Does Visual Search Performance Vary with Task Complexity as a Function of Mental Fatigue?

Richard Thripp

University of Central Florida

November 4, 2015

Abstract

This proposal explores a potential experiment exposing participants to low and high levels of mental fatigue, followed by easy and hard visual search tasks. Mental fatigue will be achieved with an easy or difficult memory task involving memorizing target letters and identifying their presence in subsequent displays of letters; this task will be employed continuously for two 45-minute periods. Easy visual search will involve looking for a circle with a peg among plain circles, while hard visual search will involve looking for circles with no pegs or two pegs among circles with one peg. Participants will be asked to commit to four testing sessions on four separate days in a 2 × 2 within-subjects design. A version of the NASA Task Load Index will be administered at the mid-point and end of each session. Interactions involving accuracy, reaction times, and subjective task load will be explored. The results may indicate that mental fatigue should be taken more seriously for motorists, air traffic controllers, and other critical activities and careers.

Keywords: visual search, task complexity, mental fatigue, vigilance, after-effects of mental effort

Does Visual Search Performance Vary with Task Complexity as a Function of Mental Fatigue?

Professionals who perform complex and important tasks, such as medical doctors and air traffic controllers, are often subject to schedules encouraging insufficient sleep and severe mental fatigue, which leads to an increased frequency of errors (Lowy, 2015). How to allow for sufficient sleep is sometimes obvious: do not schedule shifts eight hours apart, for instance. However, mental fatigue may be a more complex issue.

Bullock and Giesbrecht (2014) found that intensive aerobic activity resulted in faster reaction times for a visual search task, but no change in accuracy. Schellekens, Sijtsma, Vegeter, and Meijman (2000) conducted an impressive study on mental fatigue where each participant worked in two simulated eight-hour office workdays in easy and difficult mental-load conditions. A memory-search probe task was conducted at the beginning of the workday, the end of the workday, and 2.5 hours after the end of the workday—it was found that participants committed significantly more errors in both probe tasks after a difficult workday, with no change in reaction times. While this is reminiscent of a speed–accuracy tradeoff, their speed did not improve—fatigue is a possible reason for their reduced performance.

Faber, Maurits, and Lorist (2012) reiterated the impacts of mental fatigue on attention, finding that continuous task performance led to increased distractibility that manifested as reduced accuracy and increased reaction times over a two-hour period of performing a flanker task. With respect to vigilance tasks, Tiwari, Singh, and Singh (2009) found that higher task demand and perceived mental workload resulted in significantly lower motivation, concentration level, attention, and accuracy, albeit insignificant differences regarding reaction time. Haga, Shinoda, and Kokubun (2002) implemented a dual tracking and memory search task for a total of

30 minutes at three difficulty levels; task difficulty level was found to be a greater determinant of subjective levels of stress than time-on-task. While 30 minutes is not long, the authors say they have found similar results in subsequent 60- and 90-minute experiments (p. 142).

Overall, the research indicates that mental fatigue has detrimental effects on performance for visual attention tasks, though degradation is more consistent for accuracy as compared to reaction time. Interestingly, ocular instability has been found to increase with time-on-task regardless of task complexity; this leads to increased drift velocity and decreased saccadic and microsaccadic velocity (Stasi et al., 2013). Stasi et al. used an air traffic control task designed for laypersons, making their study particularly original and relevant. While high air traffic density is admittedly a more powerful predictor of slowed reaction times, reduced accuracy, and greater subjective mental fatigue (p. 2396), time-on-task remains important—too much time spent on a low-complexity task (e.g., low air traffic) can still be dangerous. The impact of ocular instability and other confounding factors might explain why the relationship between performance and task complexity in mental fatigue research is sometimes murky.

When combining high mental fatigue with hard visual search, inferior performance may be expected than if visual search was preceded by low mental fatigue. However, high mental fatigue may have detrimental effects even for easy visual search. This study may be the first attempt to systematically study these phenomena with contemporary, computer-based methods.

Current Study

The current study will combine two levels of mental fatigue with easy or hard visual searches in a wholly within-subjects design. Vision is arguably the most important sense, and reduced visual search performance can result in loss of life, limb, and property—air traffic control is a salient example (Stasi et al., 2013; Lowy, 2015), as is operating heavy machinery or performing delicate surgery, but driving a car is a relevant and far more common example. I could find no study directly examining visual search performance as a function of mental fatigue via a memory-search task, much less one that manipulates difficulty levels for both tasks. While Schellekens et al. (2000) manipulated an entire workday to be mentally easy or difficult and measured performance on a probe memory-search task, the probe task difficulty remained the same for all participants. Nevertheless, their probe task of remembering a list of alphabetic letters and identifying the presence of a matching letter is useful because it can be quite demanding or easy, by manipulating the compatibility of the noise letters (Eriksen & St. James, 1986). Therefore, for the mental fatigue task, I propose a hybrid between Schellekens et al. (2000) and Eriksen and St. James (1986). Subjects will be asked to remember letters and determine if displays contain a remembered letter, but the difficulty will be varied, primarily by compatibility of noise letters. In the easy condition, noise letters will be incompatible, meaning they look different from the target letters. However, in the hard condition, noise letters will be compatible, meaning they will be easy to confuse with the target letters. The hard condition, because of its ambiguity, should be mentally exhausting in comparison with the easy condition.

Participants in this study will be asked to commit to four separate days of testing. On each day, they will be assigned to one of four conditions in a within-subjects design that varies complexity and difficulty for a memory-search and visual search task. They will be exposed to all four difficulty combinations, but the order will be counter-balanced. The memory-search task will last a total of 90 minutes, in two sessions of 45 minutes interspersed with a visual search task, a subjective fatigue questionnaire (the NASA-TLX), and a short break. The visual search task will be given three times (beginning, middle, and end) per session, and the NASA-TLX will be given twice (middle and end).

Easy visual searches will involve looking for a pegged circle among a field with 12–13 non-pegged circles, while hard visual searches will involve looking for an unpegged circle *or* a circle with two pegs among a field with 11–13 circles having one peg each. It is expected the hard visual search will be of much greater difficulty, given the nature of the task and the added cognitive load of looking for two types of targets. Additionally, the pegs in the hard condition will be about half the length of the pegs in the easy condition. The idea for pegged circles is borrowed from Persuh, Genzer, and Melara (2012), though they were using them for an iconic memory task, rather than a strict visual search task. We will see how performance on each level of the visual search task varies when paired with low versus high mental fatigue tasks.

