Visual Search Experiment Report

Probability and Statistics I assignment, 2022/23

Daniel Kuchta

Visual Search

Visual search is a cognitive task that involves finding a target object within a complex visual scene. In this task, individuals are required to scan a visual display that contains multiple items to identify a target object that matches a specific criterion. The criterion could be a specific colour, shape, or size of the target object, or it could be a more abstract concept, such as a specific letter or symbol.

Visual search is a fundamental task that humans perform every day, such as finding a specific item in a cluttered room or searching for a friend in a crowded area. The task is also critical in many professions, such as airport security screening, radiology, and surveillance.

The visual search process involves several stages, including attentional selection, feature extraction, and decision-making. Attentional selection involves focusing attention on the relevant aspects of the visual scene, such as the location of the target object. Feature extraction involves extracting specific features of the target object, such as its colour or shape, to help distinguish it from the distractors. Decision-making involves comparing the extracted features with the target criterion and deciding whether the object matches or not.

Visual search can be influenced by various factors, such as the number of distractors, the similarity between the target and distractors, and the complexity of the visual scene.

Overall, the visual search task is a complex cognitive process that involves multiple stages of processing and can be influenced by various factors. Understanding the mechanisms underlying visual search can have important implications for improving performance in professions that rely on this skill and developing technologies that facilitate the task.

Experiment Design

The aim of the experiment is to measure the efficacy of visual search. The participants are searching a screen looking for a pink / among pink \ (or pink \, blue \ and blue /). See Fig. 1 for an illustration of the look of the experiment setup.

In visual search tasks, there are two main types of search paradigms: feature search and conjunction search. They differ in the complexity of the task and the cognitive processes involved.

1. **Feature Search:** In a feature search, the target object is defined by a single, unique attribute or feature that distinguishes it from the distractor objects. This feature could be colour, shape, size, orientation, or motion. For example, finding a red apple among green apples would be a feature search because the target (red apple) has a unique feature (red colour) that sets it apart from the distractors (green apples). Feature searches are typically fast and efficient, often described as "pop-out" searches because the target seems to pop out from the rest of the items. The search time is generally independent of the number of distractor items, indicating parallel processing of the visual field.

2. **Conjunction Search:** In a conjunction search, the target is defined by a combination (or conjunction) of two or more features. For example, finding a red, striped ball among red solid balls and striped green balls would be a conjunction search. The target does not have a unique feature that distinguishes it from all distractors; instead, it is the specific combination of features (red colour and striped pattern) that defines the target. Conjunction searches are typically slower and less efficient than feature searches. The search time tends to increase with the number of distractor items, suggesting a serial or semi-serial processing of the visual field.

Fig. 1 Example of screen used in a an experiment

The experiment tries to measure the effect of the set size (how many objects are there on the screen), target presence and the level of homogeneity of the set (feature/conjunction search). A participant should press F when the target is present on the screen and J otherwise.

There were 3 participants in total, each one taking the test in a calm, darkened room on a Macbook Air with a screen resolution of 2560 x 1440 on a 24’’ external display.

Results

The performance of the individuals varied quite significantly as we can see in Tab. 1. Partially, this might be explained by a different strategic approach of the participants - part. 0

seems to be focused primarily on the accuracy while part. 2 invested their efforts into speed of execution rather than accuracy.

Participant ID	0	1	2
Average RT	1642.83	1040.33	806.01
Accuracy	0.9875	0.9593	0.9156

