Running head: CATEGORICAL EFFECTS IN COLOUR SEARCH

Categorical effects in children's colour search: a cross-linguistic comparison.

Christine A. Daoutis*, Anna Franklin, Amy Riddett, Alexandra Clifford, and Ian R. L. Davies

University of Surrey, UK

*Address for correspondence:

Christine Daoutis

Department of Psychology

University of Surrey

GU2 7XH

E-mail: c.daoutis@surrey.ac.uk

Categorical effects in children's colour search: a cross-linguistic comparison.

Abstract

In adults, visual search for a colour target is facilitated if the target and distractors fall in different colour categories (e.g. Daoutis, Pilling, & Davies, 2005). The present study explored category effects in children's colour search. The relationship between linguistic colour categories and perceptual categories was addressed by comparing native speakers of languages differing in the number of colour terms. Experiment 1 compared English and Kwanyama (Namibian) children aged 4 to 7 years on visual search, using target-distractor pairs (blue-green, blue-purple, red-pink) for which the Kwanyama did not have distinct names. The presence of a category advantage in the English, but not in the Kwanyama, suggested that linguistic boundaries may affect search performance. Experiment 2 examined visual search performance in the green-yellow and the blue-green region, in English and Himba (Namibian) 6-year-olds. The number of distractors was varied to assess search efficiency. Cross-category search was more efficient than within-category search in the English group, but this advantage was absent in the Himba. Increasing the number of distractors affected search speed in the English group, but not in the Himba. Overall these findings suggest cross-language differences in categorical effects on colour search, but also in the way the children performed the search. The nature of the category effect in search is discussed with respect to these findings.

Colour perception is categorical. The physically continuous spectrum appears as seven or so qualitatively different fuzzy edged bands that are labelled in English by basic colour terms such as *red, green, yellow* and *blue*. Two colours separated by a category boundary seem more different than an equally spaced within-category pair, a marker for categorical perception (CP). While these categorical effects are well established, the answers to many fundamental questions about colour categorisation are unknown. Are perceptual colour categories innate or acquired? Where in the perceptual-cognitive processing stream do they affect colour judgements? Are they invariant structures, solely determined by the physiology of our visual systems, or are they flexible? What is the relationship between perceptual colour categories and linguistic colour categories? Do linguistic categories map onto pre-existing perceptual categories, or do linguistic categories create perceptual categories?

The current study addressed these issues in three ways. First, the effect of colour categorisation in visual search by young children was explored. In adults, visual search for a colour target is facilitated if the target is in a different category to the distractors (Daoutis, Pilling, & Davies, 2005; Franklin, Pilling & Davies, 2004; Kawai, Uchikawa, & Ujike, 1995), and this is consistent with categorical influences on early visual processes. Here, we explored whether such category effects are found in children. Second, the origin of these categorical effects was explored by comparing native speakers of languages whose linguistic categories for colour differ. If CP is due to language learning rather than to hardwired perceptual structures, then CP in search should co-vary with linguistic category structure. Third, CP was explored developmentally across two language groups. If CP is affected by language learning, then it should strengthen as linguistic competence grows and differences between the two language groups should increase with age.

Categorical perception

Discrimination of pairs of stimuli separated by a category boundary is more accurate and faster than discrimination of equally separated pairs in the same category. This pattern, which is the signature of categorical perception (Harnad, 1987), is illustrated in Figure 1. It represents six stimuli, three greens and three blues (G3, G2, G1, B1, B2, B3) with the category boundary between G1 and B1. The distances among the stimuli represent distances in a 'colour space' such as Munsell or CIE (see: Hunt, 1987; Davies & Franklin, 2002)[1]. The key test of CP is to compare performance on some task of cross-category pairs such as G1-B1, with equally separated within-category pairs (B1-B2, or G1-G2). If performance is better with the cross-category pair than with the within-category pair, this is evidence for CP. For instance, using a same-different task it is faster to decide that B1-G1 are different than to decide that B1-B2 are different (Bornstein & Korda, 1984; Özgen & Davies, 2002; Pilling, Wiggett, Özgen & Davies, 2003).

