1 ABSTRACT
This study aims to determine the effects of stress on visual search performance, specifically accuracy, and associated eye movements, saccades and fixations, obtained via eye-tracking. Twenty-one participants performed ten Spot the Difference tasks, randomized to either the intervention condition, with a time pressure stress manipulation of a visible timer, or a control condition. Eye-tracking data and accuracy reports were collected through Gazepoint Analysis and manually by researchers. Contrary to prior research, there was no difference in saccade or fixation frequency or accuracy in finding the differences between the Spot the Difference images. A variety of correlations were discovered from post-assessment analyses including strong relationships between enjoyment of the task and ignoring the timer or time limit, as well as, feelings of anxiety and the timer being viewed as helpful. The results of this experiment could be used to improve the user interface and design of computer gaming systems utilizing time pressure and activities that require a time limit, as well as to gain insight into the opinions regarding how a visual timer affects enjoyment, anxiety, and stress.

2 INTRODUCTION
Spot the Difference images have persisted as a common pastime for children and adults in newspapers and magazines. Often used for entertainment, this brain twister offers a medium to analyze visual search under a stress-induced time constraint. Students across the globe take online tests at various academic levels, such as subject tests in high school, entrance exams for college like the ACT, or specialized tests for higher education, such as the GRE or MCAT. These examinations often present a timer counting down the seconds before the student must submit their work. Additionally, workers in high pressure professions often feel time constraints or pressure to complete their assignments and duties. Workers in health care or emergency responders have to be able to assess their surroundings and isolate changes, such as a police officer clearing a house or Navy Seals watching their monitors for changes in sonar. With the addition of eye-tracking analysis, a measurement of fixation duration and saccades estimates if the addition of the timer provides a more stressful environment and distraction from the actual work. This paper explores how time pressure affects performance on the Spot the Difference visual search task and discuss time pressure’s effect on fixation duration and saccades. The goal of this paper is to illustrate the dampening effects of stress-inducing time pressure on accuracy in a testing scenario. We hypothesize that the addition of time pressure will increase the number of saccades, decrease the number of fixations, and negatively impact performance on the Spot the Difference task.
2.1 Background
Spot the Difference games double as an entertaining activity while waiting to be called to see the doctor and as a cognitive examination of visual search used to survey differences in change blindness. Change blindness is a phenomenon in which the human brain fails to notice a change, a phenomenon which occurs primarily in the dorsolateral prefrontal cortex, parietal cortex, and ventral visual pathway[2]. Examinations of change blindness have primarily been researched in terms of cognitive decline as a result of age, as with neuro-developmental disorders such as Alzheimer’s disease [17], and attention [18]. Spot the difference tasks are increasingly difficult since the observer does not know specifically what elements will be altered [18]. Additionally, while the observer may be paying attention to multiple components at once, the human brain only has the capacity to observe one change, a phenomenon called “change simultagnosia” [16].

With the increased difficulty resulting from change simultagnosia and other cognitive restrictions, time pressure further impairs the observer’s ability to accomplish their goal [16]. Time pressure is a distraction technique which increases stress by pressuring the participant to strive for speed not accuracy, thus moving their point from an ideal spot along the speed-accuracy operating characteristic [11][14]. Time pressure has been shown to impair decision making [3], inhibition accuracy [5], and face matching accuracy [6], as well as, having a dramatic effect on saccades and fixations, specifically an increase in saccade rate and decrease in saccade duration and fixation duration [1] [4] [12] [15] [19]. In addition, the combination of time pressure and clutter during a visual search has shown detrimental effects on reaction time and error rate [13].

3 METHODOLOGY
3.1 Apparatus
The program was displayed on a 22-inch desktop monitor with a resolution of 1680 x 1050. The computer used was a Dell P2213 with a refresh rate of 60 Hz. A Gazepoint GP3 eye tracker (sampling rate: 60Hz; accuracy: 0.5-1 degree), mounted at the bottom of the desktop monitor, was utilized in this experiment. The participants were seated at a distance of around 65 cm from the eye tracker. No chin rest was used; however, participants were asked to keep their heads steady throughout the experiment. A mouse and keyboard were also provided.

3.2 Stimuli
Ten Spot the Difference images were collected from an American YouTube channel called Reystar Brain Games [7][8][9][10]. All the images are real-life, color photos that have been altered to contain minor differences. Each image was converted to a resolution of 1480 x 850, to leave room for the visual timer at the top, and displayed across the center of the screen. A total of 3 differences, spread throughout the background, minor details, and main images, could be found in all ten images. Two examples of Spot the Difference images can be seen in Figure 2 and Figure 3.