Tab. 1 - Overview of performance

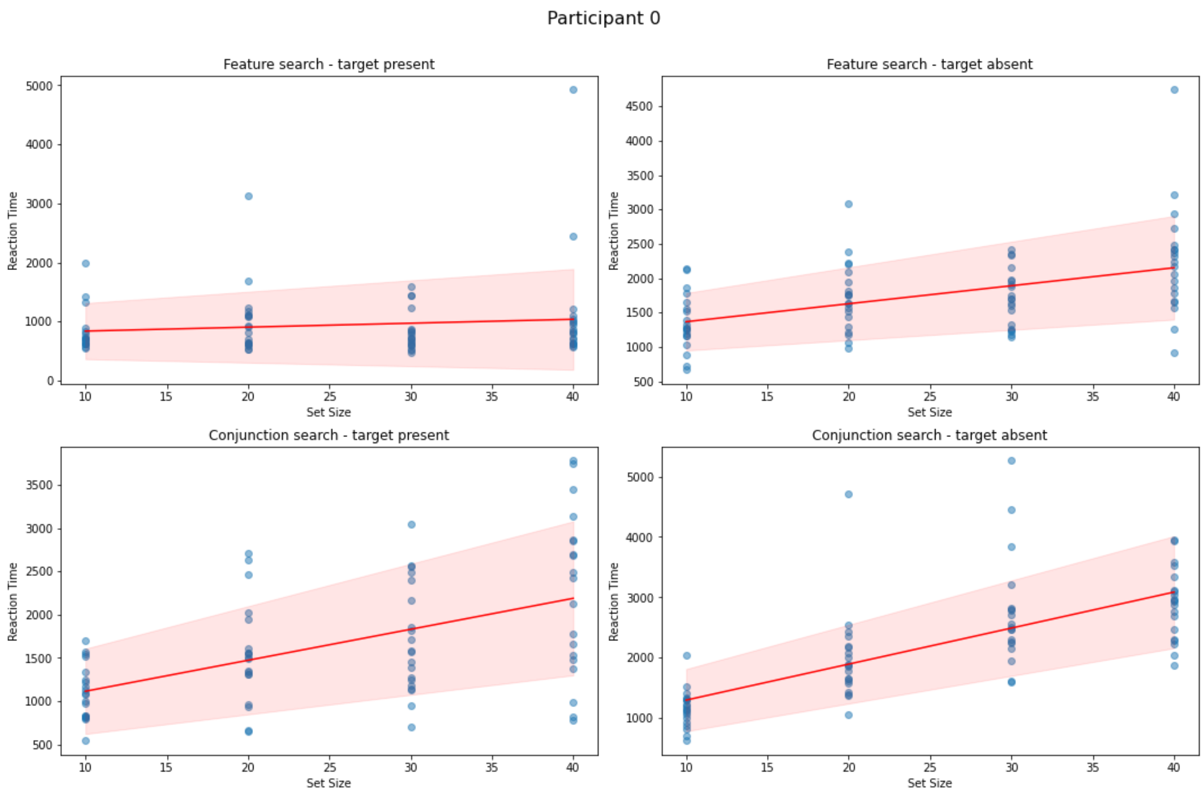
To compare their efficiency in a more detailed manner a table containing regression slope for set size under different conditions follows. The data from the participants were fitted with a Linear Regression model trained using Ordinary Least Square method.

Search type	Target present	Set Size slope	Intercept	Confidence interval - Set Size slope	Confidence interval - Intercept	Part ID
conjunction	present	35.775	758.013	[22.568, 48.981]	[397.89, 1118.135]	0
conjunction	present	8.674	875.923	[1.327, 16.021]	[678.902, 1072.944]	1
conjunction	present	0.987	840.111	[-5.927, 7.901]	[651.586, 1028.637]	2
conjunction	absent	59.661	699.7	[45.879 73.444]	[322.727 1076.673]	0
conjunction	absent	33.361	599.5	[13.17, 53.552]	[46.546, 1152.454]	1
conjunction	absent	22.193	484.897	[15.605, 28.78]	[309.873, 659.921]	2
feature	present	6.65	771.658	[-5.971, 19.271]	[424.128, 1119.187]	0
feature	present	-9.318	944.991	[-19.119, 0.482]	[675.796, 1214.185]	1

