithmetic operations involving single-digit numbers. Arithmetic problems of the D1 difficulty level consisted of an arithmetic operation involving one operation (e.g.: 7*4). In the D2 difficulty, the arithmetic problems included an extra arithmetic operation (e.g.: (7*4) +2) and in D3 there was a third arithmetic operation (e.g.: ((7*4) +2) ÷6).

In order to provide the participants with proper motivation to solve as many questions as possible and thus induce real time pressure an incentive was offered. The participants were told they would receive five dollars for their time regardless of performance and an extra five dollars if they were able to solve a certain specified number of problems in the allotted time for three out of the six experimental problems. Every participant was told that he/she successfully completed two of the first four experimental problems presented at the very last second regardless of the correct amount that they had solved at that point. This gave the participants the perception that there was barely enough time allowed, but yet it was possible. The number required to solve was not given to participants, but they were told before each condition that they needed to maintain a high level of performance throughout to have a chance to achieve the bonus. All the participants would receive the full ten dollars regardless of their performance, but the “fake” bonus imposed a heightened degree of time and performance pressure as they completed each experimental condition.

Before the beginning of the experiment, the participants were asked to put on several physiological based mental workload measuring devices. The Zephyr Bioharness3 strap was worn around the chest and collected several sets of physiological data, including heart rate and heart rate variability, via two dry electrodes. The participants were additionally asked to fill out a NASA Task Load Index survey (Hart, & Staveland, 1988)[9] after each experimental block which gave us subjective workload data to compare and contrast with the physiological data.

Display Design

For the experimental blocks that involved arithmetic problems, the problems were displayed at the top of the screen with a textbox under it where the participant's response – entered via a numerical keypad – would appear. After clicking the "submit" button on the screen, the response would be saved (no feedback was given on the correctness of the response) and the next question would appear. A dynamic timer and progress bar showed the estimated time remaining. At the beginning of the time period the background is white, but after one minute and twenty seconds have past and there are 20 seconds left, it turns to red. When there are 10 seconds left, the background starts flashing between red and white. Performance in these last twenty seconds was considered separately from the performance up until that point, in order to compare the performance effects of high time pressure.



The difficulty level and the presence of time pressure (within or outside of the final 20-second time window) were the independent variables of interest for this experiment. The main dependent variables of interest for this experiment were the speed and accuracy of answered questions, as well as the physiological and subjective measures of mental workload. For each dependent variable, we compared the participants' data that included all the independent variables as main effects, as well as their interactions.

CHAPTER III

RESULTS

After the study was performed, all the data was collected. There were two main independent variables: time pressure level, and difficulty. For the time pressure level, each participant's data was split into two different time pressure categories: the no time pressure period or the first 1:20 minutes, and the high time pressure period or the last 20 seconds when the screen turned red and started blinking. There were also three different values for the difficulty variable: D1, D2, and D3.

Because each participant had different baselines and abilities, I used a repeated measures ANOVA test that takes these inter individual differences into account and removes those sources of error. Post-hoc Tukey's tests were used to compare the means for significant effects. In order for a statistic to be considered significant, the required p-level was 0.05.

Performance Based Data Analysis

Time spent on each question:

The first performance metric that was analyzed was the amount of time spent on each question. I calculated the amounts of time the participants took since each question appeared on the screen until they entered the answer in the response box and clicked submit. This was compared across the independent variables (time pressure level, and difficulty)

The amount of time spent on each question was significantly affected by both time pressure (F (1,759) =1.42; p=0.004) and difficulty (F (2,759) =178.48; p<0.001). The interaction effect between time pressure level and difficulty was only marginal (F (2,759) =178.48; p=0.086). This effect would be expected to reach significance with a larger amount of participants.

