

Investigating impact of difficulty in a visual search task with fMRI

Authors: Simon Anneberg Merrild Hansen (201609506) in cooperation with Study Group 9 (Esther, Pernille, Nanna and Alberte)

Introduction to Cognitive Neuroscience, Cognitive Science

Professor: Mikkel Wallentin

School of Communication and Culture, University of Aarhus

Note: The paragraphs are written by me (Simon) unless other is stated

Abstract

The purpose of this study was to investigate how difference in complexity of visual search stimuli affects activity in brain. The areas of interest were primarily the visual cortex and the anterior cingulate cortex (ACC). The participant was placed in an fMRI scanner and had to decide, whether a target was present or not in 6*6 square of letters in different colours. Increased activity in visual cortex as well as left parietal cortex was observed for more demanding trials. No activity was observed in the ACC. The results from this study should be interpreted with caution, as only one participant took part in the study. The findings should therefore only be used to guide further research into the topic of visual search mechanisms.

Keywords: Visual Search, fMRI, Feature Integration Theory, 2x2 factorial design

Introduction

How do we find a friend in the crowd? How do we find our car on a parking lot? Visual searches are done by humans all the time, and it is therefore interesting to investigate the mechanism of the function. Knowledge about visual search mechanism is important, as it can lead to an optimization of search techniques and working environment in fields such as airport security, where the security staff must be able to spot dangerous items in a suitcase. It might also be possible to transfer

knowledge about, how humans perform visual searches to computers to improve recognition software, which can be used in for example surveillance cameras.

Treisman and Gelade (1980) proposed a theory for, how people combine features into an object. This theory is known as Feature Integration Theory (FIT). FIT suggests that features such as colour, shape and orientation are processed pre-attentively, while object identification requires attention. When we look at a scene, the features are extracted automatically, but we have to attend to a part of the scene to put the features together into an object. This is done by moving the spotlight of attention to the area of interest. In a visual search task, according to FIT, if a person has to identify an object based on only one feature, this can be done very rapidly. The effect is also known as the pop-out effect (Wang, Cavanagh & Green, 1994). The target object sort of “pops out” from the scene. Search tasks that involve this kind of target is known as feature searches (Treisman & Gelade, 1980). In contrast, if the person has to combine multiple features into an object to distinguish a target from the distractors, it will take longer time. This would require the person to perform the visual search in serial, meaning that the person would have to attend to one object at a time comparing it to a stored template of the target object. In the visual search paradigm, this is known as conjunction search (Treisman & Gelade, 1980). The experiment presented in this paper was designed to investigate the functional difference between feature search and conjunction search in the brain. The following hypotheses were made:

H1: An increase in task difficulty will lead to an increased activity in the visual cortex measured as BOLD response

H2: An increase in task difficulty will lead to an increased activity in the anterior cingulate cortex (ACC) measured as BOLD response

H3: It will require more time to decide, whether target is present or not for hard trials compared to easy trials. The same pattern will apply for trials, where target is absent compared to present.

H1 is based on the idea that a more complex search will require more engagement of the visual system. This has for instance been observed by Leonards, Sunaert, Van Hecke and Orban (2000). H2 is based on the observation that activity in the ACC is often seen in relation to task difficulty (Carter et al., 2000). H3 is based on FIT by Treisman and Gelade (1980).

An alternative model of visual search has been proposed by Wolfe (1994). This model is known as the guided search model, which states that, when conducting a visual search, the person first analyse multiple features across the scene. Areas containing features similar to target will yield a higher activation than areas that does not share features with target. The person will then search for the target in the area eliciting the highest activity, until target is found. In this experiment both theories would predict a tendency according to H3 in the behavioural responses. A discussion of which theory is more accurate is therefore beyond the scope of this article.

Methods and Material (written in cooperation with study group)

Functional magnetic resonance imaging (fMRI) was used for the purpose of achieving further understanding for the neural and cognitive aspects underlying visual search in different levels of complexity. Moreover, the usage of chronometric methods, allowing collection of reaction time (RT) to the different stimuli, gave further insight into the cognitive mechanisms at play. The imaging facilities used (MAGNETOM Skyra MRI scanner) were put at disposal by Aarhus University Hospital and CFIN at Aarhus University. The interscan interval used was 2 seconds.