INSERT FIGURE 1 ABOUT HERE

Equivalent results have been found using a two-alternative forced choice (2-AFC) method. Here, a target colour is followed by two test colours, one identical to the target and one different. If the target is B1, then the key comparison is B1-B2 with B1-G1. The latter cross-category choice is more accurate and faster than the within-category choice (Pilling et al., 2003; Roberson & Davidoff, 2000). Similarly, the experiments we report here compare visual search for a target among an array of same category distractors (e.g. G1 target among G2 distractors) with search when the target and distractors are in different categories (G1 among B1 distractors). As the similarity of target and distractors is the main determinant of difficulty, and the within-category and cross-category separations are matched, any difference between the two conditions implicates the effect of categorical differences (Daoutis et al., 2005; Franklin et al., 2004; Kawai et al., 1995).

The studies cited above converge to show that colour CP is a common effect. However, it is less clear what the origin of the effect is, as the characteristic signature is consistent with quite different explanations. These are considered next.

Perceptual accounts of CP

First, as the term implies, CP could be due to 'warping' of perceptual colour space. Perceptual distances in the boundary region are stretched relative to equivalent within-category regions, making cross-category discrimination easier than within-category discrimination. If CP is due to perceptual structure, this could be hardwired and driven by low-level processes which occur before language during development. Evidence for colour CP in infants at four months (Bornstein, Kessen & Weiskopf, 1976; Franklin & Davies, 2004; Franklin, Pilling & Davies, 2005) is consistent with CP not being dependent on language. Further, 2- and 3-year-olds show colour CP on a 2-AFC task and this is not related to how they name the colours (Franklin, Clifford, Williamson, & Davies, 2005). Thus, perceptual and linguistic colour categories may be independent of each other, at least in infants and young children.

On the other hand, CP could also be learned and possibly sharpened by language learning. Özgen and Davies (2002) found that colour discrimination, measured by a same-different task, could be improved by practice, and improvement was specific to the training stimuli. They argued that equivalent perceptual learning during language learning could produce CP because language directs attention to the boundary regions between categories. For instance, if a child learned the term *blue*, this is often over generalised to include neighbouring colours such as *green*. Incorrect use of *blue* for a *green* would be corrected and the child has to discover what the difference between *blue* and *green* is in order to use the terms acceptably. Equivalent generalisation within the blue category would produce a correct response that would not be corrected. Thus, the child would attend more to boundary regions than to central regions, and this could produce increased sensitivity to boundary regions, and hence CP. They showed that an analogue of this conjectured learning process could be produced in adults through training. Adults were trained to learn two new categories that split green or blue in half, with the new boundary at the centre of blue or green. Before training, discrimination in the middle of blue or green was worse than towards the boundary. After learning the new categories, discrimination was now best around the new boundary (the centre of blue or green), reversing the original pattern. That is, discrimination across the newly learned boundary was better than in the middle of the new categories: CP had been produced by a few hours category training.

Evidence from cross-cultural studies also supports the learning account of CP. Kay and Kempton (1984) and Roberson, Davies and Davidoff (2000) found that CP only occurred if the category boundaries were marked linguistically. This is consistent with language affecting perceptual colour categories, but it is also consistent with a simple direct language account of CP. Further discussion is deferred until the direct language account has been considered more fully in the next section.

The direct language account of CP

CP could be due to colour naming (e.g., Rosen & Howell, 1987). For instance, in a same-different task, judging two stimuli to be different when they are physically different and nominally different (e.g., B1-G1) may be easier than judging that a within-category pair (B1-B2) are different because comparing names is sufficient to yield correct responses in the cross-category case, but is no use in the within-category case. Pilling et al. (2003) showed that a simple model of name comparisons predicted accuracy for cross-category pairs very well, but underestimated within-category accuracy for same-different and 2-AFC tasks. It is probable that within-category judgements are based on comparing visual codes, whereas cross-category judgements may use name comparisons too. The latter is consistent with Fujisaki and Kawashima's (1971) dual-code model. A stimulus can be represented in memory at multiple levels of abstraction. The physical representation of the stimuli is useful when the task stimuli belong in the same category, while the stored symbolic representation is used for cross-category pairs.

Roberson & Davidoff (2000) tested the direct language account of CP using a successive 2-AFC-discrimination task following the usual logic outlined earlier. When the interval between target and test-stimulus (ISI) was unfilled, and when it contained visual interference, performance was more accurate for the cross-category condition than for the within-category condition (e.g. B1-G1 more accurate than B1-B2). However, with verbal interference in the ISI, there was no difference between the two conditions. They concluded that CP was eliminated because verbal interference prevented remembering of the target name. Pilling et al. (2003) replicated this result and showed that equivalent results were also found for successive same-different tasks. Thus, there is strong evidence that CP may be a misnomer; it results from labelling rather than from perceptual processes.