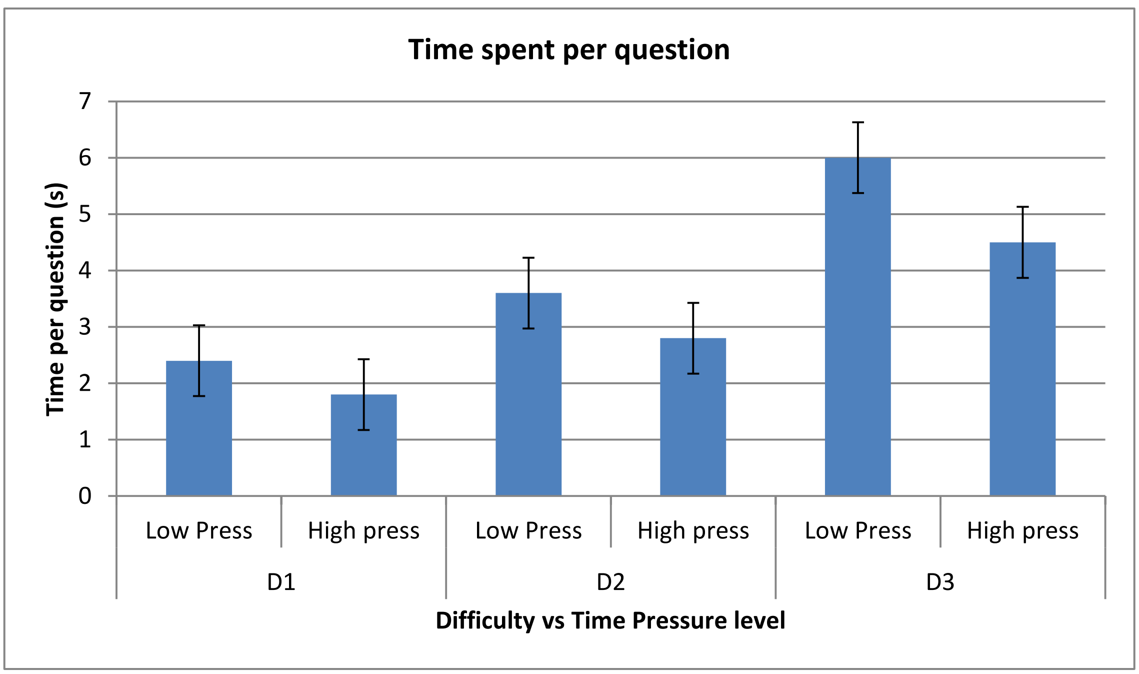


Figure 1: Time spent per question versus time pressure level and difficulty. Error bars represent standard errors

As time pressure was introduced, the time spent per question increased. Post-hoc tests for time-pressure showed significantly less time spent on questions in the high time pressure levels than the lower time pressure levels.

The percentage of correct answers:

This second metric that was tested was the percentage of responses that were correct or simply (correct responses/ total responses). This was compared across the independent variables (time pressure level, and difficulty).

The percentage of correct answers was significantly affected by difficulty ($F(2,759) = 9.29$; $p < 0.001$) but not by time pressure ($F(1,759) = 0.33$; $p = 0.563$). The interaction effect between time pressure level and difficulty was not significant ($F(1,759) = 178.48$; $p = 0.782$).