Stimuli and experimental procedure

The participant was a 27-year-old, right-handed female with no visual or cognitive deficits reported. The participant signed a consent and agreed to perform the experiment twice. She was informed that she could withdraw at any time. Instructions for the experiment were explicitly explained, before she was placed in the scanner.

The stimuli presented consisted of 36 letters positioned in a square of 6x6. The letters chosen for the stimuli was I, T and Z, all three sharing similar features of horizontal lines (T and Z) or vertical lines (I and T). Three easy distinguishable colours were included – red, yellow and blue, respectively. At all times throughout the experiment, the goal of each trial was to elicit a response of whether or not a yellow T was present, using the response box in the scanner. The other letters made up the distractors. The colour of the distractors depended on each condition, as the experiment had 4 conditions in a 2-by-2 factorial design (Figure 1). Each of the four conditions consisted of 25 trials, resulting in a total of 100 trials, presented in a randomized sequence.

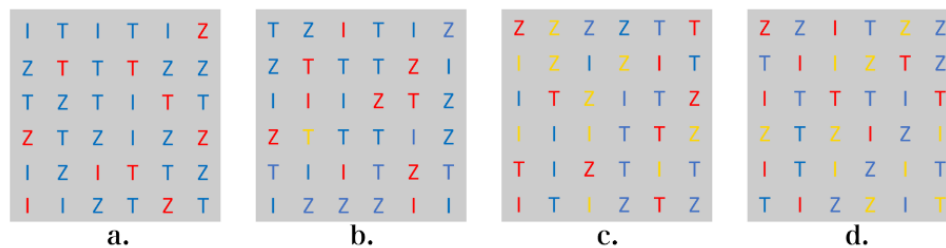


Figure 1: Overview of stimuli from the different conditions. a) Yellow T is absent, distractors are either blue or red (condition Easy-Absent, EA). b) Yellow T is present, distractors do not share colour feature with the target (condition Easy-Present, EP). c) Yellow T is absent, the remaining distractors have common colour features to a yellow T, although none of the Ts present in this condition are yellow (condition Hard-Absent, HA) d) Yellow T is present and distractors share colour feature with the target (condition Hard-Present, HP)

Extensive control of the stimuli was exerted for the purpose of controlling the outcome data. A series of criteria for the stimuli was made: *i*) Each letter had 12 representations in each trial, *ii*) 10 red letters in each condition *iii*) In condition HP and HA 10 yellow letters were present, *iv*) The yellow T was placed six different spots in each of the 3x3 squares within the stimuli, as it made up a total of 24 trials, the 25th target T of the conditions EP and HP were placed at random. Each trial was shown for 3 seconds separated by 4 seconds of rest with a fixation cross on the screen. The experiment took 12 minutes (24 minutes for both sessions). This resulted in a total of 38 minutes, including an external experimental paradigm between the two sessions and a structural scan prior to the first session.

Analysis (written in cooperation with study group)

On the behavioural data a multiple regression including the interaction effect was carried out using RStudio (RStudio Team, 2015). The model predicted reaction time by difficulty and target (present/absent). Both correct and incorrect responses were included in the analysis. Two trials were excluded due to missing responses, resulting in a total of 198 trials. A t-test comparison between the two repetitive sessions of the experimental paradigm referred, showed no significant difference between reaction times across sessions, $b = -0.07$, $SE = 0.07$, $t(196) = -1.05$, $p = 0.30$. Therefore, the two sessions are combined in the analysis.

The fMRI analysis was carried out using Statistical Parametric Mapping 12 (The FIL Methods group, 2016). In general, standard settings were applied. The images were realigned to correct for movement. As coregistration was carried out, the reference structural image was compared to the mean functional image, and each functional image was compared to the reference structural image.

Segmentation was carried out, where each voxel was classified as different tissues. All images were spatially normalized into MNI space. In the last pre-processing step the images were spatially smoothed. The analysis was carried out using the functions: fMRI model specification, Model estimation and Contrast Manager. Four contrasts were chosen on the basis of the experimental condition; Absent > Present, Hard > Easy, as well as two contrasts created to investigate the interaction effect.