Proposed Methods

Participants. Prospective participants will be required to have normal or corrected-to-normal vision, normal sleep patterns, to be without psychiatric or neurological illness, and to not be taking medications that might confound results. They will be asked to commit to 4 days of testing and to select possible dates and times they may be available. Participants will not be allowed to schedule more than one session in a calendar day. Participants will be asked to not consume alcohol for 24 hours before each session nor caffeine for 12 hours—caffeine has been found to have specific effects on information processing (Lorist, Snel, Kok, & Mulder, 1994). Participants will be encouraged to schedule sessions between 9:00–12:00 in the morning to avoid fatigue from daily life or diurnal rhythms impacting results, but it is not expected that all participants will be able to accommodate this. Compensation will be offered on a sliding scale to minimize attrition: for each of the four sessions, participants will receive, in order, $30, $40, $55, and $75. If an appointment is missed, participants will have the option of re-scheduling so they can still complete all four sessions, but this point will be omitted in the briefing session to discourage missed appointments, unless participants specifically ask. The air temperature of the testing space will be 72 F. Since the proportion of participants who attrite is difficult to predict, the initial sample size should be somewhat liberal to ensure enough participants complete all four conditions—perhaps 25–30 participants.

Materials and procedures. Participants will be seated at a computer desk in a lab with overhead fluorescent tube lighting and no natural lighting, 75 cm from a 19” true-color LED-backlit LCD monitor with a 16:9 aspect ratio, resolution of 1920 × 1080 pixels, and refresh rate of 60 Hz. The monitor will be calibrated to a color temperature of 6500 K, display gamma of 2.2, and luminance of 125 cd/m^2 with a ColorVision Spyder colorimeter. Both types of tasks will entail binary responses; participants will textually be instructed to be both fast and accurate and no further elaboration will be given. Participants will press the “J” key on a QWERTY wired USB keyboard for “yes” and the “F” key for “no.” At all times, a computer program will record their reaction times and correctness, which will subsequently be statistically analyzed by the researcher. Between trials, a red fixation cross will be displayed for 200 ms on visual search trials; masking “%” percentage signs (like the asterisks used by Schellekens et al., 2000, p. 40) will be shown for 200 ms on mental fatigue tasks. When participants respond correctly, the fixation cross or masking percentage signs will be immediately displayed with no feedback; for incorrect responses, a blank screen saying “Incorrect” will first be displayed for 100 ms, followed by the fixation cross or masking percentage signs for 200 ms. Trials where participants fail to respond within 1500 ms for low mental load or easy visual searches, or within 2500 ms for high mental load or hard visual searches, will be counted as failures to respond for purposes of data analysis, but participants will not know about this. All text will be in Arial font. Backgrounds will be dark gray (#7A7B82) and cued items will be white (#FFFFFF) in all instructions and tasks, except the NASA-TLX questionnaire (see figure 15).

Mental fatigue tasks. The mental fatigue tasks will be similar to the probe task used by Schellekens et al. (2000), but will employ the concept of compatible and incompatible noise letters exemplified by Eriksen and St. James (1986). High mental fatigue tasks will involve participants being shown four underlined letters for 950 ms. On 2–5 subsequent displays, participants will have to respond indicating whether one of the four target letters is among the four displayed letters. Easy mental fatigue tasks will show three underlined letters for 1400 ms, and on exactly 3 subsequent displays, participants will respond indicating whether one of the three cue letters is among the three displayed letters. No time limits will be imposed on any responses.

Eriksen and St. James (1986) did a visual search experiment requiring little or no memory, so they simply duplicated the cue letters for compatible noise. For this experiment, I developed a list of 19 letters that look easily confusable with other letters (figure 1).

Figure 1: Letters and compatible noise letters employed for the high mental fatigue condition. 19 of 26 letters are included; excluded are A, C, G, H, L, T, and W. Note that most relationships are reciprocal, except D/B, R/P, and Y/V. Unfortunately, the author came up with this list in Microsoft Notepad in about 30 minutes; he has no empirical evidence to support it.

B	P	M	N	U	V
D	B	N	M	V	U
E	F	O	Q	X	K
F	E	P	B	Y	V
I	J	Q	O	Z	S
J	I	R	P		
K	X	S	Z		

The low mental fatigue condition will only use the letters the author thinks are less confusable as target letters: A, C, G, H, L, T, and W. There will only be three letters per display, and predictably, three tasks per trial. While any of the 26 letters may be displayed in tasks, there will be a 25% chance that the same letter will be displayed in 2 out of 3 spots.

The high mental fatigue condition will use the 19 confusable letters from figure 1 as target letters. There will be four letters per display, and participants will be given two, three, four, or five tasks per trial, randomly; therefore, they will not know whether a new trial has started until the masking percentage signs appear for 200 ms, following which they will have only 950 ms to commit the newly displayed letters to memory. There will be a 60% chance that the compatible noise letter for a target letter will be selected as a displayed letter on each task; otherwise, a random letter will be selected. Positions will be shuffled on all trials.

It is predicted that the hard visual search task will be very distressing and mentally fatiguing, given the additional load, randomness, confusing letters, and reduced time for memorization (950 ms versus 1400 ms).

Figure 2: Instructions screen for both easy and hard visual search tasks.