Cross-linguistic studies

Languages vary in how basic many colour terms they have, and in the positions of category boundaries (Berlin & Kay, 1969). When a language has relatively few colour terms, these terms often cover all of colour space, and thus category sizes are larger than in languages such as English which has eleven basic colour terms (the maximum found according to Berlin & Kay, 1969). For instance, in Otjihimba (one of the languages studied here) red, pink, some orange and some purple are in one linguistic category (*otji-serandu* 'red') and blue and green are in another, *otji-mburou* 'blue-green'. Comparing languages with different categorical structures allows a test of the linguistic relativity hypothesis (Whorf, 1956). It also allows some insight into the origin of CP. If the distribution of CP does not vary across languages, this would suggest that CP is due to universal perceptual categories. On the other hand, if no CP is shown other than at a boundary marked by language, this could be interpreted in different ways. CP may be a perceptual effect that emerges from language learning; or it could be the result of direct labelling. Note that that both the universal and relativist account of CP can be valid. CP could be innate, as the infant studies suggest (e.g. Franklin & Davies, 2004), and later modified by language.

The first cross-language studies, although not directly testing CP, seemed to suggest that perceptual structures were hardwired and universal (Heider, 1972; Heider & Olivier, 1972; Rosch, 1973). The Dani have only two basic colour terms, but, for instance, remembered focal examples of red, green, yellow and blue, better than non-focal examples, just as the English speaking comparison group did. Rosch argued that the foci of Berlin & Kay's (1969) universal categories were perceptually salient universally. However, the findings from recent studies of possible effects of language differences on perception and memory have suggested that Rosch's strong universalist conclusions need to be qualified with a small relativist component (Davies, 1998; Davies & Corbett, 1997, 1998; Davies, Sowden, Jerrett & Corbett, 1998; Kay & Kempton, 1984; Pilling, 2001; Roberson, Davies & Davidoff, 2000). These studies used a range of tasks for instance, 2-AFC discrimination, colour grouping (free-sorting by similarity), triads ('which is the most different?') and recognition memory, but all found broad cross-language similarities consistent with universalism, modulated by small language-related differences, consistent with relativism. Equivalent results have also been found in cross-language studies of children from 4-7 years of age (Boyles, 2001).

Two of the above studies investigated CP specifically. First, Kay and Kempton (1984) compared speakers of Tarahumara (Mexico), whose language has a single term for blue and green, with English speakers, using a triads task. Three colours were presented and the one least like the other two had to be chosen. The colours were either all blue or all green (within-category) or two blues and one green, or two greens and one blue. The Tarahumara were unaffected by which English category the colours were in, but the English speakers tended to choose the nominal isolate on cross-category triads. According to Kay and Kempton, the Tarahumara choices were based purely on perceptual separations, whereas the linguistic differences affected the English groups choices, albeit implicitly.

Roberson, Davies & Davidoff (2000) compared speakers of Berinmo with English speakers using a successive 2-AFC-discrimination task. Berinmo has no blue-green boundary, but its boundary for *nol* ‘green’ and *wor* ‘yellow’ is in a different position to the English green-yellow boundary. Using the standard 2-AFC method for testing for CP, the English showed CP for the blue-green boundary, but not for the *nol-wor* boundary, while the Berinmo showed the opposite pattern. Both studies of CP are inconsistent with a simple hardwired perceptual account of CP, but they are consistent with both the learned perceptual account (stretching of cross-category distances make the nominal isolate the perceptual isolate as well) and the direct language account (the English use nominal differences when available to guide choice).

The present study

No simple answers to the questions about the origin and nature of CP emerge from the foregoing brief review. There is evidence that perceptual colour categories are hardwired (Bornstein et al., 1976; Franklin & Davies, 2004; Franklin, Pilling & Davies, 2005), that they can be learned (Özgen & Davies, 2002), that CP is due to labelling (Roberson & Davidoff, 2000) and that it is due to perceptual warping (Daoutis et al., 2005; Davies, Daoutis, Pilling & Wiggett, 2003; Kawai et al., 1995). Moreover, in most studies of CP, perceptual and direct language accounts cannot be disentangled. In the present study we chose to extend the use of visual search as a way of exploring CP. The task was to detect a colour target among a number of distractors that fell in the same or in a different linguistic category from the target.

