An ERP study of visual search showing early detection of feature conjunctions

Patrick Bourke[1], Elena Georgiadi[2], Steve Wilson[1], Vince di Lollo[3] & Mario Liotti[3]

University of Lincoln, UK[1], University of Nottingham, UK[2], Simon Fraser University, Canada[3]

UNIVERSITY OF LINCOLN

Introduction

• Visual search is the process of deciding if a specific object is present in a cluttered visual scene.

• This requires that input be matched to a target representation. How is this achieved ?

• The initial processing of input is thought to be by the activation of specific feature maps. For example, a green feature map would encode the presence and topographical location of green items in the input. This map encodes no other feature of the objects.

• In theories of visual search, how targets are represented and how the matching to visual input is implemented is underspecified.

• One possibility is that the matching of input to a target representation is achieved by lowering thresholds in the relevant feature map. For example if the task will be to decide if there is a green object - thresholds in the green detector map would be lowered. Green input would then be processed earlier than other input. In a race model this would allow the target to be rapidly detected, if present.

• As the feature maps are thought to cross-refer via a master map of location, the same mechanism could exist for objects which are composed of a conjunction of features e.g. green X. Thresholds in the two or more relevant maps would be lowered.

• If correct there should be evidence early in processing of a difference between displays with a target present in the visual input and those where the target is absent from the input.

• This should be the case for objects defined by a single feature and objects defined by a conjunction of features.

• Suggestive evidence comes from neurophysiology and brain imaging studies.

• ERP studies have reported evidence of activation of a location map and single feature maps, consistent with the implementation of target matching early in visual processing.

• In the current study the neurological signature of matching of visual input to the target representation was studied by varying (1) the presence of the target, (2) the complexity of the target (1, 2 or 3 features).

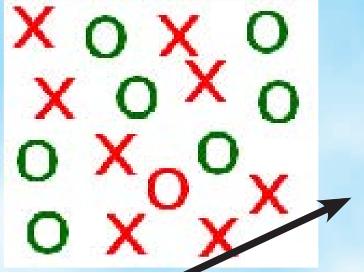
• Whether the representation is coded visually or more generally was also studied by having the search task done either alone or concurrently with a story listening task

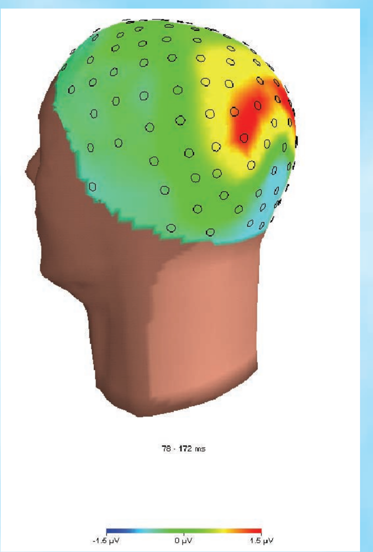
Method

Subjects
17 paid volunteers were tested. Their mean age was 25.8 with a standard deviation of 8.05. There were 3 males. All were right handed.

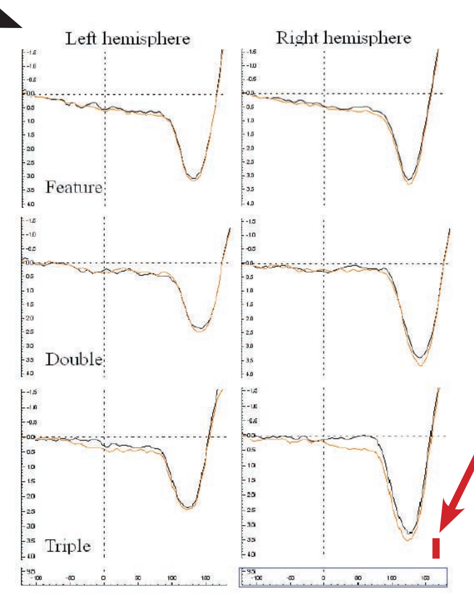
Tasks
There were two tasks which were done alone and in combination. The first task was a text memory task in which participants were played audio recordings of a selection of monologues on various issues or stories. The second task was a visual search task of which there were three versions, a feature search, double conjunction search (classic conjunction search) and triple conjunction search. In the feature search, participants searched for a red O amongst green Xs. Displays could be of all large or all small stimuli. In the double conjunction search participants searched for the target - a red O, amongst two types of distracters - a green O and a red X. In the triple conjunction search participants searched for the target – a large red O amongst distracters - a large green X, a small green O and a small red x. A block started with a screen of written instructions. If the task was feature search then the instructions asked participants to search for ‘a red letter’ and to click the left button if was present and to click the right button if it was not. If the task was a double conjunction then the instructions asked participants to search for ‘a red circle’. If the task was a triple conjunction then they were instructed to search for ‘a big red circle’