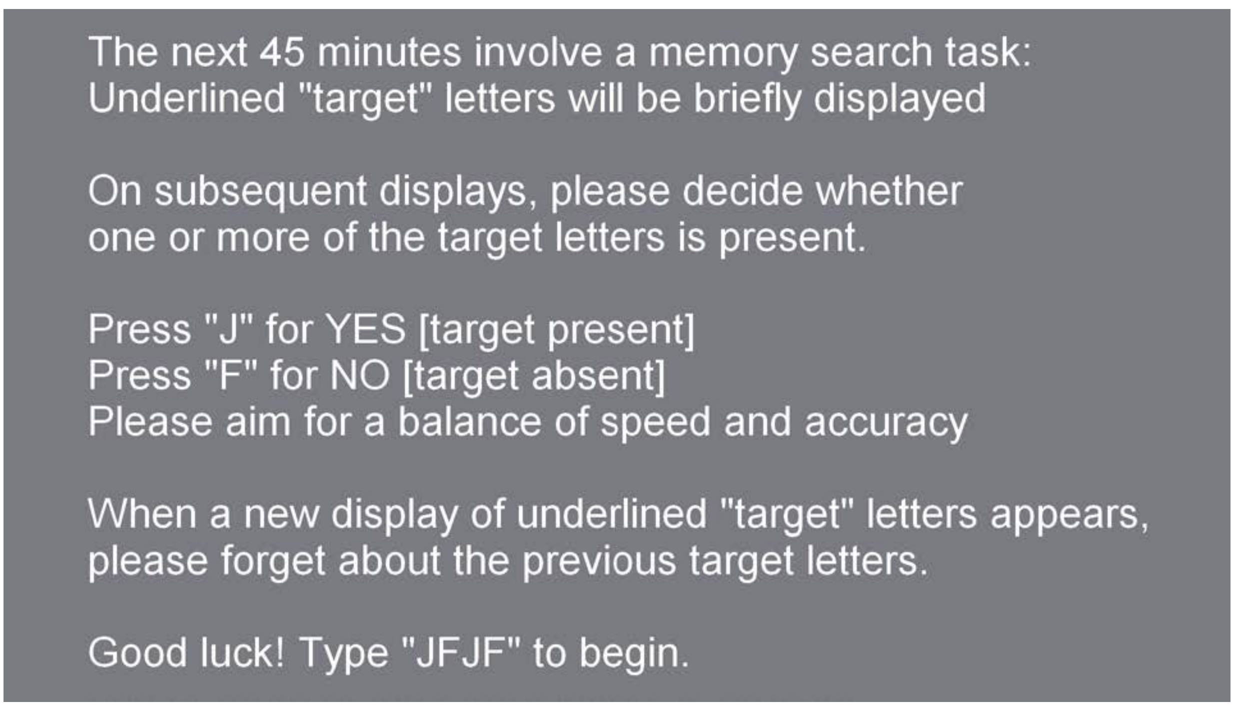


Figure 3: Example of a low mental fatigue display of target letters.

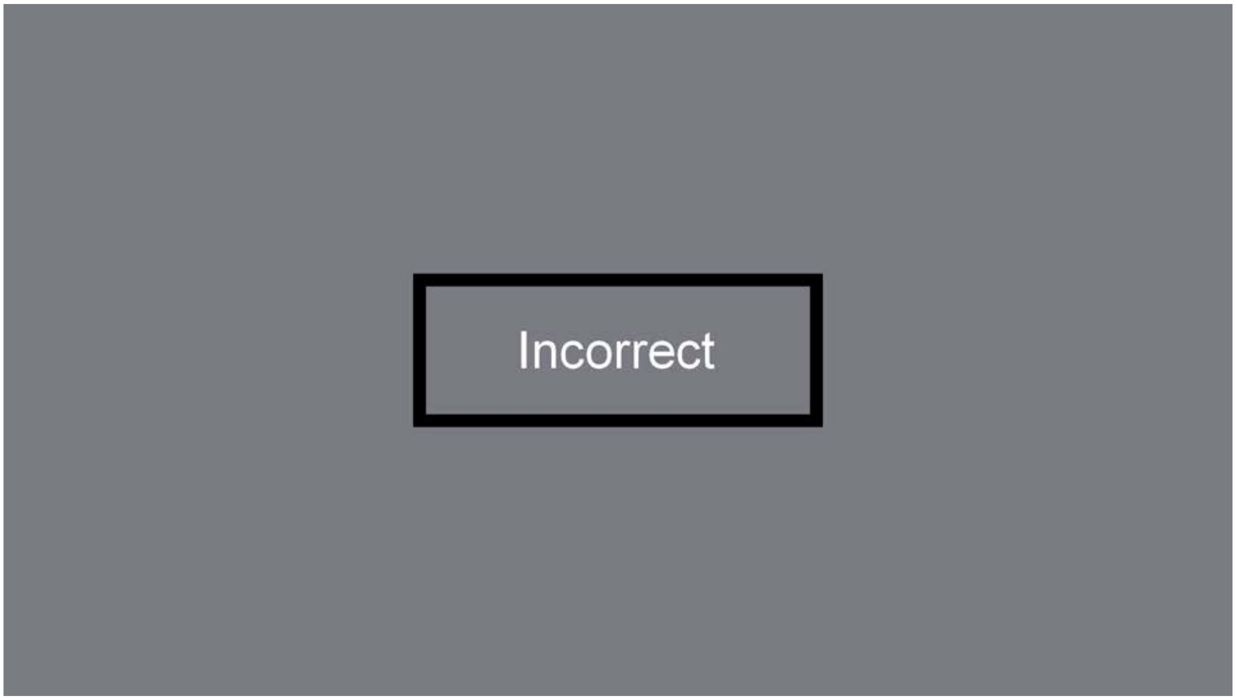
Figure 4: Example of a low mental fatigue task. Two letters are duplicated in this display (occurs 25% of the time with the hope of making the task even easier).

Figure 5: Example of a high mental fatigue display of target letters.

Figure 6: Example of a high mental fatigue task where compatible noise letters for all target letters from figure 5 are displayed.

Figure 7: Masking percentage signs shown for 200 ms between high mental fatigue trials. Masks for low mental fatigue trials have three percentage signs instead of four.

Figure 8: Incorrect notification given for 100 ms in all tasks following an incorrect answer.



Visual search tasks. The visual search tasks will borrow the idea of pegged circles from Persuh, Genzer, and Melara (2012). In the easy condition, participants will be looking for a circle with a long peg in any of four 90º orientations among a display containing 13 circles, of which either 12 or 13 are unpegged. In the hard condition, participants will be looking circles with no pegs or two pegs among a display containing 13 circles, of which 11, 12, or 13 have one peg; pegs in the hard condition may be in any of eight 45º orientations; therefore, 28 variations of circles with two pegs may appear, as shown by this combinatory expression that disallows repetition:

$$\frac{n!}{(n-r)!(r!)}$$

Therefore, the total numbers of potential targets types and potential distractors are 29 and 8, respectively, in the hard visual search condition, as compared to 4 and 1, respectively, in the easy visual search condition. Time limits will not be imposed in any visual search tasks.