The advantage of visual search is that when the target-distractor difference is on a single feature, in this case colour, search is very efficient, and therefore the use of labelling is not necessary. In this case, the prime determinant of search performance is target-distractor similarity (Duncan & Humphreys, 1989). Thus, if target-distractor similarities are equated in distances in colour space, and there is an advantage of a categorical difference between target and distractor, this would implicate perceptual processes, rather than labelling, as the origin of CP. We could therefore claim that a simple feature visual search task should be driven by the perceptual attributes of the stimuli, and should be unlikely to require labelling.

We present two cross-linguistic experiments which addressed the origin of CP. In both experiments, we compared native speakers of languages that differ in their categorical structures of colour, on a visual search task. For example, the language of the Kwanyama tribe in Namibia (Experiment 1) has no basic colour terms for the English orange, purple or pink, or for the English blue and green. These colours are included in large generic categories; for instance pink is included in red, and there is a compound term for blue-green. Similarly, the language of the Himba tribe in Namibia (Experiment 2) has a single linguistic category encompassing the English blue and green, and a single category for the English red and pink. The questions of interest were whether the speakers of different languages would perform differently on a perceptual task such as visual search, and whether their patterns of performance would mirror their respective linguistic colour categories. In particular, we looked for CP around English boundaries that are not present in either Namibian language. This finding would provide converging evidence that CP develops through perceptual learning, guided by language.

The subjects were young children from each language group. This was done for several reasons. First, we wanted to know whether children, like adults, would show CP in search tasks. Second, it is common in rural Africa for people to be familiar with several languages, including, in our case, English and Afrikaans. Young children are much less likely to have been exposed to additional languages, and to be pure monolinguals. This should ensure that the findings would not be confounded by the knowledge of foreign terms. We compared children from 4 to 7 years of age so that any changes in CP with age could be detected (Experiment 1). For instance, it could be that CP strengthens with linguistic competence. This might be particularly prominent for our rural African samples where colour term learning may not be complete until about 7 years of age (see e.g., Davies, Corbett, McGurk, & Jerrett, 1994).

The experiments reported below used paper-and-pencil search tasks, suitable for fieldwork and for school visits. The children had to tick the colours that matched the target colour in a search array, as fast and as accurately as they could. During search, the target was always visible at the top of the search array; the children did not have to remember what the target was. The speeded nature and reduced memory load of the task further discouraged any use of labelling (see Boyles, 2001). In addition, colour naming data were collected to examine the relationship between linguistic categories and colour search performance.

The colour stimuli were chosen to be perceptually equidistant within each condition. To achieve this, we chose colours equidistant in CIE space. The CIE space (1976; see Hunt, 1987) is a perceptually uniform space, intended to ensure that equal distances (ΔE) between colours correspond to equal discriminability. Although it should be noted that the CIELUV space is not perfectly uniform, it still is the best candidate for use in colour research. Moreover, since the questions of interest focused on differences between language groups, any deviations from uniformity should affect both language groups in the same way.

As mentioned earlier, both experiments reported here addressed the same main question: if two languages differ in the way they partition colour space, are these language differences mirrored in the performance on a perceptual task that is unlikely to require the direct use of language? In addition, each experiment explored different questions related to the main issues. Experiment 1, as already mentioned, looked at age differences in naming and search performance, to examine whether CP changes with age as a function of linguistic competence. Experiment 2 varied the number of distractors in the search task, to unravel possible differences in search efficiency between conditions in the two language groups. These additional questions will be presented in more detail with respect to each experiment below.

EXPERIMENT 1

Experiment 1 compared English speaking children with Kwanyama children. The Kwanyama live in Ovamboland in Northern Namibia near the border with Angola. The Kwanyama language has five basic colour terms: *oshitoka* ‘white’; *oshilaula* ‘black’; *oshitilyana* ‘red’ (and some darker pinks and oranges); *oshunga shei* ‘yellow’ (and some oranges) and *oshitwima* ‘blue or green’. *Oshitwima* is sometimes qualified by *eulu* ‘sky’ to denote blue and *omafo* ‘grass’ to denote green.