ERP Methods
Brain electrical activity was recorded continuously using a standard 128 electrode cap (Biosemi Activa 2 system) and six additional electrodes (four eye movement channels plus the two mastoids), digitized at a sampling rate of 256 Hz. A Band-pass filter (0.5-70 Hz) and a Notch filter (50Hz) were applied offline. Correction for eye movements employed the Gratton & Coles method (using a bipolar derivation between electrodes above the left orbit and below the right orbit). All other artefacts were automatically rejected on a single channel basis when voltage over a 200 ms period exceeded 100 μV or was less than 0.50 μV. After this, the EEG was referenced to the average reference. The continuous EEG recordings were segmented in epochs from -200msec before stimulus onset to 2,000 ms after stimulus onset. Subject ERP averages (time-locked to stimulus onset) were computed for each combination of Dual Condition (Alone, Dual-Task), Visual search condition (Feature, Double, Triple), Target (Present, Absent) and Set Size (4, 8 or 16 items). Only correct hits were included. ERP amplitudes were aligned to a 200 msec pre-stimulus baseline period. Grandaverages were then computed across subjects. To facilitate visualization, scalp voltage topographic distributions were obtained using spherical spline interpolation (Perrin et al, 1989).





P1



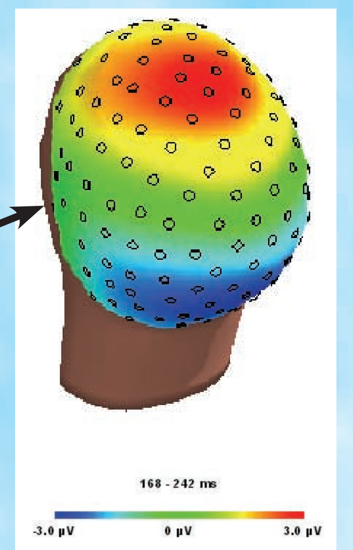
ERP waveforms for target present (black) and target absent (orange) for each visual search condition. Activity recorded from parietal electrode sites P3 and P4. Note greater amplitudes for target absent, particularly over right hemisphere.

Behavioural Results

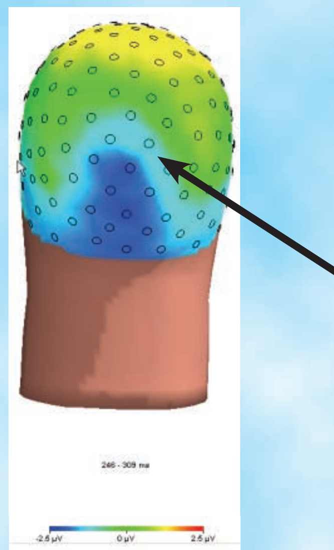
Story Memory
Performance in the story memory task was better when it was executed alone than in conjunction with the visual search task, as measured in terms of ideas recalled, Friedman 2 = 20.65, p =.004. There was however no effect of type of visual search task (feature, double or triple conjunction) on ideas recalled.

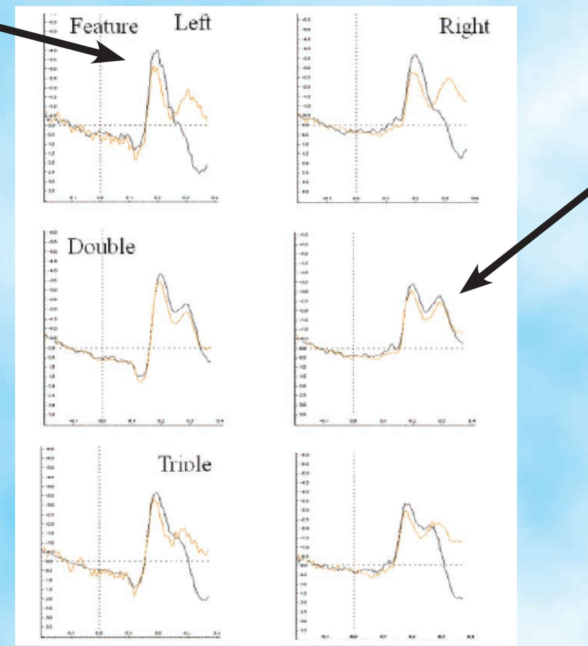
Visual Search Task
All four factors produced significant main effects. (1) There was a main effect of target, with ‘target present’ being faster than ‘target absent’ (408 vs. 517msec). (2) There was a main effect of visual search condition, feature (440 msec), triple (474msec) and double (592 msec). There was a significant interaction of target by visual search condition, with the effect of the presence or absence of the target being greatest in the double conjunction search, less in the triple conjunction and not present in the feature search. (3) There was a main effect of whether the task was done alone or under dual task conditions, F (1, 16) = 9.24, p = .008, with faster reaction times when visual search was done alone (493 msec vs. 512), Error data were consistent with the reaction time results.

