

# The Presence of Timers and their Impact on Team Communications During High-Stress Scenarios

Logan Fine

March 18, 2016

## **Abstract**

This research focused on the impact of a timer on the team dynamic in high-stress scenarios. Participants were asked to defuse a virtual bomb while communicating via headset audio. Only one participant could see and interact with the bomb, while the other read and interpreted instructions from the bomb's defusal manual. Results seem to suggest that groups where both participants can see a countdown timer for the bomb perform better than groups where only the bomb defuser could see it, due to higher rates of task completion, completion speed, and more effective communication. There is not enough data to be statistically significant.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Background and Related Work</b>	<b>3</b>
<b>3</b>	<b>Methods</b>	<b>5</b>
<b>4</b>	<b>Results</b>	<b>11</b>
<b>5</b>	<b>Discussions and Future Work</b>	<b>15</b>

# 1 Introduction

Teamwork is prevalent in any working environment, and often can be the deciding factor in success. Teamwork tasks can range from simple day-to-day affairs to life-or-death scenarios where even small mistakes have large consequences. Furthermore, many situations have tight time limitations that can turn even simple tasks into difficult races against the clock. To better understand these situations, it is imperative to know if the time limitations affects not just the task difficulty, but also the quality of teamwork of those attempting to complete it. What can we expect to see happen when teams are put under this pressure?

Previous studies on the effects of timers tend to focus on test-taking environments for students. These standardized tests can heavily affect a student's future, and as a result create a high-pressure time-sensitive environment that produces test scores that can be used as a measure of success. Such studies on learning-disabled students show that extended time increases performance across the board, but showed decreased performance in some areas for non-learning-disabled students [5]. In contrast to this, other studies have shown that non-learning-disabled students can succeed when given extended time limits when they are told to work more slowly and carefully [1]. This would suggest that the timer itself may not be the deciding factor for the success of those students, but instead it is how the timer is presented.

This research focuses on the team aspect of high-pressure timed scenarios. More specifically, this research looks into the communicative aspects of a team to see if the perception of a tight time limit hinders or helps team efficiency and performance. We hypothesized that communication will begin to break down, and errors will increase as inexperienced participants attempt to rush against the clock.

# 2 Background and Related Work

## Experimental Semiotics

The field of experimental semiotics was a large source of information regarding the important elements of interpersonal task performance. Relevant studies primarily focused on shared visual workspaces and the subsequent effects on team performance. In several studies, it has been shown that sharing a workspace either visually or physically, helps teams work together by facilitating the effectiveness of communication.

The shared workspaces in the studies provided an easier means for information to be passed along between participants, and ultimately made all parties involved more efficient. [3, 2, 8, 7] The benefits come largely from the additional visual feedback on top of audio communications in order to better establish conversational grounding, which allows any workers involved to better understand the problem at hand, as well as to better portray how each worker aims to proceed with solving a task.

## **Timing Students**

To better understand the effects of timers on task performance, research was done on the effects of timing on students in standardized test environments. Many studies tended to focus on the results for learning-disabled (LD) students, however, non-learning-disabled (NLD) students were often included as well.

Ofiesh et al.[5] conducted a study on college students taking a test to assess vocabulary and comprehension. Most of LD students tested improved their scores, while the same could not be said for their NLD counterparts. On younger age groups from 1st grade to High school, studies conducted by Brooks et al.[1] showed that NLD students have either no benefit from extra time, or had it negatively impact their performance.

Further studies on 6th grade students by Huesmen et al.[4] showed the opposite affect, but only when the specific directions given to students taking the test differed between groups. Students told to take their time along with extended time improved their scores noticeably, whereas their counterpart told to work quickly did not have any significant improvement. On top of this, a study by Powers et. al.[6] on graduate students taking the GRE writing test showed improved scores for those given 60 minutes over those who had the 40 minute alternative.