In both conditions, 50% of trials will have target(s) present and 50% will have target(s) absent. For the hard visual search condition, 15% of trials will have both targets present and 35% will have one target present. This means that 30% of target-present trials in the hard condition will be easier, since participants can press “J” for YES as soon as they find one of the two targets. This is intended to add flair to the hard condition; it should not impact results since the conditions are not designed to be compared to each other directly, and the hard condition is already so much harder than the easy condition.

Figure 9: Instructions screen for easy visual search condition.

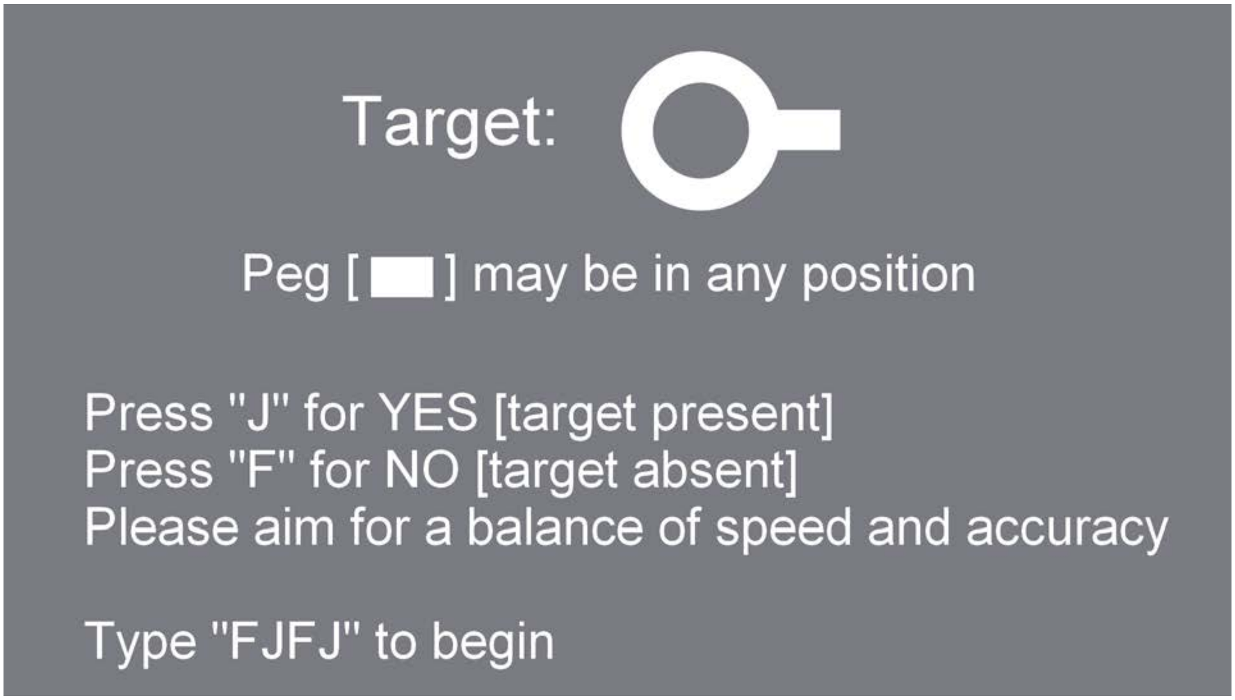


Figure 10: Example of an easy visual search task, target present.

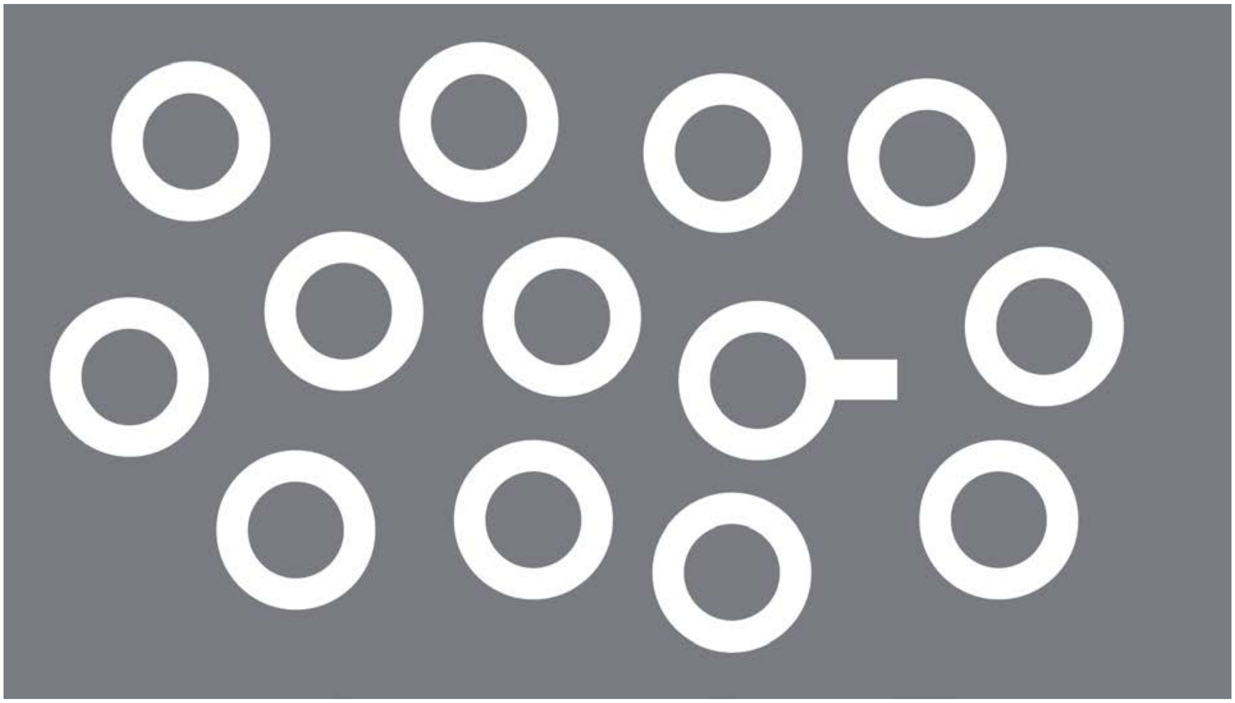


Figure 11: Instructions screen for hard visual search condition.

