**Nationaal Lucht- en Ruimtevaartlaboratorium**
National Aerospace Laboratory NLR

NLR-TP-2004-010

# Studying the use of colour for military cockpit displays

H.G.M. Bohnen

This report has been based on a paper presented at the Annual conference of the Ergonomics Society, Swansea, United Kingdom, 14 - 16 April 2004.




| | |
|---|---|
| Customer: | National Aerospace Laboratory NLR |
| Working Plan number: | AT.1.F.2 |
| Owner: | National Aerospace Laboratory NLR |
| Division: | Air Transport |
| Distribution: | Unlimited |
| Classification title: | Unclassified |
| | January 2004 |


| Approved by author: | Approved by project manager: | Approved by project managing department: |
|---|---|---|
| | | |

**Summary**

Three successive experiments with increasing task fidelity were conducted to study the effect of colour coding on military cockpit displays. The experiments were performed using a PC-based tabletop arrangement, a part-task flight simulator and a full-task flight simulator.
The single-task experiment showed that subjects could better detect and identify display objects in multi-chrome display versions than in a monochrome display version. Significant performance differences were also found among the multi-chrome display versions. The experiments with higher task fidelity showed smaller (objective) performance effects.
More degrees of freedom in behaviour with an increase in task fidelity could be an explanation for the latter phenomenon. Therewith behaviour is more difficult to assess. Furthermore, subtle differences in detection performance could get bogged down as many tasks were performed simultaneously.

## Contents



(10 pages in total)

## 1 Introduction

Due to the large amount of information needed by a fighter pilot to perform his task, military cockpit displays are receptive to clutter. Colour capability is an option to accommodate high symbol densities. Grouping information by colour, such as red for threats and green for friendly aircraft helps the pilot to sort the information (Murch and Taylor, 1986). Colour also increases the rate of information transfer to the pilot. As a consequence search times for targets decrease, resulting in more accurate target detection.
Multi-chrome displays are developed for tracking radar display pages and navigational display pages (Fig. 1) present on two separate multifunction displays. Head-down in the cockpit these pages provide a large quantity of tactical and navigational information. Colour coding is applied as a display feature redundant to symbol shape, especially for display objects that have a variable position on the displays. The tracking radar display provides radar, friend or foe identification and data link information. For air-to-air missions the display presents information for single or multiple targets. The navigational display presents tactical and navigational data using a lateral view. The tactical overview consists of data link communication and friend or foe information. The navigational overview provides routes and waypoint information with a 360-degree coverage area around the aircraft.
The starting point of the design of the multi-chrome display versions is the ability of a pilot to quickly detect and search for tactical information such as the presence and position of friendly, unknown and hostile aircraft symbols and data link information. The design of two multi-chrome displays is primarily based on subject matter expertise. These display versions are tailored to the design of the aircraft. For the design of three other multi-chrome displays, theoretical considerations for the use of colour are applied as well as a model of an operator in his working environment. In addition, standards and regulations regarding the use of colour on cockpit displays are taken into consideration. Each version emphasises some of the design rules that often conflict.
Three successive experiments with an increasing level of task fidelity are performed. The rationale is an evolving understanding of the effects of colour design choices. First, the colour effects are studied in a simple task environment followed by tests in increasingly realistic task environments. Changing the task of the subject from the simple, PC-based search and identification task to a realistic fighter pilot task increases fidelity. The major aim is to assess whether colour shortens target search times and whether that effect is greater at higher display densities. The second experiment is performed in a part-task flight simulator. Several missions are flown by fighter pilots. Pilot task related performance parameters such as head-out time and

time-to-a-kill and mental workload are used to study the effect of colour coding. The final experiment is conducted in a full–task flight simulator to validate the findings of the earlier experiments. The same means as described for the second experiment are used to assess pilot performance.

## 2 Methods

### 2.1 Subjects

Six subjects participated in each experiment. University students performed the tabletop experiment. The part-task and full-task flight experiments were performed by (former) fighter pilots. All subjects had normal vision. None of the subjects participated in more than one experiment.

### 2.2 Apparatus and display versions

The tabletop arrangement consisted of a PC plus screen. The part-task flight experiment made use of a fighter pilot station. The multifunction displays were equipped with CRT's whereas an elementary outside view was generated on which a virtual head-up display was projected. The full-task flight experiment made use of a real cockpit equipped with LCD's. A dome was used to project the out-of-the-window view.