## **Current Study**

Because the studies reached contradictory conclusion, it may be possible that the timer itself was not a deciding factor in test performances, but rather the presentation of the scenario and the time limit that impacted student behavior and testing capabilities. To learn about team dynamics and creating effective team settings, it was appropriate for us to look into whether timers in a team setting would have similar

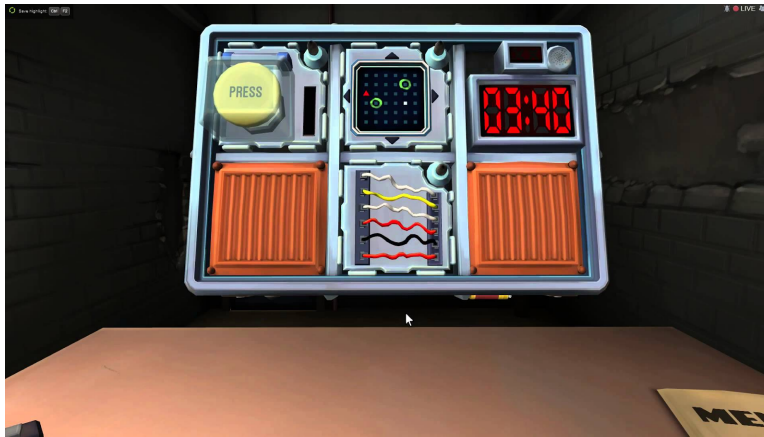


Figure 1: The Virtual Bomb in, "Keep Talking and Nobody Explodes"

effects. In order to investigate this, a new experiment based around a team dynamic in a stressful scenario was necessary.

### 3 Methods

The ideal experimental setting required a task that could be completed by two or more participants who would act as a team. The task would also need to have easily measurable checkpoints towards success as well as be under timed conditions such that participants would feel minor levels of stress when attempting to solve it.

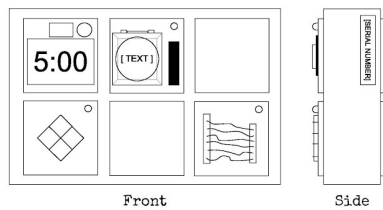
The game, "Keep Talking and Nobody Explodes," was chosen for the experiment. It provided a team scenario where players would need to communicate under time-pressure in order to solve somewhat complex tasks to defuse a virtual bomb. The bomb also provided a way to measure success, as it is defused by solving small puzzles on the bomb's face using directions taken from a manual. However many puzzles that were solved in working towards complete defusal were a measurable indicator of how close participants were to finishing the task and winning the game.

Participants were given a short questionnaire asking if they have had prior experience with games, or if they have previously worked in any team-oriented settings with each other, and what team related

## Defusing Bombs

A bomb will explode when its countdown timer reaches 0:00 or when too many strikes have been recorded. The only way to defuse a bomb is to disarm all of its modules before its countdown timer expires.

Example Bomb



### Modules

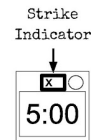
Each bomb will include up to 11 modules that must be disarmed. Each module is discrete and can be disarmed in any order.

Instructions for disarming modules can be found in Section 1. "Needy" modules present a special case and are described in Section 2.

### Strikes

When the Defuser makes a mistake the bomb will record a strike which will be displayed on the indicator above the countdown timer. Bombs with a strike indicator will explode upon the third strike. The timer will begin to count down faster after a strike has been recorded.

If no strike indicator is present above the countdown timer, the bomb will explode upon the first strike, leaving no room for error.



### Gathering Information

Some disarming instructions will require specific information about the bomb, such as the serial number. This type of information can typically be found on the top, bottom, or sides of the bomb casing. See Appendix A, B, and C for identification instructions that will be useful in disarming certain modules.

Figure 2: An excerpt from the defusal manual from, "Keep Talking and Nobody Explodes"

experiences they have had in the past. This aims to identify any validity threats prior to the experiment being run.

One participant played the role of the "defuser," and was tasked with interacting with the virtual bomb in order to defuse it. This participant had access only to the computer to play the game, and a headset to communicate with the other participant.