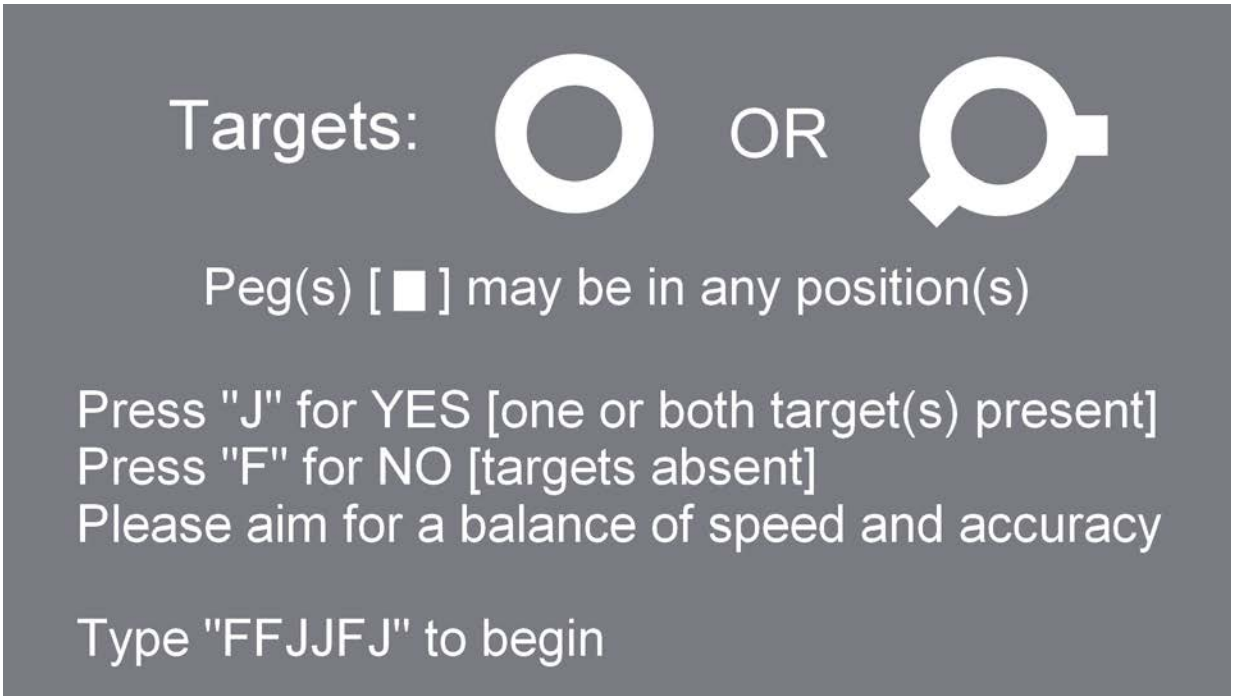


Figure 12: Example of a hard visual search task, both targets present (occurs in 15% of trials).