The monochrome version of the displays consisted of three intensities of green. The use of green (for border text and some symbols) and yellow and white for other symbols characterised the first expert-based display version (I). The use of white (text and some symbols) and blue and yellow (symbols) the second one (II). In one rule-based display version priority was given to stereotype ideas people had about the meaning of colours. Furthermore, text and several symbols were displayed in white (III). Another rule-based version concentrated on consistency in the use of colour for text (green) with the use of colour on other cockpit displays (IV). In addition, yellow, blue and white were used for The last rule-based version (V) took advantage of the design principles of display versions III and IV and text was displayed in green. All five multi-chrome display versions contained intensity modulated colours.

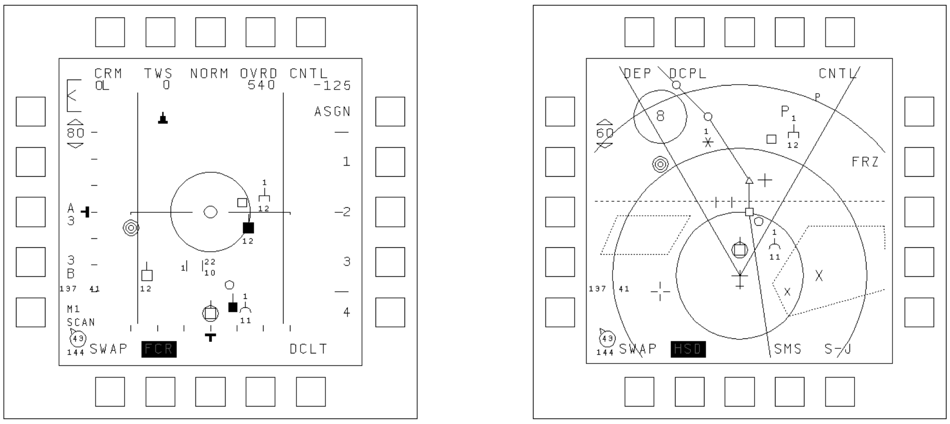


*Fig. 1 Tracking radar display (left) and navigational display (right)*

### 2.3 Experimental procedures

Subjects were extensively trained in the use of the apparatus (PC-configuration and/or flight simulators) and studied one by one the colour display versions. Pilots also studied the meaning of all display objects. Flights to get used to each colour display version preceded the experimental flights. All three experiments used a within-subjects design with colour coding (monochrome and multi-chrome display versions) and display density (number of targets) as factor. The number of colour display versions ranged from five in the single-task experiment and four in the part-task flight experiment (version V excluded) to two in the full-task flight experiment (version II, IV and V excluded).

## 3 Results

### 3.1 Single-task experiment

Search times for targets were higher for the monochrome display (1757 ms) when compared to all multi-chrome display versions (1304 ms). The maximum difference of 517 ms was found for version III. Among the different multi-chrome displays significant differences were found too. One of the subject matter expert-based display versions (II) and one of the rule-based display versions (V) showed the worst performance. No effect of display density was found. Evidence was present for the fact that search times increased for targets with a colour identical to fixed-position display objects. This was most clear for targets coded green in cases where text was coded green too (I/IV and V). Counting targets in an exhaustive search measurement took about 4000 ms for the monochrome and about 3000 ms for the multi-chrome display versions.

### 3.2 Part-task flight experiment

The percentage of killed aircraft was statistically not different for the various display versions. However, a difference of 5% in favour of the multi-chrome displays was present. The time interval from target pop-up to target hit was slightly longer for the multi-chrome display version designed for consistency (IV) than for the other display versions (138 s versus 130s). Colour coding did neither influence the amount of viewing time spent by the pilot at one of the displays nor the frequency with which the pilot looked at one of the cockpit displays or outside the cockpit nor mental workload. The monochrome display was rated less useful than the multi-chrome display versions. This was independent of the tasks of the pilot. The display that applied colour stereotypes and used the colour white for text (III) was rated most useful. Likely because grouping was applied properly. The coding of hostile objects was rated less useful in the expert-based display versions (I/II) than in the rule-based display versions (III/IV). The multi-chromatic display that applied stereotypes (III) was preferred to the monochrome display version.

### 3.3 Full-task flight experiment

Colour coding did not affect any of the assessed performance parameters, neither the pilot's viewing behaviour nor mental workload. Both multi-chrome display versions (I/III) were rated as significantly more useful than the monochrome display and both were rated equally useful. No differences were found between both multi-chrome display versions. Especially the coding of target objects (5 in total), mainly done by allotting yellow was rated in favour of the multi-chrome display versions. Coding of cursors by white (III) is more useful than coding by green such as applied in the monochrome display version. All pilots preferred the multi-chrome display versions to the monochrome display version. Half of them preferred version I the other half version III.