Results (written in cooperation with study group)

Behavioural results

The overall hit rate for the participant resulted in a total of 96.5 % correct responses. According to the multiple regression analysis the reaction time was significantly greater for the hard conditions than for the easy conditions, $b = 0.76$, $SE = 0.06$, $t(194) = 11.85$, $p < .001$ (figure 2 and 3). Though there was no significant effect on the reaction time for, whether the target was present or not, $b = -0.12$, $SE = 0.06$, $t(194) = -1.81$, $p = 0.07$. The same model showed a significant interaction effect between difficulty and, whether or not target was present, $b = -0.30$, $SE = 0.09$, $t(194) = -3.28$, $p < .01$. F-statistics proved the significance of the model, $F(3,194) = 75.25$, $p < .001$, adjusted $R^2 = 0.53$.

Condition	Mean reaction time (seconds)	Standard deviation
Easy Present	0.75	0.16
Easy Absent	0.86	0.30
Hard Present	1.21	0.36
Hard Absent	1.62	0.41

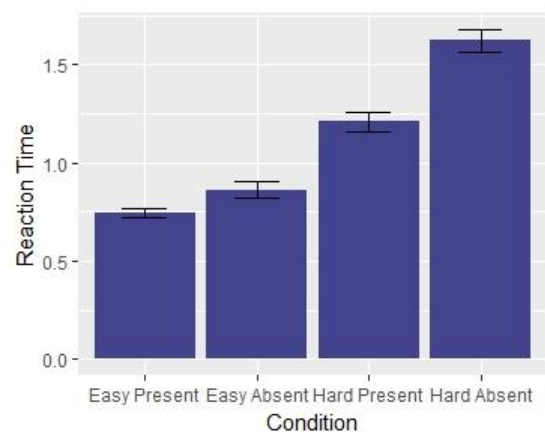


Figure 2: Table of the mean reaction time for each condition

Figure 3: Bar plot of reaction time from behavioural data

fMRI results

The analysis carried out revealed nothing according to the hypotheses using a 0.05 FWE corrected p-value. Therefore, an 0.001 uncorrected p-value was used instead, confirming increased activation in visual cortex in line with H1 (figure 4.a). The activation was higher, when the target was absent

than when present (figure 5). Nothing validated H2 of activation in ACC. Another area of interest was discovered by significant results in the parietal cortex (figure 4.b). The activation in the parietal cortex increased with the difficulty of the task, showing a similar tendency as the behavioural data (figure 3 and 6).

For all contrasts, the BOLD signal was weaker in the first session compared to the second. As the same pattern of activation appeared in both sessions, the analyses of the two sessions were combined.

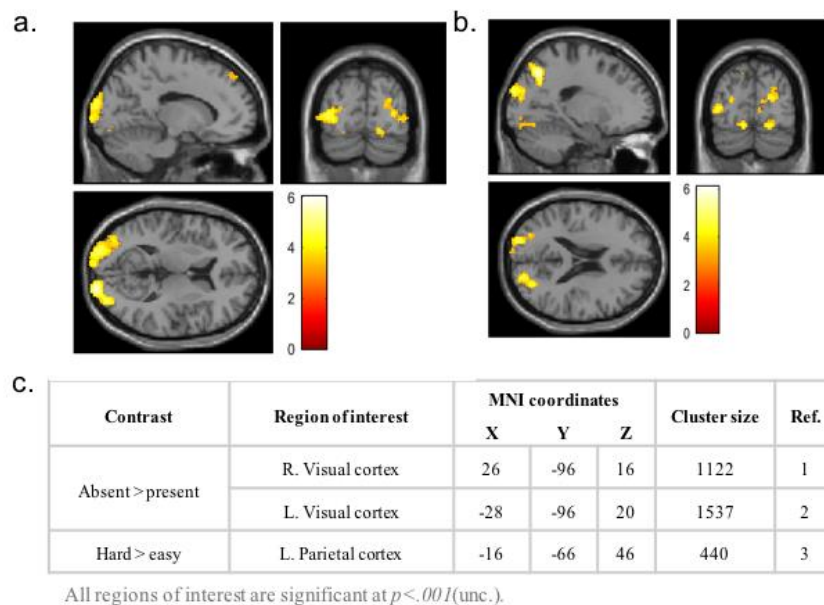


Figure 4: 4.a shows increased activation in the visual cortex for contrast Absent > Present. 4.b shows increased activation in the parietal cortex for contrast Hard > Easy. 4.c Table of MNI coordinates for regions of interest. These coordinates are similar to other findings of same areas in the brain. (Kristjánsson, Vuilleumier, Schwartz, Macaluso & Driver, 2007; Schumacher, Hendricks & D'Esposito, 2005)