Both the intervention and control used the same order of Spot the Difference images. The sole difference was the intervention condition used a visible timer at the top-middle of the screen between the two images. The timer was a rectangle with a black background and white digital numbers counting down from forty to zero. The control condition did not contain the timer. The timer can be seen in Figure 2 and Figure 3.

3.3 Participants
A total of 21 college students were recruited for this study. Participants were contacted in person or through email and willingly agreed to partake in the study. Additionally, participants were blinded to the condition and purpose of the study to prevent the Hawthorne Effect. All participants had self-reported normal or corrected to normal vision. This study was approved by the university’s Institutional Review Board.

3.4 Procedure
Participants were first randomly assigned to either the intervention or control condition; the randomization list was provided by an online generator. Participants were then seated in front of the desktop and asked to complete a short pre-assessment questionnaire collecting general demographics, gender and age, and information regarding their eyesight. After completion of the pre-assessment, participants were asked for verbal consent to participate in the study. Once consent was verbally given, participants were instructed to complete a short calibration and validation activity through Gaze
Point Analysis using a five-point calibration grid. After calibration, participants were reminded to keep their head position as still as possible without causing discomfort. Participants were then informed that they would complete ten Spot the Difference tasks with each image containing three differences. Additionally, they were encouraged to locate all three differences in each photo quickly and informed of their 40 second time limit. In the stress condition, an additional instruction notified them that a timer at the top-middle of the screen would remind them of their remaining time. Once a difference was discovered, they were instructed to point at the difference with their right hand using their pointer finger. Any questions regarding instructions were answered, and the participants began the tasks.

After the conclusion of the final Spot the Difference image, participants completed a post-assessment questionnaire regarding their stress and anxiety levels, difficulty and comprehension of the task, and experience with the Spot the Difference tasks. Participants in the stress condition received questions regarding the timer, while participants in the control condition received questions about the time limit. Once the post-assessment questionnaire was completed, the participants were thanked and given the researcher’s contact information in case there were questions or concerns.

3.5 Experimental Design

This study was an experimental design investigating the effect of a two-level stress variable, timer condition vs. control, on accuracy, number of saccades, and number of fixations. Accuracy was determined based on how many differences were correctly identified for each of the ten images. Extended fixations over the differences in addition to verbal confirmation and pointing from the participant to the difference on the computer monitor would constitute an accurate acknowledgment that it was located. Utilizing the Gaze-point software, saccades and fixations were recorded, stored, and exported through HDFView. Data extraction and statistical analyses were performed using SPSS.

4 RESULTS

While a total of 21 participants were recruited for this study, only 20 participants were analyzed due to incomplete or corrupted data (7 females and 13 males; \(M_{age} = 23.7; SD_{age} = 2.70\)). An independent samples t-test was run through the statistical program SPSS. When analyzing the data, none of the predicted hypotheses proved to be statistically significant. Regarding the accuracy in finding the differences between the two pictures, there was no statistical difference between the two conditions, meaning both groups shared similar success, finding around 1-2 differences per image (\(M_{intervention} = 0.60; SD_{intervention} = 0.29; M_{control} = 0.59; SD_{control} = 0.28\). \(t(18) = 0.14; p = 0.89\)). See Figure 4 for a visualization of these results.

In terms of the number of saccades, the stress condition produced a slightly higher number of saccades (\(M = 106.3; SD = 3.7\)), but not enough show a significant difference from the control condition (\(M = 104.3; SD = 14.00\). \(t(18) = -0.41, p = 0.69\)). This suggests that the addition of a timer did not increase the number of times participants increased the number of eye movements as predicted. See Figure 5 for a visualization of these results.

With the number of fixations within the areas of interest (AOIs), the intervention condition produced a slightly higher average (\(M = 5.40; SD = 2.10\)) compared to the control condition (\(M = 4.67; SD = 1.63\)), but no statistically significant difference was discovered (\(t(68) = -1.58, p = 0.12\)). Additionally, no significant differences were discovered between the average number of fixations for each image, (\(t(18) = -1.05, p = 0.31\)). See Figure 6 and Figure 7 for visualization of these results.
It was also discovered that there was no statistical difference in terms of stress ($M_{\text{intervention}} = 3.44; SD_{\text{intervention}} = 1.24; M_{\text{control}} = 2.90; SD_{\text{control}} = 1.45$) or anxiety ($M_{\text{intervention}} = 3.33; SD_{\text{intervention}} = 1.32; M_{\text{control}} = 2.80; SD_{\text{control}} = 1.32$) in both the timer ($t(18) = 1.05; p=0.31$) and control conditions ($t(18) = 0.72, p=0.48$). While participants in the timer condition reported that the time limit made the task more difficult ($M=1.89$) and less helpful ($M=2.22$) than the control condition ($M=2.09; M=2.45$), neither statement reaches significant levels between conditions ($p=0.586; p=0.632$). See Figure 8 for a visualization of these results.