The other participant was the role of the "bomb expert," and was tasked with reading and understanding a manual that explains how to solve the bomb's puzzles to defuse it. This participant only had access to the defusal manual and the headset used to communicate with their teammate. They were not able to see the bomb.

It was expected that one participant will take a leader-role throughout the duration of this experiment, as success relied heavily on the participant's ability to communicate what needed to be relayed to him/or, and what the defuser needed to do in response to events happening in the game.

## **Participants**

There were 8 groups of two union college students for this experiment. Each participant pair had experience working with one another in team-related tasks, and had experience with games prior to the experiment. However, no participants had any experience with "Keep Talking and Nobody Explodes."

## **Treatments**

- The first treatment was set up such that both bomb defuser and bomb expert saw a timer that accurately depicted the time remaining left on the bomb before detonation. This would be the only information that both participants would have access to without needing to communicate about it.
- The second treatment was set up such that only the bomb defuser had access to the timer. This would be a scenario where additional information would need to be passed along between participants, and may impact the performance of the bomb expert.

Because these two treatments had only the expert's timer as a variable, it allows for examination of the teams' performance quality under different time pressures. We hypothesized that the second treatment's setup will yield a lower success rate in terms of completing the task. This would be due to the possibility that communication about the timer may slow down the solution process, or may be ignored completely, leading to the bomb expert moving too slowly in providing directions to the defuser. It is possible, however, that being unaware of the time restriction may allow the bomb expert to be more relaxed, and may result in better, more efficient performance overall.

## **Task**

The task was for both participants to work and coordinate together to successfully solve all puzzles on the virtual bomb within the allotted time. Puzzles appeared on the bomb as individual panels that fill one out of six square slots on the face of the bomb. Potential puzzles that were on the bomb are as follows:

- Wires: The defuser had to cut one correct wire on the bomb that depends on what wires already appear on the bomb.
- The Button: The defuser had to hold a button on the bomb for a specific amount of time depending on what physical attributes the bomb and button has.
- Keypads: The defuser had to press the keypads in a specific order depending on what symbols appear on the keys themselves.
- Simon Says: The defuser had to press flashing buttons in a specific order depending on what buttons flash when
- Who's on first: The defuser had to press a button with a certain word depending on what word is displayed on the puzzle's screen
- Memory: The defuser had to press a specific button pattern depending on what appears on the puzzle's display screen



Figure 3: The Bomb displays the time, task completion, and strikes accumulated

These will be the only puzzles that can appear on the bomb during this experiment. Each group will be given the same set of puzzles to eliminate any potential validity threats that would arise otherwise.

## Procedure

Participants were invited into the Union College Human Computer Interaction Lab. Here, the subjects were separated into separate rooms, and asked to fill out questionnaires regarding details of their previous experience with each other, games, and team-related activities. After this, each participant were read a prompt that explains that they will be defusing a virtual bomb, and what their respective roles are between defusing and reading from the manual. Following these briefings, a timer was started by the experimenter and the participants began their task. Video of the game itself as well as audio of the two participants were recorded. After the task was finished, the students were debriefed on the experiment and let go from the lab.

## Measurements

To determine whether or not the timer had an effect on the participants, we looked at several different aspects of the performances. Measurements were chosen in order to best assess overall success in task completion, success in communication of information, participant efficiency, and what type of team dynamic

was established during the test. Only the footage of the game screen and audio from the conversation of the participants was recorded. What was recorded and noted from the experiment were as follows:

- From the game:
  - Mistakes made in attempting to solve a puzzle were taken from in-game, “strikes” on the bomb. These strikes appeared on the bomb when the defuser clicked on a wrong answer on any of the given puzzles. This measurement was used to measure the efficiency of teams who participated, as mistakes could have been made from either end. The defuser could have misclicked, misheard the expert and accidentally clicked the wrong choice, or the expert may have given incorrect information to the defuser, leading to the wrong choice.
  - The total number of puzzles solved throughout the task were recorded as a measurement of how close the team came to completing the task. Because each test group had the same number of puzzles to solve in order to defuse the bomb, measuring this allowed us to note and categorize levels of success between teams.
  - The time in which it took to solve individual puzzles were also recorded, as it provided a measurement that can be applied to groups who complete puzzles. For example, some teams may have only completed two out of three puzzles, but they may have completed those two quickly and efficiently whereas a group with the same amount of puzzles solved may have used the entire time to do so.
- From the audio:
  - The total number of words each participant spoke in conversation was recorded to look into the dynamic between speakers. This was used to see the conversational dynamic between defuser and manual expert. It was of note to see if one participant type spoke more or less between the different treatments, as there may be behavior changes.
  - The number of Conversational turns between participants was measured to assess the amount of times information was exchanged between two groups. Changes in this behavior between

treatment groups would indicate that the added timer has some affect on the dynamic between participants.

These measurements will be tallied to see if there is any significant differences in totals between the two treatments.

## 4 Results

Because of the small sample size, the results found were not enough to provide statistical significance. There were, however, noticeable differences between the performance of the two groups, as well as performances between sessions.

Between both participants, the amount of turns taken remained roughly the same, as the conversation would usually continue back and forth with no additional interruptions. In session one, the average turns for the one-timer groups was 34.75 turns. The average turns for the two-timer groups was 42.25 turns. In session two, the one-timer groups took on average 28 Turns, and the two-timer groups took an average of 40 turns.

The calculated average Word ratio between groups showed no noticeable difference between sessions 1 and 2 of the two-timer groups, where the expert spoke more words at roughly a 9:8 average word ratio. The one-timer groups, however, had a drastic ratio change. In session 1, the ratio between expert and defuser was 14:10, but changed to 8:10 in favor of the defuser.

The total tasks completed were in favor of the two-timer groups, as they on average completed 1.5 tasks between both sessions, while the one-timer groups completed 1.1 tasks on average. It is also important to note that in the one-timer group, some participant groups solved no tasks at all in some sessions, whereas the two-timer groups always solved at least 1 task. The two-timer groups also had the only participant pair who successfully completed three tasks and disarmed the bomb.

Average Task Completion Time In terms of speed, the two timer groups also outperformed their counterparts, with an average task solution time of 97.5 seconds between both sessions, while the one-timer groups took on average 121 seconds.