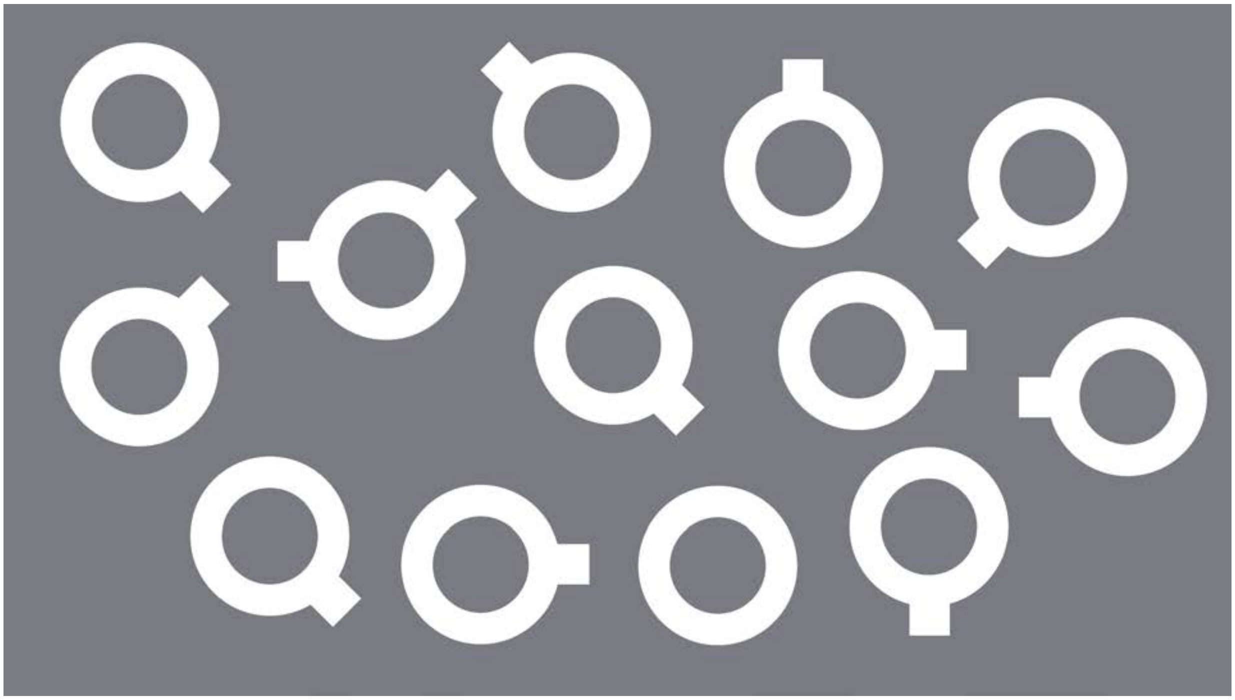
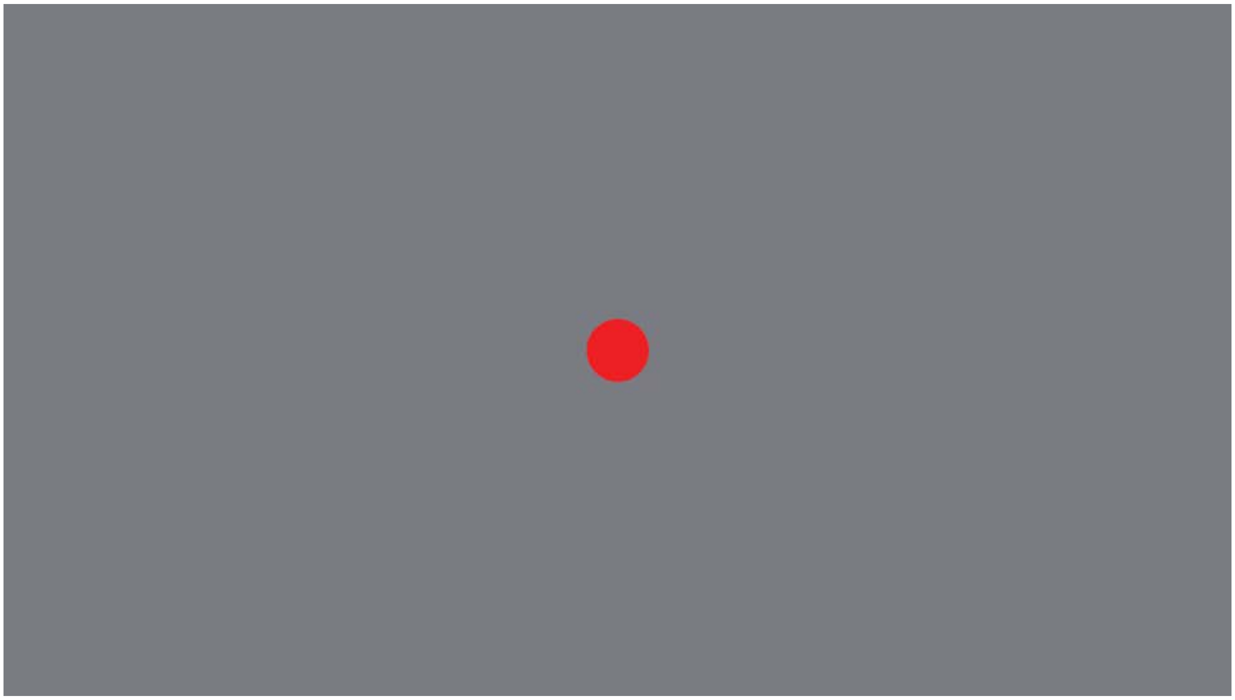


Figure 13: Fixation cross shown for 200 ms between visual search trials.



Further procedural details. The visual search task will be five minutes long and will be given continuously for 10 minutes at the start of all sessions (at the same difficulty level that will be used throughout the session), with the hope of standardizing practice effects so all participants are operating at a roughly equal level of practice in the pre-test, mid-test, and post-test. After a five-minute break, the visual search task will be given again as the official pre-test. During the 90-minute mental fatigue task, there will be an intermission after 45 minutes where the visual search task will immediately be given again (mid-test), followed by a fatigue questionnaire, and a 10-minute break. Upon return from break, the same sequence will repeat, and the session will conclude after the visual search task is administered for the 3rd time and the fatigue questionnaire is administered for the 2nd time.

Task complexity will be varied at two levels for both the mental fatigue task and subsequent visual search task. A within-subjects design will be used, where participants will be exposed to all four levels of the experiment on four separate days.

All tasks (including before and after the intermission) will be at the same complexity level for each session and day. The following four conditions will be used:

Condition 1: Low mental load, easy visual search

Condition 2: High mental load, easy visual search

Condition 3: Low mental load, hard visual search

Condition 4: High mental load, hard visual search

Conditions will be counter-balanced so that approximately a quarter of participants start with each condition in their first session. Subsequent sessions will be similarly counter-balanced.

In sustained tasks exceeding 30 minutes, maintaining arousal level is important to prevent boredom (Haga et a., 2002, p. 142); however, the author believes the easy mental load task is sufficient to maintain arousal, because of the additional memory requirement. Further, due to the brevity of the visual search tasks (less than 30 minutes), it is not anticipated that arousal level will be a problem with easy tasks.

Fatigue questionnaire. The NASA Task Load Index (TLX) will be used as our fatigue questionnaire (Hart & Staveland, 1988). A computerized version will be administered, in a "raw TLX" mode where pairwise comparisons are eliminated for brevity, and where the "physical demand" subscale is dropped. The other five subscales will be retained, including "temporal demand"—even though there is no time pressure placed on participants, it may be interesting to see whether more temporal demand is erroneously reported in the hard conditions. Participants will use a wired USB optical mouse to adjust the sliders on the subscales to their desired ratings—100 possible responses for each subscale will be possible, with each of the 20 displayed hatch marks being 5 points apart (though no numerical indication will be provided). Participants will be instructed to consider *both* the preceding mental fatigue tasks *and* visual search tasks as a unit, since the NASA-TLX will not be administered in between the two tasks.

Figure 14: Instructions screen for NASA-TLX questionnaire.