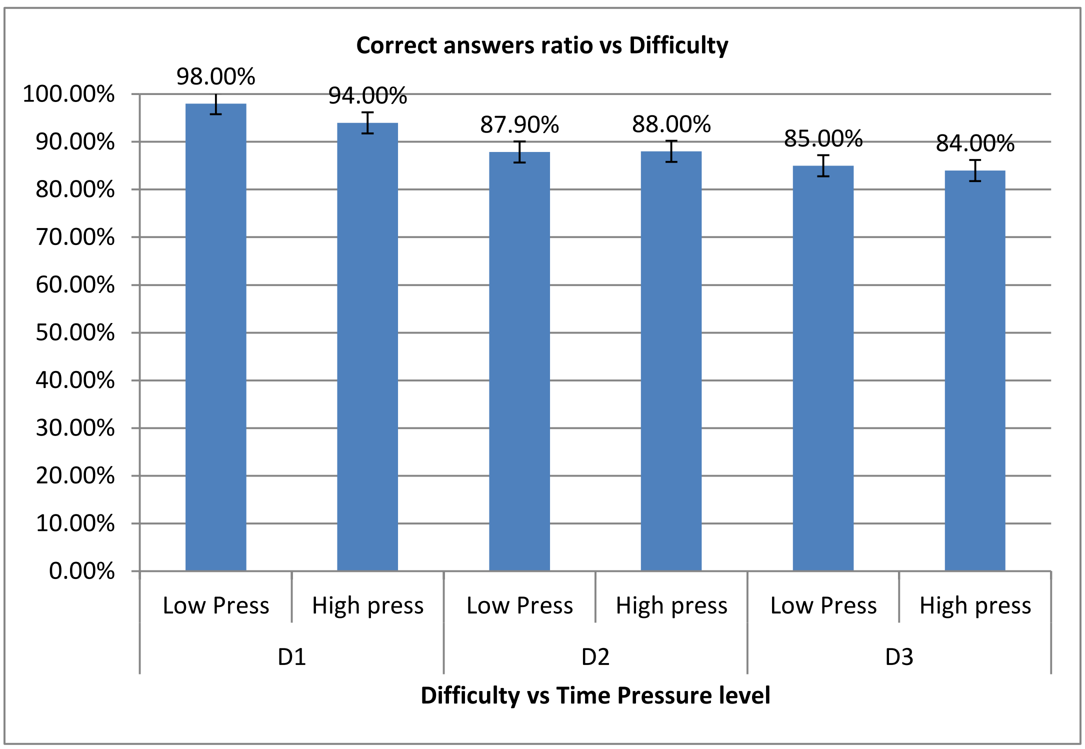


Figure 2: Correct answer ratio versus time pressure level and difficulty. Error bars represent standard errors

Physiological Data Analysis

After the study was completed, I extracted the physiological data from the devices. This data was analyzed similarly to the performance data and the same independent variables were used. For this particular study, only heart rate variability was analyzed but in future follow up studies, skin conductance levels and EEG levels will also be collected and analyzed.

Heart rate variability was not significantly affected by neither difficulty ($F(2, 47) = 0.13$; $p = .719$) or by time pressure ($F(1, 47) = 1.10$; $p = 0.343$).

Subjective Data Analysis

After the participant completed each experimental block, they were asked to fill in a NASA-TLX survey. After the study was completed, their responses were tabulated and key categoreies were then compared across the different difficulties. In future follow up studies, the participants will be asked to take an additional survey that asks for detailed questions about time pressure to give me a subjective means of analyzing the effects of time pressure.

Mental Demand phrased as "How mentally demanding was the task" was significantly affected by the difficulty level ($F(2, 23) = 8.56$; $p = 0.002$). Temporal Demand phrased as "How hurried or rushed was the pace of the task" was significantly affected by the difficulty level ($F(2, 23) = 15.32$; $p < 0.001$).

CHAPTER IV

CONCLUSION

Inducing time pressure affects different tasks in different ways that that are not yet completely understood and whether we intend it or not, time pressure is exerted on us every day. The main goal of this ongoing research is to better understand the effects of time pressure and the first step in doing so is to find ways to reliably quantify time pressure. Mental workload has been well examined and several established methods already exist. The main goal that I was hoping to accomplish in this study was to obtain an initial understanding of how successful the different mental workload measures would be in capturing differences in time pressure.

The results of this study very clearly showed how performance measures offer a promising way of measuring time pressure. There is a clear and strong correlation between time pressure and the time spent on each question. As soon as time pressure was introduced and the background turned red, participants started answering questions at a quicker rate. While the ratio of correct responses to total responses saw no statistically significant change, this was no surprise. This was a relatively simple task and even under a lot of pressure, the participants were still able to solve them correctly. This reinforces the notion that as long as a task is relatively not challenging, inducing time pressure will provide added motivation to the operator, and thus, increasing the rate at which the task is completed without sacrificing performance.