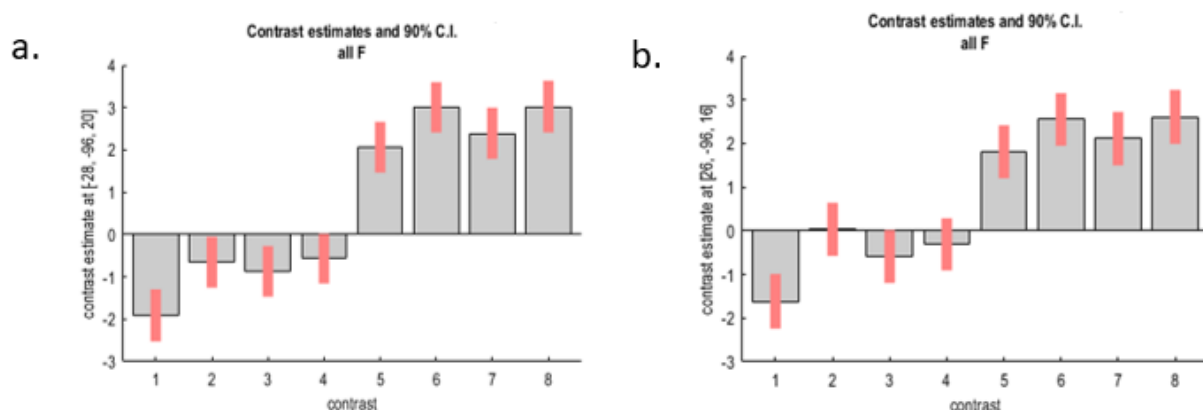


Figure 5: Graph of the activity for the Absent > Present contrast in the left visual cortex (5.a) and the right visual cortex (5.b). Contrasts 1-4: 1st round, contrasts 5-8: 2nd round. The order of the conditions is as follows: EP, EA, HP, HA in both rounds. The MNI coordinates correspond to those presented previously (figure 4.c1 and 4.c2).

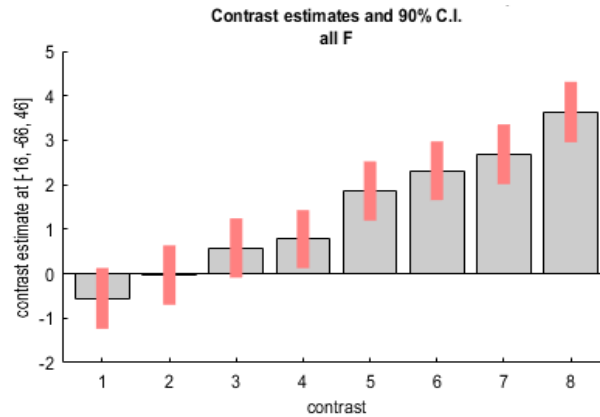


Figure 6: Graph of the activity for the Hard > Easy contrast in the parietal cortex. Contrasts 1-4: 1st round, contrasts 5-8: 2nd round. The order of the conditions is as follows: EP, EA, HP, HA in both rounds. The MNI coordinates correspond to those presented previously (figure 4.c3).

Discussion

The results in general seem to support H1, as there is an increased activity in the visual cortex, when target is absent compared to when the target is present. This is also in line with the results produced by Treisman and Gelade (1980). There also seem to be an increase in activity in the visual cortex in the hard condition compared to the easy condition. This effect is, however, much weaker. The behavioural data showed support for H3 and are thereby supporting Treisman and Gelade's (1980) FIT.

While the results seem to support H1 and H3 there was no significant increase in activity in the ACC, as the task difficulty increased. The experimental data are therefore unable to support H2. This could, of course, be due to the fact that we do not have a sufficient amount of data, but it is more likely due to other reasons. In a study by Pardo, Pardo, Janer and Raichle (1990), the participants had to perform the Stroop task. The participants underwent two conditions: In the congruent condition the colour of the word matched the colour name. In the incongruent condition there was a mismatch between colour name and colour of the word. Increased activation in the ACC was observed for the incongruent trials. Based on this, Pardo et al. (1990) suggested that ACC is important for selecting and recruiting centres appropriate for the task.