Possible category effects were investigated using a visual search task that varied the categorical relationship of target and distractors at constant target to distractor perceptual CIE distances. There was one type of target and one type of distractor. The task involved scanning an array of colours and marking targets as they were found. There were multiple targets to ensure that search was long enough to be timed accurately. Targets and distractors were either from the same English category (within-category) or from different categories in English (cross-category). Thus, for each region, one search array was cross-category for

English, but within-category for the Kwanyama: (e.g., B1 target among G1 distractors) and one search array was within-category for both groups (e.g., B1 target among B2 distractors). Categorical perception was evidenced if cross-category arrays were completed faster or with greater accuracy than within-category arrays. If colour terms warp perceptual colour space, then, CP should be evident for the English children, but not for the Kwanyama children. CP should be shown as faster or more accurate search when the search stimuli straddle category boundaries (e.g. B1-G1) than when they do not (e.g.B1-B2).

The colour terms described above are based on adult naming data and so a naming task was included here to test children's colour term knowledge. For the naming task, the experimental stimuli were presented individually and the child was asked to name the colour. This information allowed us to examine our central question, namely, the relationship between colour language and performance on a colour perception task. In addition, Experiment 1 examined whether, and how, this relationship changes with age**.** As learning of colour terms tends to occur later in rural African children, particularly those who have not been to school (Davies et al., 1994), it was likely that the younger children would not know all the colour terms. If this was the case, then it would be possible to test whether the size of the category effect increased as colour term knowledge increased.

Categorical effects were tested in three regions of colour space: blue-purple, blue-green and red-pink. Whether or not there were absolute differences in search times among these regions was not directly pertinent to the current issues. It should be noted, however, that while the target to distractor perceptual distances were equated within a colour set, they were smaller in the red-pink set than in the other two. Thus, as search should get harder as target-distractor distances reduce (Duncan & Humphreys, 1989), search should be slowest for the red-pink set. This therefore provides a test of whether perceptual distance affects children's search as in adults.

Absolute differences between the two language groups were not directly relevant to the current issues either. The two groups could differ in many irrelevant ways, such as facility in using a felt tip pen. Rather, tests of the main questions would be provided by significant interactions between language and category, category and age, and possibly by language category and age. The data analysis will therefore pay particular attention to these.

Method

Participants

Sixty-four English children and 45 Kwanyama children took part in the study. The age range of the participants was 4 years to 7 years. In the Kwanyama sample there were 12 4-year-olds, 9 5-year-olds, 12 6-year-olds, and 12 7-year-olds. For the English sample there were 16 children in each age group. There were equal numbers of boys and girls in the English sample, and 40% of the Kwanyama sample were male. The English participants were sampled from schools in Surrey, England and the Namibian participants were sampled from Kwanyama speaking villages in rural Namibia.

Stimuli and Design

Visual search. Each search array consisted of 48 circular colour patches, 2cm in diameter, arranged in an equally spaced grid of 6 columns and 8 rows on an A4 sheet of card. At the top centre of the page was a single target colour and there were 16 target circles, and 32 distractor circles. Stimuli were produced by a Hewlett Packard colour proofer, and checked by a Minolta-CS-100 colorimeter. Table 1 gives the expected colour terms for each language group. Within a colour set, the perceptual distances in CIE (ΔE) between adjacent pairs were equal. The perceptual distances of the stimulus pairs in CIE units (ΔE) were 30 ΔE for the blue-purple and blue-green sets and 20 ΔE for the red-pink set. Therefore, the separation sizes for the red-pink set were smaller than for the other sets.

For the English participants, the tasks were completed under lighting conditions that approximated natural daylight (colour temperature 6500ºK, at 810-1880 lux). For the Namibian participants the tasks were completed outside, not in direct sunlight or deep shade (colour temperature 5500 – 7000ºK as indicated by a Gossen colormaster 3F). For the Namibian data, the experimenter was a student nurse at the University of Namibia, Windhoek. Her native language was Kwanyama, and the data were collected in and near to her home village. Training on the task procedure was provided and translations of the task instructions were discussed to ensure that they were satisfactory.

Each participant completed six search arrays, two for each colour set. One array was within-category (A, B) and one was cross-category (B, C). Category membership was defined by English adult naming. There were two versions of each stimulus combination; one version had, for instance, A as the target and B as the distractor, and the other version had B as the target and A as the distractor. Half of the participants in each language and age group were given one version and the other half were given the other.

INSERT TABLE 1 HERE

Naming task. Two cm^2 coloured squares of each stimulus were mounted on $4cm^2$ squares of white coloured card. These were used for individual stimulus naming.