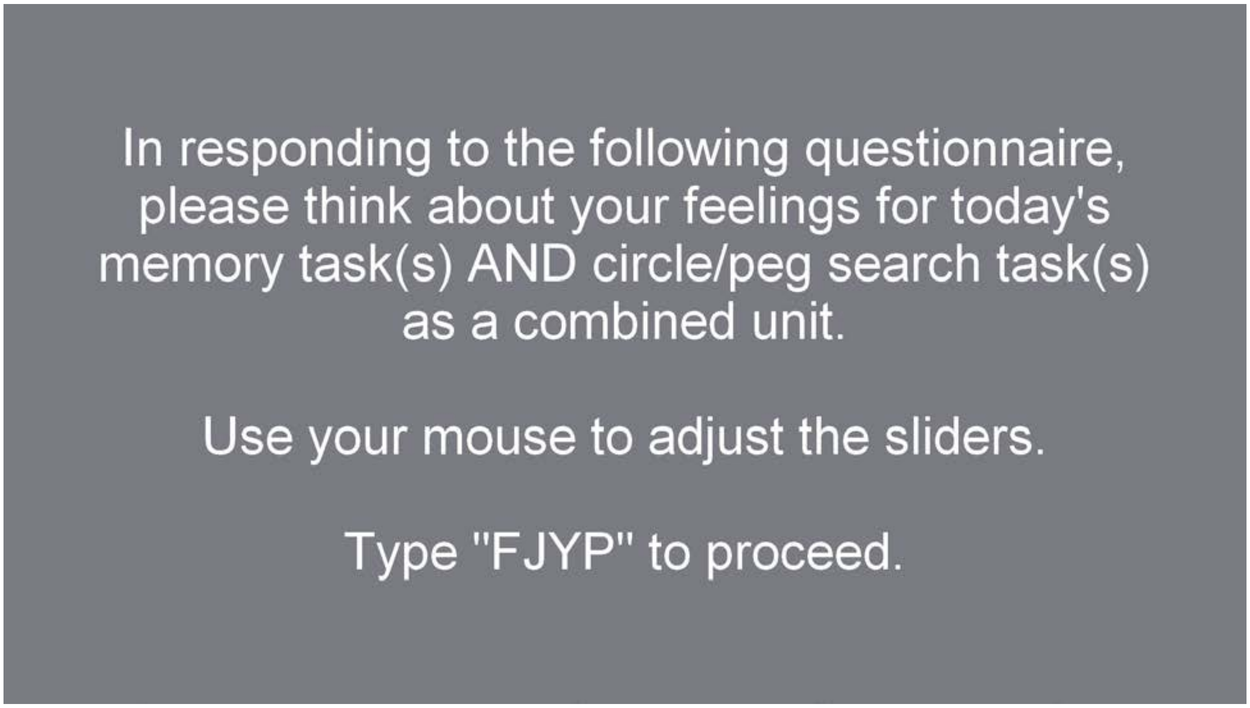
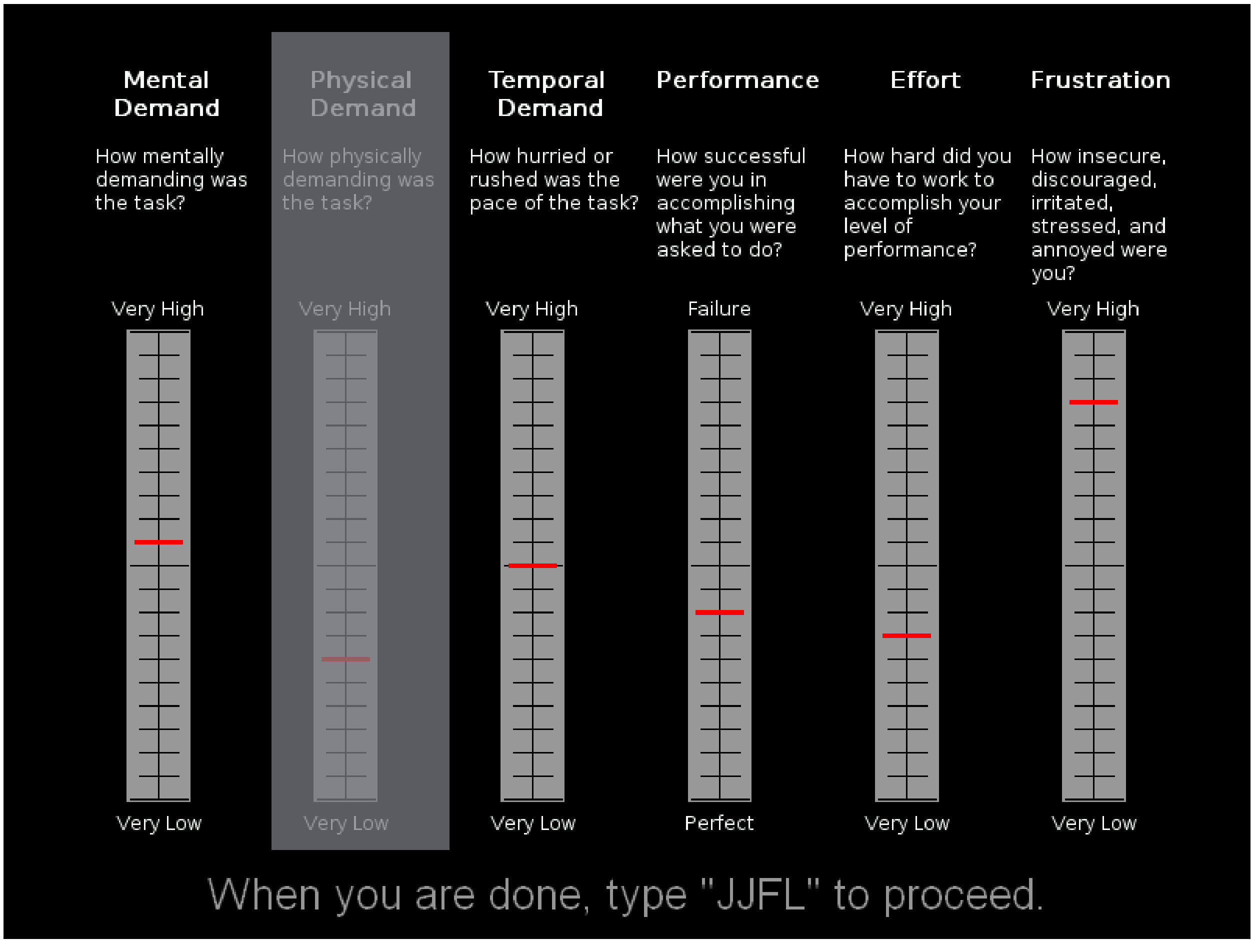


Figure 15: The modified NASA-TLX questionnaire that will be used. Note that the "physical demand" subscale will not be used or shown in the final implementation. Figure adapted from "Screenshot of the PEBL TLX scale" (2012).