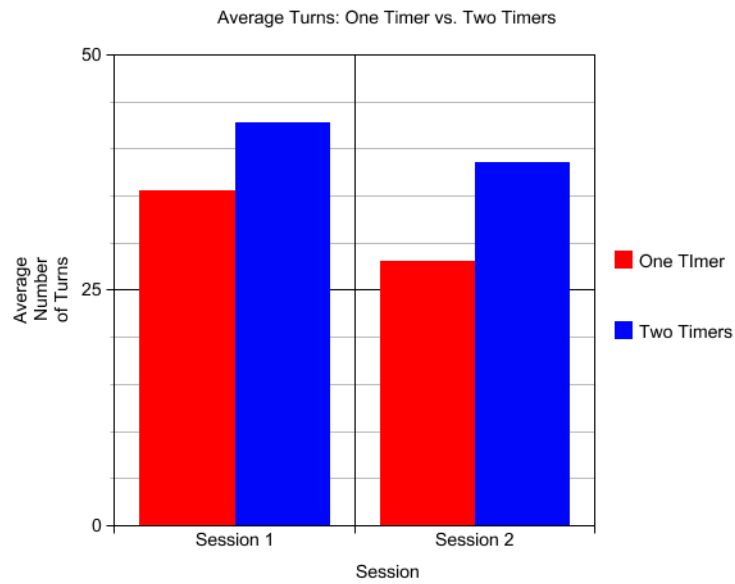


Figure 4: Participants in the two-timer groups take more turns in both sessions

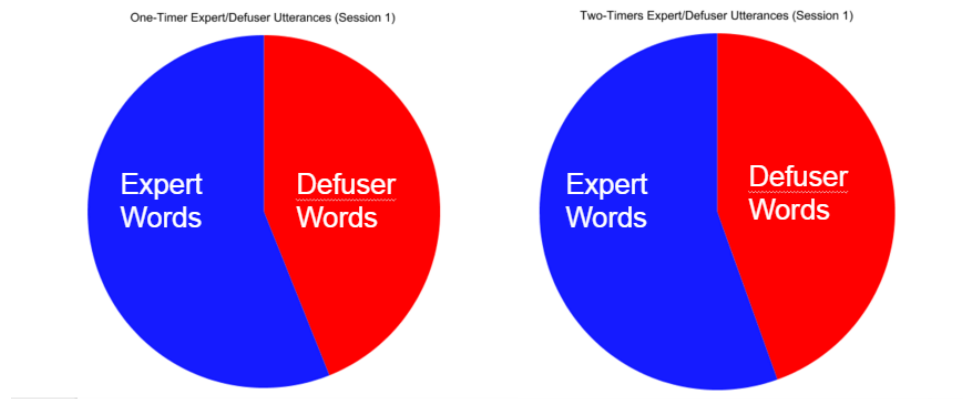


Figure 5: Word count ratios are roughly the same between groups during session 1



Figure 6: During session 2, the defusers of the one-timer groups spoke more than the expert

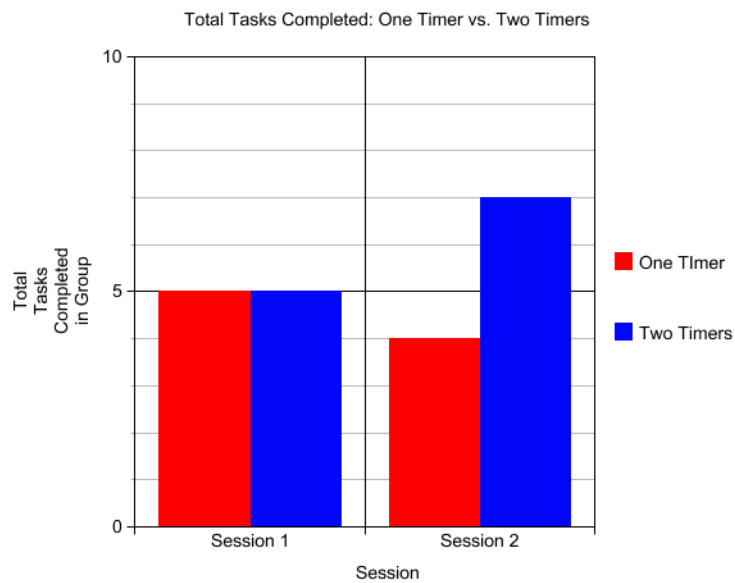


Figure 7: Participants in the two-timer groups completed more tasks overall than the one-timer groups

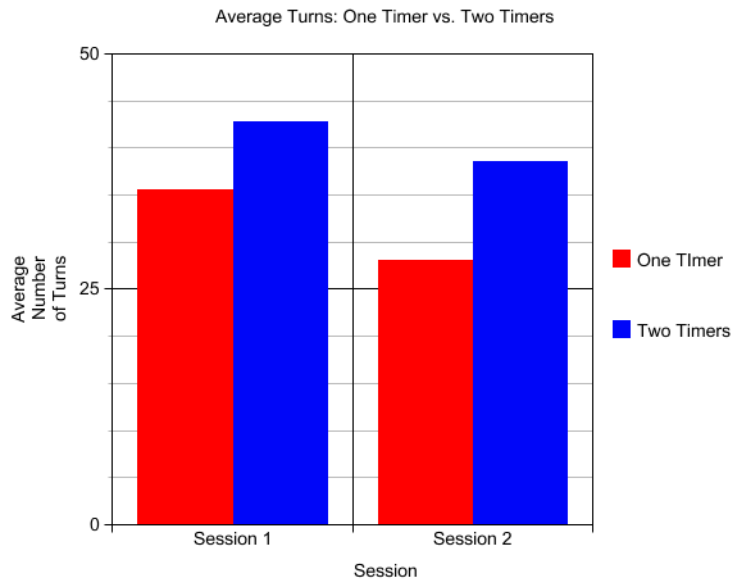


Figure 8: Participants in the two-timer groups performed faster on average than the one-timer groups