Procedure

Visual search task. Each language group was tested by a native speaker of the appropriate language. The English version of the instructions was: “In this game you have to find all the dots that are the same colour as this dot at the top. So when I say ‘go’ can you mark all the dots that are the same colour as this dot at the top. When you have finished put your pen in the empty circle on the bottom of the page. Do it as fast as you can." A practice array (green target, yellow distractors) was given first. Once it was clear that the task was understood, the six experimental arrays were completed in a random order. The time taken from 'go' to pen in the finishing circle was measured with a stop-watch.

Naming task. After completing the visual search task, each stimulus was presented individually, in random order, and had to be named.

Results

Naming

Table 2 shows, for each language group, the terms given to each stimulus, and the percentage of the sample that used that term. The expected term is shown in bold. For the English sample the majority term was 96–100% consistent with the expected term, and thus, as intended, there was a linguistic boundary between stimuli B and C for each colour set. For the Kwanyama there was lower agreement, with only 39–61% agreeing on the majority colour term. The pattern of naming for the Kwanyama was similar for each age group and the frequency of 'I don't know' responses did not go down with age, F (3, 45) = 1.77, p = 0.17. References to specific items were offered as names for the stimuli. For example, the words *ombidi* ('like traditional spinach') and omufyaati ('leaves of the mopane trees') were offered for the stimuli from the blue-green and blue-purple sets, and the word *olaole* ('earthworm') was offered for the stimuli from the red-pink set. The majority term for all the stimuli of the blue-green-and blue-purple sets, except C in blue-purple, was *oshitwima*. The majority term for A and B of the pink-red set was *oshitilyana*. The low level of naming agreement meant that there was no clear name-name boundary for the blue-purple and pink-red sets. However, if 'I don't know' is treated as indicating non-membership in a category, then there is higher consensus. Based on the majority response, there is a boundary between

B and C for the blue-purple set (*oshitwima, oshitwima,* 'no name') and for the red-pink-set (*oshitilyana, oshitilyana,* 'no name') but there is no boundary for the blue-green set (all *oshitwima*).

INSERT TABLE 2 HERE

Visual Search

As less than 1% errors were made by either language group, only search time was analysed. Table 3 gives the mean times in seconds taken to complete each array.

INSERT TABLE 3 HERE

A four factor mixed effects ANOVA on category (cross-/within), language (English/Kwanyama), set (blue-purple/blue-green/pink-red) and age (4/5/6/7 year-olds) was conducted on search times. Significant interactions involving language were explored with separate ANOVAs for each language group. Other post-hoc tests were Bonferroni corrected *t* tests.

Main effects

All main effects were significant. Within-category search was slower than cross-category search (means: 28.36 (9.5), 25.02 (8.8) secs respectively), F (1,101) = 57.62, $p < .001$. The English group were slower than the Kwanyama group (means: 29.2 secs (8.72) and 24.1(8.32) secs respectively), F (1,101) = 11.48, $p < .001$. Search times reduced with age (means: 31.1 (10.4), 32.1 (6.9), 26.0 (7.5), 20.0 (3.8) secs respectively), F (3,101) = 16.65, $p < .001$. There was no difference between 4- and 5-year-olds, but 5- and 6-year-olds and 6-

and 7-year-olds both differed, (minimum t (54) = 2.98, $p < .05$). The colour set effect (F (2,202) = 20.11, $p < .001$) was due to the red-pink set (29.0 secs. (10.3)) being slower than both the blue-green (25.4 secs. (9.4)) and the blue-purple set (25.6 secs. (9.4)), minimum t (108) = 4.87, $p < .001$.

Interactions

There were also significant interactions between language and category (F (1,101) = 22.85, $p < .001$), language and age (F (3,101) = 3.89, $p < .05$), language and set (F (2,202) = 9.40, $p < .001$), category and set (F (2,202) = 8.32, $p < .001$), and category, set and language (F (6,202) = 5.01, $p < .01$). No further significant interactions were found (largest F = 1.24, smallest p = 0.29).

Category by language interaction. Figure 2 shows the mean within- and cross-category search times for the Kwanyama and the English children.

INSERT FIGURE 2 HERE

It seems that the category effect was larger for English than for Kwanyama. This impression was supported by ANOVAs on each language that showed a significant category effect for the English children, F (1, 60) = 87.9, $p < .001$, but not for the Kwanyama, F (1, 41) = 3.69, p = 0.06.