During the analysis, an interesting discovery was made. An increased activity in the parietal cortex was observed in the hard condition compared to the easy condition. As the parietal cortex has been suggested to play a role in attentional processes (Behrmann, Geng, & Shomstein, 2004; Treisman, 1996) it makes sense that activity increases, as the task becomes more demanding. In addition, Treisman (1996) found activity in the parietal areas in tasks requiring binding of features. Parietal

activity has also been associated with eye movement as well as peripheral attention (Corbetta, 1998). As it is proposed that the participant would have to scan the stimuli with the attentional spotlight to decide, whether target is present or not in the hard condition, it seems very sensible that the hard condition will lead to an increase in activity in parietal areas.

The results obtained in this study are generally in line with the theory on the field. However, one should be cautious when interpreting the findings. The primary reason for this concern is the lack of data. As only one participant completed the experiment, the amount of data is not sufficient for confirmation of our hypotheses. In addition to this, an uncorrected p-value was used, which makes the data even more unreliable, as it increases the risk of making Type I errors. The reason for using this threshold was due to the lack of data, which makes it difficult to detect changes in the BOLD signal. The data should therefore not be used to confirm or reject any hypotheses, but instead be used to guide further research. It indeed seems like there is an effect, which can easily be investigated by letting more participants complete the experiment.

If one examines figure 5 and 6, it is easy to see that there is a large difference in the registered BOLD response between the two different sessions. This could be the result of the participant completing the experiment twice, but due to the large difference, it might have something to do with the scanner equipment. Interestingly both rounds display the same pattern of activation for the different conditions (figure 5 and 6). Hence, we can be somehow confident that the data is reliable. It is also very intriguing that the results in figure 6 show the same pattern of activation as the behavioural data in figure 3. This, of course, strengthens our results, since two methods of investigating the same hypotheses show the same pattern. As both methods are indirect measures of, what is actually the measurement of interest, it is especially important that the results are in line with one another.

As there are multiple ways of increasing the difficulty between conditions, further experiments should be carried out to test the difference between variables, such as number of distractors, the relationship between the features of the target and the features of the distractors and the orientation of the target/distractors. This might produce a different pattern of activation in the visual cortex, as the different parts of the visual cortex are thought to process different features such as colour, orientation and motion (Gazzaniga, Ivry & Mangun, 2014, p. 190).

References

- Behrmann, M., Geng, J. J., & Shomstein, S. (2004). Parietal cortex and attention. *Current opinion in neurobiology*, 14(2), 212-217
- Carter, C. S., Macdonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., Noll, D., & Cohen, J. D. (2000). Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, 97(4), 1944-1948
- Corbetta, M. (1998). Frontoparietal cortical networks for directing attention and the eye to visual locations: Identical, independent, or overlapping neural systems?. *Proceedings of the National Academy of Sciences*, 95(3), 831-838
- Gazzaniga, M., Ivry, R. & Mangun, G. (2014). *Cognitive Neuroscience: The Biology of the Mind* (4th ed.). New York, NY: W. W. Norton & Company
- Kristjánsson, Á., Vuilleumier, P., Schwartz, S., Macaluso, E., & Driver, J. (2007). Neural basis for priming of pop-out during visual search revealed with fMRI. *Cerebral cortex*, 17(7), 1612-1624
- Leonards, U., Sunaert, S., Van Hecke, P., & Orban, G. A. (2000). Attention mechanisms in visual search—an fMRI study. *Journal of Cognitive Neuroscience*, 12(2), 61-75
- Pardo, J., Pardo, P., Janer, K., & Raichle, M. (1990). The Anterior Cingulate Cortex Mediates Processing Selection in the Stroop Attentional Conflict Paradigm. *Proceedings of the National Academy of Sciences of the United States of America*, 87(1), 256-259
- RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Retrieved from <http://www.rstudio.com/>
- Schumacher, E. H., Hendricks, M. J., & D'esposito, M. (2005). Sustained involvement of a frontal–parietal network for spatial response selection with practice of a spatial choice–reaction task. *Neuropsychologia*, 43(10), 1444-1455
- The FIL Methods group (2016, October 20). Statistic Parametric Mapping (Version 12). Retrieved from <http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>
- Treisman, A. (1996). The binding problem. *Current opinion in neurobiology*, 6(2), 171-178.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1), 97-136
- Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Attention, Perception, & Psychophysics*, 56(5), 495-500
- Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic bulletin & review*, 1(2), 202-238.