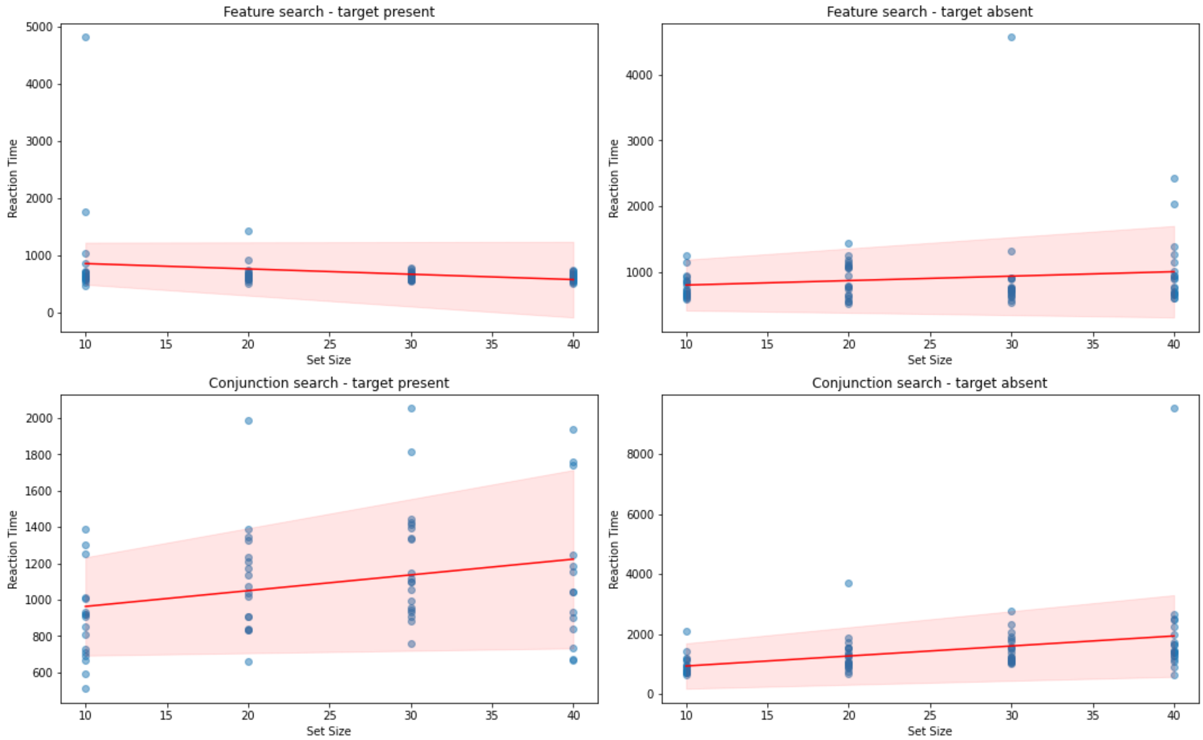
feature	present	-0.605	625.447	[-3.51, 2.301]	[545.772, 705.123]	2
feature	absent	26.162	1104.825	[14.982, 37.341]	[798.67, 1410.98]	0
feature	absent	6.781	737.275	[-3.517, 17.078]	[455.257, 1019.293]	1
feature	absent	13.472	445.38	[6.894, 20.051]	[264.235, 626.526]	2