In order to address a floor effect, participants were asked to rate the difficulty of the 10 spot the difference images. Overall, most participants viewed all images as very difficult or difficult with the no image averaging above "neither easy nor difficult". The most difficult image, seen in Figure 2, had an average rate of 0.35 (finding one difference out of three), and the easiest image, seen in Figure 3, had an average of 0.77 (consistently finding at least two out of three differences). In terms of conditions, the average difficulty rating of spotting the differences were slightly higher in the control condition ($M = 2.42; SD = 0.52$) compared to the intervention condition ($M = 2.27; SD = 0.37$), however, no statistical differences emerged between the two groups ($t(17) = -0.71, p = 0.48$). See Figure 9 for a visualization of these results.

Additional correlations were performed to assess whether there was a relationship between the participant’s opinions on the task, such as the task was hard to understand, the timer made them anxious or stressed, the timer was distracting, the timer made the task easier, the timer was helpful, and whether or not participants ignored the timer. Additionally, a correlation between number of fixations on the timer and difficulty ratings of the task was performed. Strong positive correlations were found between the following:

- The participant’s level of enjoyment and how much they ignored the timer or time limit ($r=1; p<0.000$).
- The participant’s opinion that the task was hard to understand and the timer made them anxious ($r=-0.69, p<0.000$) and stressed ($r=0.85, p<0.000$).
- The participant’s opinion that the timer or time limit was stressful and the opinion that the timer or time limit made them anxious ($r = 0.69, p = 0.001$).
- The participant’s opinion that the timer was distracting and viewing the task as easy ($r = 0.57, p = 0.008$).

Additionally, a strong negative correlation was discovered between the following:

- The opinion that the task was hard to understand and the timer was helpful ($r=-0.69, p = 0.001$).
- The opinion that the task was hard to understand and the timer was helpful ($r=-0.69, p = 0.001$).
- The timer implemented feelings of anxiety and the timer was helpful ($r=-0.69, p = 0.001$).
- The timer implemented feeling of stress and the timer was helpful ($r=-0.69, p = 0.001$).

Lastly there were moderate positive correlations between:

- The task being easy and participants ignoring the timer ($r = 0.46, p = 0.04$).
- The task being easy and the enjoyment of the task that also reached significance ($r = 0.46, p = 0.04$).
- The number of fixations on the timer in the ten Spot the Difference images and the difficulty of finding differences in the images in the intervention condition ($r = 0.65, p = 0.04$).

See Figure 10 and Figure 11 for the correlation tables.
5 DISCUSSION

Overall, the expectation that time pressure would impact the participant’s accuracy, number of saccades, and number of fixations proved inaccurate. Both the timer condition and the control condition experienced similar levels of accuracy when trying to find the three differences in the ten different spot the difference images. In addition, the length of fixations and number of saccades also did not significantly differ between the two conditions. While our results differ from those previously mentioned and found in prior research [1] [4] [12] [15] [19], this suggests the possibility that the addition of a timer does not further add pressure to situations where neither reward nor punishment are given at the end. The strong relationship between enjoyment of the task and ignoring the timer suggests that as participants felt more immersed in the task having fun, the less they consulted to the timer or thought about the time limit. Additionally, the relationship between the number of fixations on the timer in the ten Spot the Difference images and the difficulty of finding differences in the images in the intervention condition suggests that participants glanced at the timer in the top-middle of the screen more often with images where they had more difficulty finding differences.