Both groups received on average about the same amount of strikes. The two-timer groups had 1.5 strikes on average during session one, and increased to 2 strikes for the second session. The one-timer groups had 1 strike on average during session 1, and also increased to average 2 strikes during session 2.

## Observations

:

Other non-measurable observations were made during the proceedings of the experiment. Many participant groups carried out explicit confirmation between one another, where one participant would describe something, for example, "There are three red wires," and in order to clarify and ensure understanding, the other participant would repeat "There are three red wires," and then continue on with whatever else they had to say. This short exchange of information allowed for more clarity, as group members were able to ensure that what was being discussed was understood by both parties.

## 5 Discussions and Future Work

In regards to the results, it would appear that the addition of the second timer does aid in task performance. The additional information provided to the expert allows for them to better organize themselves and take a more active role in getting the bomb defused. Though both one-timer and two-timer groups would communicate about the time left, groups with no timer for the expert ran into issues where the expert was unaware that they were moving too slowly. This may also explain why the defusers in the one-timer groups began to speak more than the experts in session 2, as they may have realized that they needed to better control the workflow since they were the most aware of the time. There was one participant pair in particular that allowed time to run out without even picking up the bomb, with the expert stating that he was unaware that the time was already ticking.

In order to better understand this however, it would be necessary to continue running the experiment to get a larger sample size. As far as the data is concerned, it is entirely possible that the apparent trends are random and do not accurately represent any differences between the groups. A further threat to the validity of the experiment is the relationship between participant pairs. It may be the case that certain teams performed better simply because they work better as a team than others. The only participant pair to solve all three tasks and defuse the bomb stated that they have been close friends since childhood. Whether or not it is directly related to their success, it may be an indicator that there needs to be more rigidity in participant selection in future works.

Another route to follow would be to remodel this experiment with a third treatment, where the participants have no timers at all. One of the most common items mentioned between participants in their dialogues was the exact amount of time remaining. Without this information available, what would take the place of this dialogue, and how would teams organize themselves to take on the bomb?

## References

- [1] Case B. J. Brooks, T. E. and M. J. Young. Timed versus untimed testing conditions and student performance. Technical report, Pearson Education, Inc, 2003.

- [2] Herbert H. Clark and Susan E. Brennan. Grounding in communication. In *Perspective on Socially Shared Cognition*, 1991.
- [3] Gary M. Olson Xiaolan Fu Elizabeth S. Veinott, Judith Olson. Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1999.
- [4] Jr. & Frisbie D. A. Huesman, R. L. The validity of the itbs reading comprehension test scores for learning disabled and non learning disabled students under extended-time conditions. Paper presented at the Annual Meeting of the American Education Research Association, New Orleans LA., 2000.
- [5] N. S. Ofiesh. Using processing speed tests to predict the benefit of extended test time for university students with learning disabilities. *Journal of PostSecondary Education and Disability*, pages 39–56, 14.
- [6] & Fowles M. E. Powers, D. E. Effects of applying different time limits to a proposed gre writing test. *Journal of Educational Measurement*, pages 433–452, 1996.
- [7] Susan R. Fussell Robert E. Kraut, Darren Gergle. The use of visual information in shared visual spaces: Informing the development of virtual co-presence. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, 2002.
- [8] Jane Siegel Susan R. Fussell, Robert E. Kraut. Coordination of communication: Effects of shared visual context on collaborative work. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, 2000.