Tab. 2 - Overview of performance

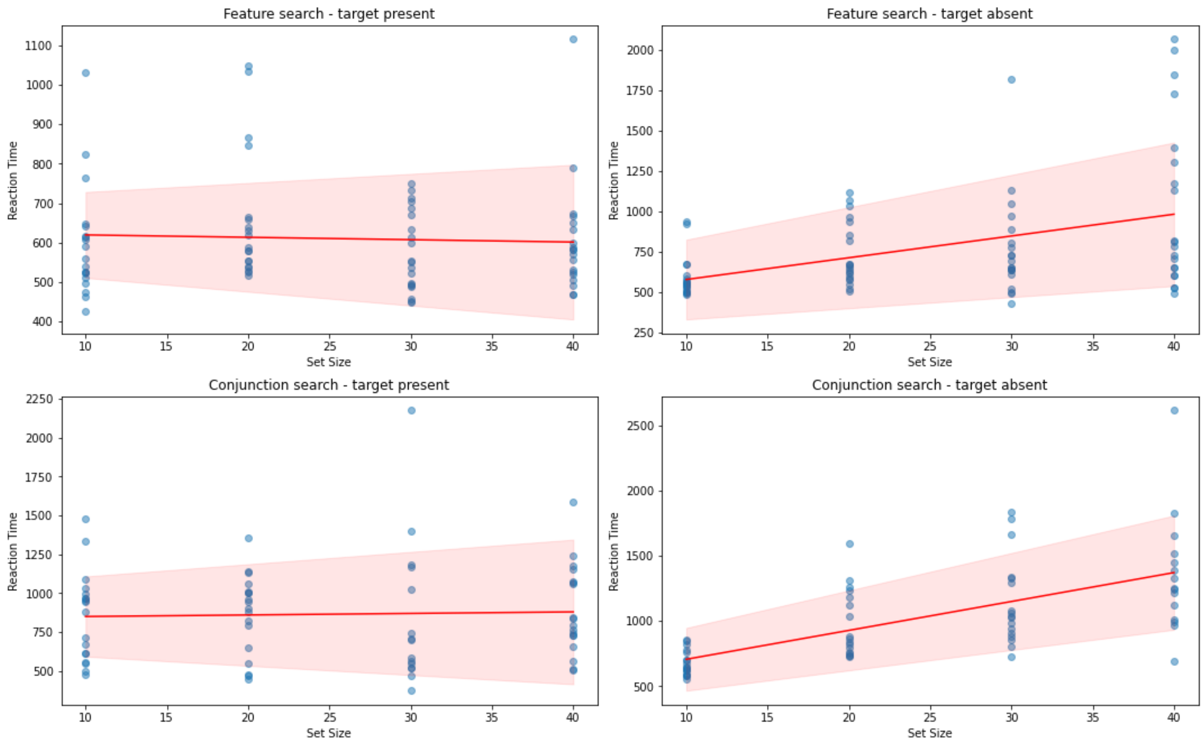
For better visual clarity, here is an overview of data points and regression lines for each individual participant:



Participant 1



Participant 2



Search type	Target present	Set Size slope	Intercept	Confidence interval - Set Size slope	Confidence interval - Intercept
conjunction	present	16.224	818.217	[8.95, 23.499]	[620.901, 1015.534]
feature	present	-1.09	782.212	[-6.738, 4.559]	[627.016, 937.409]
conjunction	absent	40.28	576.612	[29.835, 50.725]	[293.221, 860.003]
feature	absent	15.301	768.178	[7.858, 22.743]	[563.988, 972.368]

Tab. 2 - Overview of performance across all participants

The data suggests that conjunctions of features can not be processed in parallel and need to be processed in a serial manner with deliberate attention. Relaxing this to only a single feature significantly improves the reaction time of the subjects (as expected).

However, when we look at the coefficients in Tab. 3 we can see that difference between conjunction/feature is less significant when the target is absent - this is due to the fact that absence of the target requires checking most of the individual items on the screen carefully while finding the target (when it's present) is on average much faster - in the latter case we need only existential predicate to be fulfilled while in the former we require a universal one.

A technical note:

The confidence interval for the parameters of a linear regression model is calculated using the standard errors and the t-distribution. The standard error measures the accuracy of the coefficient estimate. It is calculated as the square root of the variance of the coefficient estimate.

The t-distribution is used to construct the confidence interval because the true standard deviation of the coefficient estimate is usually unknown, and the t-distribution approximates the normal distribution when the sample size is large.

The confidence interval is calculated as:

```
[estimate - t_value * standard_error, estimate + t_value * standard_error]
```

where `estimate` is the estimated value of the coefficient, `t_value` is the critical value from the t-distribution for the desired confidence level (for example, 1.96 for a 95% confidence interval), and `standard_error` is the standard error of the estimate.

Conclusions

Overall, the data seems to confirm the common sense hypothesis that more features will make it harder for the participant to find the target (when it’s present), effectively blocking the ability to process the visual stream in parallel using preattentive preprocessing while the effect of this is not so significant in case of the absence of the target which requires careful inspection of most the elements on the screen (or at least a significantly larger portion when compared to the situation with a target present). The findings are consistent across all 3 subjects but a larger sample would be required to be able to say this confidently.