5.1 Limitations

This study contained several limitations. Firstly, the small, unbalanced sample size ($N_{\text{intervention}} = 9$, $N_{\text{control}} = 11$) could have reduced the power and affected the significance of the findings. Secondly, all participants verbally stated that the find the Spot the Difference puzzles were difficult. In addition to these verbal ques, the data showed that none of the participants found all three differences in all the images. Of all the trials, two participants, one in the control condition and one in the intervention condition, tied for most differences spotted. These participants identified all three differences in five images out of ten ($M_{P7} = 0.73$, $M_{P10} = 0.77$), which suggests a floor effect. Additionally, in order to ascertain accuracy, researchers stood behind or to the side of the participant. Because of this data collection method, participants could have fallen victim to the Hawthorne effect. Additionally, potential mental energy which would have been used to focus on the task could have been unconsciously exerted towards paying attention to the researcher in the participant’s periphery. Furthermore, while the experiment was not necessarily long, lasting only around twelve minutes, the researchers did ask participants to focus their attention for a prolonged period of time without breaks. This lack of breaks and solely using all their effort towards remaining still and looking for minor changes in realistic images, could have potentially lead to a form of burnout. While participants unanimously agree that the experiment was enjoyable, the prolonged extension of effort could have been tiring. Lastly, the experiment was conducted in a loud and busy lab, which could have been distracting to participants as they were trying to focus and complete their task in a timed fashion.

5.2 Future Directions

In terms of future directions, possible experiments could re-conduct the current experiment utilizing easier spot the difference images in a quiet room. Additionally, the software could be programmed to record accuracy in the form of extended fixations or double blinks to select the difference. Additionally, to further investigate change blindness, the original image could be displayed on both screens for a short amount of time before replacing one of the images with an altered version. This pre-existing knowledge of the original picture helpful suggests that people who did not understand the goal of the task or how to successfully accomplish their goal did not find the time limit and visual timer helpful. These findings support the idea that a lack of instructions combined with a time limit to complete a task could increase feelings of anxiety and stress. Lastly, participants viewed the task as easy also ignored the timer or time limit more often and enjoyed the task more. This suggests that people who found the differences in the pictures faster, did not have to concern themselves with the time limit. Additionally, this ease of task in a difficult assignment could lead to an increase in enjoyment potentially arising from feelings of excellence or mastery. Lastly, the relationship between the number of fixations on the timer in the ten Spot the Difference images and the difficulty of finding differences in the images in the intervention condition suggests that participants glanced at the timer in the top-middle of the screen more often with images where they had more difficulty finding differences.

Figure 10: Post-Assessment Correlations

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Difficulty</th>
<th>Fixations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty</td>
<td>Pearson Correlation</td>
<td>.653*</td>
</tr>
<tr>
<td>N</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

*Correlation is significant at the .05 level (2-tailed).

Figure 11: Timer Fixation and Difficulty Correlation

### Correlations

<table>
<thead>
<tr>
<th>Correlations</th>
<th>N</th>
<th>10</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast/Later</td>
<td>.849*</td>
<td>.087*</td>
<td>.385</td>
</tr>
<tr>
<td>Sig. (interv)</td>
<td>.300</td>
<td>.009</td>
<td>.306</td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Fast/Later</th>
<th>Anxious</th>
<th>Disturbing</th>
<th>Easy</th>
<th>Help</th>
<th>Ignor</th>
<th>Enjoy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>.849*</td>
<td>.087*</td>
<td>.385</td>
<td>.300</td>
<td>.009</td>
<td>.306</td>
<td></td>
</tr>
<tr>
<td>Sig. (interv)</td>
<td>.009</td>
<td>.009</td>
<td>.009</td>
<td>.009</td>
<td>.009</td>
<td>.009</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Fast/Later</th>
<th>Anxious</th>
<th>Disturbing</th>
<th>Easy</th>
<th>Help</th>
<th>Ignor</th>
<th>Enjoy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>.849*</td>
<td>.087*</td>
<td>.385</td>
<td>.300</td>
<td>.009</td>
<td>.306</td>
<td></td>
</tr>
<tr>
<td>Sig. (interv)</td>
<td>.009</td>
<td>.009</td>
<td>.009</td>
<td>.009</td>
<td>.009</td>
<td>.009</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>
6 CONCLUSION

Overall, both the timer condition and control condition performed about the same in terms of accuracy. The lack of significant difference in terms of fixations and saccades contradicts prior research but continue to prove the difficulty in noticing small changes in a format which has been marginally altered, also known as inattentional change blindness. Given the important nature in noticing minor detail changes in a variety of high stress jobs, such as fighter pilots in the sky, scientists working at mission control, and surgeons in an operating room, it is important to research the human mind and how it reacts to this change blindness. The addition of eye-tracking technology allows for a better understanding of the search and recognition process. As mentioned above, future research can focus on developing techniques such as increased salience and knowledge of differences to help reduce change blindness.

REFERENCES