N2a

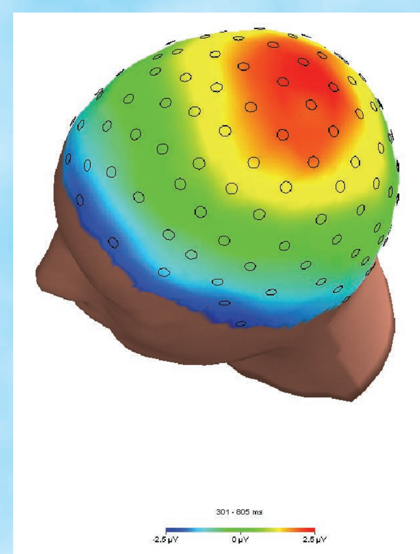


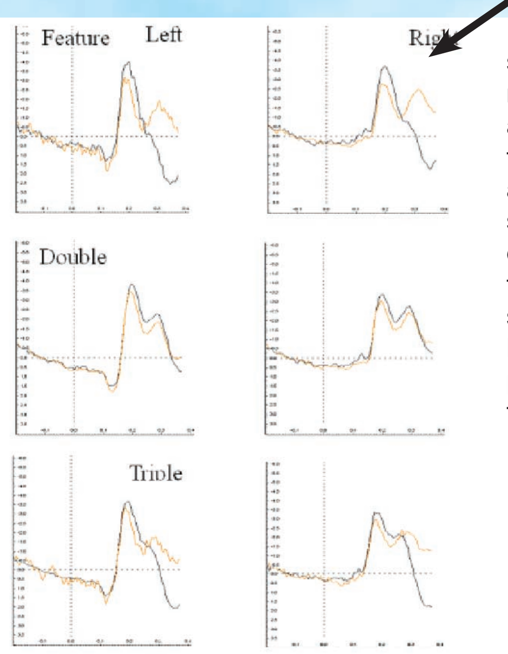
N2b





P3





ERP waveforms showing activity recorded from anterior scalp (left frontopolar site) and posterior scalp (right centroparietal site) for the three visual search conditions. In black, target present; in orange, target absent.

ERP Results

Inspection of the grandaverage waveforms and sequential topographical maps for Feature, Double and Triple trials all revealed the typical succession of P1, N2, and P3 components of the visual ERP. Due to the use of an average reference, the ERP waves displayed bipolar scalp distributions.

After careful inspection of grand average waveforms and sequential topographical maps for each experimental condition, time windows around the peak amplitudes of each ERP component were selected for further analysis. For the P1 analysis it was evident that the grand-average peak latencies for the Feature and Double conditions were shifted relative to that for the Triple condition, (mean peak latencies for feature =128 msec, double = 130 msec, triple = 119 msec). Therefore, different time windows were chosen: Triple: 85-155, Feature and Double 100-170. Time windows of 170-240 msec and 250-330 msec where chosen for the N2a and N2b, respectively. As with the P1, inspection of the grand averages showed that the P3 waves were similarly shifted in time (mean peak latencies for Feature= 421 msec, Triple= 441 msec, and Double= 476 msec). To account for such P3 latency differences, distinct time windows were selected for each type of visual search, i.e. for Feature: 320-520 ms; for Triple: 340-540 ms, and for Double: 380-580 msec).

For all ERP waves the dependent variable used for statistical analysis was mean voltage amplitude. Based on the inspection of scalp voltage distributions, regions of interest (ROIs) were selected by collapsing together neighbouring electrode sites of maximum amplitude. For the P1 analysis, two ROIs were chosen (collapsing three sites each), one left parieto-occipital, and one right parieto-occipital. For the N2a analyses, three more ventral ROIs were selected (collapsing four sites each), one left parieto-occipital one right parieto-occipital and one midline frontocentral. For the N2b analysis a medial parieto-occipital ROI was created by combining activity at three sites. Finally, for the P300, four ROIs were chosen, one left centro-parietal, one right centro-parietal, one left frontopolar, and one right frontopolar.