## 4 Discussion

The results of the single-task experiment clearly show that search times for 'target' objects are shorter when colour coding is applied (about 25%). This effect was irrespective of display density, which confirms the parallel processing theory of colour. Also in line with earlier findings are the search times for 'target' objects that are longer in the presence of display objects with the same colour.

At the $p<0.05$ statistical significance level, colour coding did not influence the behaviour of the pilots in the part-task experiment. However, colour coding has influence on behaviour when considering $p<0.30$ (Wickens, 1998). With colour coding, pilots spend slightly more time looking outside and less on those displays. This indicates that it is easier to retrieve information.

Colour coding according to stereotypes (e.g. friendly aircraft in green) together with text in a neutral colour (white) makes that targets can be distinguished more easily. Colour coded displays are sampled less frequently at higher symbol densities, which suggests that search times are shortened (also see Backs and Walrath, 1992). Disregarding, even when considering $p<0.30$ colour coding has no effect on pilot workload. The less strict p-level also does not change the earlier conclusion regarding no effect at all of colour coding on performance in the full-task experiment.

The approach to perform three successive experiments with an increase in task fidelity did not result in convergence of objective evidence. The opposite is true as performance effects decrease with an increase in task fidelity. The pilots that performed the part-task experiment also conducted the single-task experiment for training, and their results are in line with those of the students. This brings us to the question why only some small performance effects were found in the part-task experiment. A theoretical explanation is that colour is not always the best-recalled coding format under high task load. The utility of colour is task specific, being suitable for segmentation, but not identification. Segmentation may be viewed as a primarily perceptual use of colour. Identification is dependent on central cognitive processing. According to Dudfield and Hughes (1993) colour serves a purpose in tasks concerned with searching for or detecting information at unknown spatial locations in complex and cluttered displays. On the other hand, colour coding is often of less value in monitoring tasks or identification tasks, in which the spatial location of the target is cued to the operator, as is partly the case in the part-task experiment. An important difference between the single-task and the multiple-task experiments is of course that the first requires only the performance of one single task whereas the latter require simultaneous performance on multiple tasks. This explanation is in agreement with suggestions that the use of colour is well suited for a predominantly perceptual task -as the single-task experiment - but may not confer an advantage for concurrent tasks. Furthermore, the beneficial effects of colour coding on central processing may be limited in high-density displays under time-constrained conditions (Backs and Walrath, 1992) as present in the multiple-task experiments.

In both multiple-task experiments pilots favoured the use of colour, most likely due to esthetical appeal. Its likely that colour also positively influences the pilots' perception of the overall system quality.

Reising and Aretz (1984) state that displays that are multifaceted in nature will benefit from the use of colour and that one key to potential benefits of colour lies in time-constrained conditions. They conducted a high-fidelity flight experiment to study colour on cockpit displays and found no differences in performance. One of their explanations is that colour had only a subtle effect on the display they studied. This explanation does not hold for the present study as the tracking radar and navigational displays are far more complex than their display was. Reising and Aretz state that these types of studies should apply more than one metric to tease out subtle performance differences. With the increase in fidelity from single to multiple-task environment the means to assess performance in this study increased in number and complexity. In the part-task experiment few performance benefits were found. In the full-task experiment no effects were present even though performance was assessed in all its dimensions. Therefore, the main relevance of gathering performance data in the multiple-task experiments is maybe that colour did not prove to be counter-productive for the tasks of the pilots. Together with the positive effects found in the single-task experiment and the appreciation of the pilots, this may be sufficient reason to introduce colour in the fighter cockpit.

## 5 References

Backs R.W. and Walrath L.C. 1992, Eye movement and pupillary response indices of mental workload during visual search of symbolic displays, *Applied Ergonomics*, **23**, 243-254

Dudfield H.J. and Hughes P.K. 1993, *The use of redundant colour coding as an alerting mechanism in a fast jet head-up display*. Technical report: DRA/AS/FS/TR 93018/1.

Murch G.M. and Taylor J.M. 1986, *Effective use of colour on avionics displays*. Human Factor Research Textronix, Inc.

Reising J.M. and Aretz A.J. 1984. Colour coding in fighter cockpits: it isn't black and white. *Proceedings on the second Symposium on Aviation Psychology*, 1-7

Wickens C.D. 1998, Common sense Statistics, *Ergonomics in Design*, **6** (Santa Monica), 18-22