The physiological data did not show any correlations with time pressure. Physiological data tends to be the most difficult data to collect accurately and reliably. When participants yawned or physically moved, that tended to introduce noise into the data and make it more unreliable. Participants would also get acclimated to the task and this would make the time pressure's physiological effect harder to notice. In future follow up studies, more measures will be taken to reduce this noise and more detailed analysis of the data will be done.

While observing the participants, I noticed that participants would almost always start at a relatively quick pace, but eventually fatigue and boredom would set in and the time spent per question would reach its highest point immediately before the time pressure was induced. As soon as the screen turned red, participants would be reminded of the approaching time limit and in turn returned to the quick pace they started off with.

The collected subjective data also seems to be a very promising method of identifying time pressure. Participants seemed to be aware of the induced time pressure and communicated that through the NASA TLX surveys. The temporal demand question had mostly positive responses which showed that the participants truly felt "hurried or rushed". This feeling also significantly increased across the different difficulties.

Because this study only involved 4 participants, it has its obvious limitations. Even though each participant ran 6 experimental trials, having a limited amount of participants led to a lack of a representative picture of the population. It is also important to note that this study only investigated these questions for a very specific type of task at a very easy difficulty. Even the hardest difficulty of the three tested levels is an easy task that most people can do with almost no real difficulty in normal conditions. Follow up studies will be conducted that test different types and structures of tasks to be solved at a wider range of difficulties.

In conclusion, this study showed what potential benefits of understanding time pressure and how it affects the different types of tasks. It can be used to motivate and positively improve performance and understanding it can give us an advantage in designing any user interface. Also the results proved that mental workload measuring techniques can be used to also measure time pressure. The results from this study will be used in future studies where different tasks will be tested in different situations.

REFERENCES

1: Wickens C. D., Hollands J. G., Banbury S., Parasuraman R. (2012). Engineering psychology and human performance (4th ed.). New York, NY: Pearson.

2: Ordonez, L., & Benson III, L. (1997). Decisions under time pressure: How time constraint affects risky decision making. *Organizational Behavior and Human Decision Processes, 71*(2), 121-140.

3: Reimer, B., Mehler, B., Coughlin, J. F., Godfrey, K. M., & Tan, C. (2009, September). An on-road assessment of the impact of cognitive workload on physiological arousal in young adult drivers. In *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 115-118). ACM.

4: Andrews, F. M., & Farris, G. F. (1972). Time pressure and performance of scientists and engineers: A five-year panel study. Organizational Behavior and Human Performance, 8(2), 185-200.

5: Murray, N. P., & Janelle, C. M. (2003). Anxiety and Performance: A Visual Search Examination of the Processing Efficiency Theory. *Journal of Sport & Exercise Psychology*, *25*(2).

6: Tsang, P. S., & Vidulich, M. A. (2006). Mental workload and situation awareness. *Handbook of Human Factors and Ergonomics, Third Edition*, 243-268.

7: Eggert, C., Lara, O. D., & Labrador, M. A. (2013, April). Recognizing mental stress in chess players using vital sign data. In Southeastcon, 2013 Proceedings of IEEE (pp. 1-4). IEEE.

8: O'Donnell, C. R., and Eggemeier, F. T. (1986). Workload assessment methodology. In K. R. Boff, L. Kaufman, & J. P. Thomas (eds.), Handbook of perception and human performance: Vol. II. Cognitive processes and performance (pp. 42.1-42.29). New York: Wiley

9: Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human mental workload, 139-183. Amsterdam, The Netherlands: Elsevier.

CONTACT INFORMATION

Name:	Mahmoud Mohsen El-Sherif
Email Address:	melsheri@tamu.edu
Education:	B.S., Industrial and Systems Engineering, Texas A&M University, May 2012 Undergraduate Research Scholar