Expected Results

It is predicted that accuracy and reaction times will be better on the hard visual search task when it follows low mental load rather than high mental load. Performance should be better on the easy visual search if mental fatigue is low; however, the effect of mental fatigue may be more pronounced with hard visual searches since they are more cognitively taxing. With the high mental fatigue conditions, visual search performance should be highest in the pre-test, lower on the mid-test, and lowest on the post-test, similar to findings by Tiwari et al. (2009) on stress and vigilance. Given that Haga et al. (2002) found that difficulty level, *not* time-on-task determined task load levels in their dual tracking and memory search task experiments, results for the NASA-TLX may be largely the same for the mid-administration as compared to the post-administration. Overall, NASA-TLX results should indicate low stress on the low mental load / easy visual search condition, and relatively higher stress on the high mental load / hard visual search condition, particularly on the effort and mental demand subscales. However, Haga et al. did find that *other* workload measures were sensitive to time-on-task and surmised the relationships to be complex and hard to dissect (p. 142). Therefore, as the session drags on, degradation in performance on the high mental fatigue task would be unsurprising.

The above predictions assume that participants who do not complete all four sessions will be excluded from statistical analyses. Counter-balancing will be used to ensure an approximately equal number of participants receive each permutation of conditions, which should compensate for many possible extraneous variables. The high mental load condition may be particularly frustrating for some participants—it may be necessary to make the task less frustrating to reduce attrition, but the verdict is unclear at this preliminary stage.

Since the easy and hard conditions within each category (mental fatigue and visual search) are different from each other in multiple ways, it is not anticipated that they will be directly comparable to each other—they may only be comparable to themselves as a function of difficulty level in the other category.

Many within-subject analyses of variance (ANOVAs) will be conducted on the statistical data to look for main effects between mean error rates and mean accuracy rates for the different combinations of tasks, as well as interactions with subjective fatigue reports on the five selected subscales of the NASA-TLX questionnaire, both as a whole and individually. Additionally, the real-time data for response times and accuracy rates for mental fatigue tasks will analyzed temporally, by breaking each 45-minute block down into nine 5-minute chunks and looking for trends, both for individual subjects and across multiple subjects. Since this is a complicated experiment, there are many interactions and combinations to be examined. Therefore, it is expected that at least several significant interactions will be found.

Discussion and Implications

This study will help elucidate whether mental fatigue has a relationship with visual search performance as a function of task complexity. If mental fatigue leads to poorer visual search performance, and these results withstand scrutiny and are amiable to replication, it should be recommended that motorists and others limit their mental fatigue prior to driving or performing other dangerous tasks.

The consequences of mental fatigue may be more severe when coupled with poor quality sleep, distracting emotions, alcohol and drug use, and difficult conditions such as heavy rain while driving or high air traffic volume for air traffic controllers (Stasi et al., 2013; Lowy, 2015)—further research is needed on these combinations. Studying mental fatigue in the laboratory may produce conservative results, particularly because we exclude participants with sleep aberrations, alcohol or caffeine addictions (Lorist et al., 1994), drug addictions or dependencies, and uncorrected vision. While these exclusions are necessary to produce homogenous samples, they are not naturalistic. Therefore, the costs and consequences of mental fatigue may be higher and more pronounced in the real world, where these factors are present.

How do you prevent prolonged mental fatigue? Frequent breaks are in order. Furthermore, structuring the workday in such a way that mental fatigue is spread out, interspersed with mundane activity such as basic data entry or housekeeping, may help. Research studies that identify mental fatigue as a real problem, as this one hopes to, will help us influence industries and individuals to account for it in their daily operations and lives.

References

Bullock, T., & Giesbrecht, B. (2014). Acute exercise and aerobic fitness influence selective attention during visual search. *Frontiers in Psychology*, *5*(1290), 1–11. doi:10.3389/fpsyg.2014.01290

Di Stasi, L. L., McCamy, M. B., Catena, A., Macknik, S. L., Cañas, J. J., & Martinez-Conde, S. (2013). Microsaccade and drift dynamics reflect mental fatigue. *European Journal of Neuroscience*, *38*, 2389–2398. doi:10.1111/ejn.12248

Eriksen, C., & St. James, J. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, *40*(4), 225–240. doi:10.3758/BF03211502

Faber, L. G., Maurits, N. M., & Lorist, M. M. (2012). Mental fatigue affects visual selective attention. *Plos One*, *7*(10), 1–10. doi:10.1371/journal.pone.0048073

Haga, S., Shinoda, H., & Kokubun, M. (2002). Effects of task difficulty and time-on-task on mental workload. *Japanese Psychological Research*, *44*(3), 134–143.

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.), *Human mental workload*. Amsterdam: North Holland.

Lorist, M. M., Snel, J., Kok, A., & Mulder, G. (1994). Influence of caffeine on selective attention in well-rested and fatigued subjects. *Psychophysiology*, *31*, 525–534.

Lowy, J. (2015, August 11). Air traffic controllers are dangerously overworked. *Business Insider*. Retrieved November 4, 2015, from http://www.businessinsider.com/air-traffic-controllers-are-dangerously-overworked-2015-8

Persuh, M., Genzer, B., & Melara, R. D. (2012). Iconic memory requires attention. *Frontiers in Human Neuroscience*, *6*(126), 1–8. doi:10.3389/fnhum.2012.00126

Schellekens, J. M. H., Sijtsma, G. J., Vegter, E., & Meijman, T. F. (2000). Immediate and delayed after-effects of long lasting mentally demanding work. *Biological Psychology*, *53*, 37–56. doi:10.1016/S0301-0511(00)00039-9

Screenshot of the PEBL TLX scale, an implementation of the NASA-TLX subjective workload instrument. (2012, November 26). In *Wikipedia*. Retrieved November 4, 2015, from https://en.wikipedia.org/wiki/File:Tlx.pbl.png

Tiwari, T., Singh, A. L., & Singh, I. L. (2009). Task demand and workload: Effects on vigilance performance and stress. *Journal of the Indian Academy of Applied Psychology*, *35*(2), 265–275.