Language by age interaction. Figure 3 shows the mean search times for each age group, for each language.

INSERT FIGURE 3 HERE

It appears that the interaction is due to the difference between the two language groups varying unsystematically with age. Tests of simple main effects using ANOVAs for each language group, show significant effects of age for both languages (minimum $F = 7.06$, largest $p < 0.01$). However, consistent with the visual impression from Figure 3, the reason for the language by age interaction was that the two languages did not differ in speed at the ages of 5 and 7 years (maximum $t = 0.84$, $p = 0.41$, but did at 4 years (t (26) = 3.0, $p < 0.01$) and at 6 years (t (26) = 2.98, $p < 0.01$).

Language and set. Figure 4 shows the mean search times for each set, for each language group.

INSERT FIGURE 4 HERE

It appears that the difference between the two language groups is particularly marked for the red-pink set. This was supported by separate ANOVAs for each group that showed a significant set effect for English, F (2, 60) = 29.5, $p < .001$, but not for the Kwanyama, F (2, 41) = 2.07, $p = 0.13$. For English, the red-pink set was slower than both the other sets (minimum t (63) =5.70, $p <.001$) and there was no other difference[2].

Category by language by set interaction. Figure 5 shows the mean within- and cross-category search times for the English children (a) and the Kwanyama children (b) for each set.

INSERT FIGURE 5 (5a AND 5b) HERE

Comparison of the two sets of data suggests that the category effect is particularly marked for the red-pink set for the English children. This impression was supported by the separate language ANOVAs which showed a significant interaction between category and set for the

English children, $F(2,120) = 12.38$, $p < .001$, but not for the Kwanyama children, ($F = 1.80$, $p = 0.17$). Direct comparisons of the difference between within-category and cross-category conditions between the three sets showed that for the English children the red-pink category effect was significantly larger than the blue-purple ($t(63) = 4.11$, $p < .001$) and the blue-green set ($t(63) = 4.29$, $p < .001$). This may be because of differences in discriminability across sets. It is possible, for example, that the size of the category effect was masked by ceiling performance (floor search times) in the blue-green and blue-purple sets, while the greater perceptual difficulty in the red-pink set made the category effect more pronounced.

Effects of naming on search

If the category effect in the English children was due to naming, then the Kwanyama who showed the same pattern of naming as the English should show a category effect too. The size of the category effect was investigated for those Kwanyama children who gave the within-category stimuli the same name and the cross-category stimuli different names when 'no name' responses were treated as a distinct label. There were 15 children showing this pattern of naming for the blue-purple boundary, 10 for the blue-green boundary and 17 for the pink-red boundary. Table 4 gives the mean search times for these children for within- and cross-category arrays.

INSERT TABLE 4 HERE

Paired comparison *t* tests revealed no significant category effect for these children on any set: blue-purple, $t(14) = 1.84$, $p = 0.09$; blue-green, $t(13) = 0.95$, $p = 0.36$; pink-red, $t(15) = 0.96$, $p = 0.35$.

Discussion of Experiment 1

The findings of interest in Experiment 1 are outlined here. First, both language groups could perform the task very accurately, and search speed increased with age. Second, the English children were sensitive to the perceptual discriminability between sets. They were slower when the perceptual distance between the colours was smaller (pink-red set: 20 ΔE), than when the perceptual distance was larger (blue-green and blue-purple sets: 30 ΔE). This difference in performance between colour sets was absent in the Kwanyama. Third, and perhaps most importantly, cross-category search was faster than within category search in the English group, but this advantage was absent in the Kwanyama. The category advantage in the English group was consistent with the presence of a linguistic boundary in the blue-green, blue-purple, and red-pink sets. The absence of the effect in the Kwanyama also was consistent with the absence of a name-name boundary. The Kwanyama children who had a name-‘I don’t know’ boundary did not show the category effect.

As highlighted earlier, the search task was chosen on the assumption that it is driven primarily by the discriminability between the stimuli and does not require labelling. If this is indeed the case, then our results are consistent with the perceptual warping account of CP. Learning names for the categories could lead to cross-category stretching of colour space and thus increased discriminability between stimuli that cross the category boundary; or, most in accordance with the evidence from infant studies, the ability to perceive categorically, already present at birth, could be enhanced or attenuated by the presence or absence of a linguistic boundary. Thus, having the linguistic boundary should lead to faster search in the cross-category conditions (Duncan & Humphreys, 1989); the absence of the linguistic boundary, on the other hand, should be reflected in no category advantage.