P1 Results

70 – 170 msec after display presentation a clear peak in amplitude was seen. The size varied. There was a main effect of target, F (1, 15) = 16.24, p= .001, with higher amplitudes seen when the target was absent (no-target = 1.56 μV ± .29, target =1.39 μV ± .28). There was a main effect of visual search condition, F (2, 30) = 5.71, p= .008 (Feature = 1.27 μV ± .30, Double 1.61 μV ± .27, Triple 1.55 μV ± .31). The contrast of Feature vs Double search was the only significant difference determined by Bonferroni comparisons. There was a main effect of whether the search task was done alone or concurrently, F (1, 15) = 5.03, p= .04 with higher amplitudes when the task was done concurrently (concurrent = 1.53 μV ± .28, alone = 1.42 μV ± .29). There was no main effect of hemisphere. Visual search condition interacted with target and hemisphere, F (2, 30) = 5.76, p =.008, To clarify these interactions the visual search conditions were considered separately. For Feature search, there was a lower amplitude when the target was present (1.20 μV ± 0.31) than when it was absent (1.35 μV ± .29), F (1, 15) = 5.46, p= .034. For Double conjunction search, there was again a main effect of the presence of the target with lower amplitudes for target present (1.53 μV ± .27) than target absent (1.69 μV ±.27), F (1, 15) = 8.28, p = .011. However, here the effect was significantly larger over the right hemisphere, (target present =1.53 μV, target absent =1.78 μV) than the left (target present =1.53 μV, target absent =1.60 μV), F (1, 15) = 7.39, p = .016. For Triple conjunction search, the effect of the presence or absence of the target showed a similar trend to that of the double but this did not reach significance F (1, 15) = 3.81, p = .070, again with larger amplitudes seen when the target was absent (1.65 μV ±.33) than when it was present (1.45 μV ± .29).

N2a Results.

170 – 240 msec after stimulus presentation, there was a main effect of target, F (1, 16) = 41.72, p = .000, but unlike in the P1 the order is reversed with higher amplitudes seen when the target was present, (target present = -2.51 ± .43, target absent = -1.99 ± .40). This interacted with two other factors. Firstly, it interacted with hemisphere, F (1, 16) = 6.72, p=.020 with higher amplitudes been seen over the left hemisphere posterior site, when the target was present (left present = -2.59 ± .51, left absent = -1.99 ± .46, right present = -2.43 ± .43, right absent = -1.99 ± .40).

Secondly, the presence of the target interacted with the type of search being performed, F (2, 32) = 4.72, p = .016. The interaction appears to arise due to the difference in the extent of the effect, being largest for feature, followed by triple followed by double.

In addition there was an interaction between type of search and whether the task was done alone or concurrently, F (2, 32) = 3.60, P= .039. Separate ANOVAs showed that only in the feature search condition were amplitudes different (concurrent -2.11 ± .49, alone -2.55 ± .42) when the search task was being done alone rather than concurrently, F (1, 16) = 6.27, p = .024.

N2b Results

250 -300 msec after target onset, there was a separate wave to the N2a. There was a main effect of visual search condition, F (2, 32) =10.14, p= .000. There was also a main effect of target, F (1, 16) = 30.72, p =.000. These two factors interacted, F (2, 32) = 46.12, p= .000. This reflects the fact that a wave clearly seen when the target is present in double conjunction search is completely absent in the feature condition and only weakly present in the triple conjunction search condition. In all 3 search conditions the wave is seen when the target is absent.

P3 Results

300 -800 msec after stimulus onset, there was a main effect of target, with target present displays producing greater P3 amplitudes than target absent displays, F (1, 16) = 46.54, p=.0000, (Target present= 3.79 ± .30 μV; Target Absent=3.13 ± .34 μV). This effect varied in strength with anatomical location. The highest amplitude differences between target presence and target absence were over anterior right frontopolar scalp. There was a main effect of Visual search condition, F (2, 32) = 12.98, p<.0001. Bonferroni comparisons showed a significant difference between triple conjunction (3.79 ±.36 μV) and double conjunction search (3.07 ± .31 μV), p= .001, and between feature search (3.51 ± .31 μV) and double conjunction search, p= .036

Anterior		Posterior		Anterior		Posterior	
Present Left	Absent left	Present left	Absent left	Present right	Absent right	Present right	Absent right
3.99	3.19	3.70	3.24	3.67	2.55	3.79	3.53
(.41)	(.40)	(.28)	(.38)	(.28)	(.28)	(.39)	(.45)

Conclusions

Within 70 - 170 msec after target presentation, a clear difference is seen between cluttered displays in which there is a target compared to those where there is no target.

These results show early matching of input to target, even objects that are defined by a conjunctions of 2 features.

Despite the early detection of targets this is insufficient to allow a response. Further neurological processing occurs whose time and approximate location can be identified but which is less obviously tied to specific cognitive functions.