This conclusion, however, may be weakened by those Kwanyama children with a name-‘I don’t know’ boundary who did not show CP. If we assume that a name-‘I don’t know’ boundary in the Kwanyama is equivalent to a name-name boundary in English, then it is unclear why the Kwanyama did not show a category effect, as the English group did. Alternatively, it is possible that a name-'I don't know' boundary is not adequate to produce perceptual change. The stimulus that cannot be named may still share many features with the named stimulus. This, in turn, may reduce the likelihood of perceptual stretching around the two stimuli and hence the possibility of a category effect. Thus, we can only claim relationships between language and perception in the Kwanyama if we treat the name-'I don't know' boundary as a case of no boundary. For this reason, claims for a correspondence between linguistic categories and perceptual performance should be treated with caution in the Kwanyama. To draw more confident conclusions, it would be necessary to include a condition with a name-name boundary in our African group. If we were able to show CP in this case, then we could make stronger claims about perceptual warping.

Although, as already noted, the visual search task should not require or encourage labelling, we cannot completely rule out the possibility that the category effect shown in the English group were due, not to perceptual warping, but to the use of a direct naming strategy. If the children name the target and then search for stimuli with the same name, this could facilitate cross-category search, but not within-category search. The English search times are equivalent to about 600ms per stimulus, which is probably too fast for all the stimuli to be named. If the English were using a naming strategy, whereas the Kwanyama relied more on the perceptual attributes of the stimuli, this might also explain why the English were slower than the Kwanyama.

In short, Experiment 1 provided evidence for a category effect in the English, but not the Kwanyama group. In the English group, the presence and size of the category effect was consistent with naming patterns, as well as with discriminability between the stimuli. In the Kwanyama group, the absence of a category effect was consistent with the lack of a name-

name boundary. These results taken together support a relationship between language and performance on the colour search task; and more specifically, suggest equivalence between naming patterns and evidence for perceptual warping. However, this conclusion would be stronger if it were supported by a category effect around a name-name boundary in the Kwanyama. One of our future aims should be to identify a linguistic boundary in this group, while keeping the perceptual distances within the desired range for the search task.

As will be seen, Experiment 2 approached the same question of the relationship between colour naming and search performance in another African tribe, the Himba. Although the central question was the same, Experiment 2 addressed further questions related to the nature of CP; it also attempted to identify a name-name boundary for the African group.

EXPERIMENT 2

As in Experiment 1, our main question in Experiment 2 related to potential differences in colour search performance between two language groups that partition the colour space differently. To address this, we compared an English group and a Himba group from the Himba tribe in Northern Namibia, on a colour search task. The Himba language, Otjihimba, has five basic colour terms: *otjivapa* (white), *otjizoozu* (black), *otjiserandu*, (red, orange and pink), *otjidumbu* (yellow), and *otjimburou* (blue, green, and some purples). It also has a less frequent number of secondary terms and a limited number of borrowed terms, namely *otjingirine* (green) and *otjipinke* (pink). As there is a single term for blue and green, *otjimburou*, it was predicted that, if colour naming affects colour perception, search for a blue (B) target among green (G) distractors, and vice versa, would be faster and more accurate than search of a B1 target among B2 distractors or a G1 target among G2 distractors. This effect, however, would be present only in the English group that has distinct linguistic categories for blue and green. For the Himba group, all these search conditions should be within-category, and thus no category effect should be expected. This main prediction was therefore the same as our hypotheses in Experiment 1. If the findings of Experiment 1 are robust, then, despite small differences in the samples and methods, the results should be replicated in Experiment 2.

Experiment 2 addressed some further issues. As mentioned in the Discussion of Experiment 1, to make a strong claim that colour language is related to colour perception, we should be able to show a category effect with a name-name boundary and the absence of the effect when there is no boundary, within the same language group. One of our aims was therefore to use colours that crossed a linguistic boundary in Otjihimba. To achieve this, we included a green-yellow set for which the Himba were likely to have separate terms, possibly *otjimburou* and *otjindumbu.* It should be noted, however, that to achieve the desired perceptual distance between targets and distractors (see below) we had to choose light greens that were not likely to be best category examples, and for which we had no previous naming data. For this reason we were not sure whether this set would really include a linguistic boundary for the Himba; its choice was only tentative.

The search task used here was slightly more demanding than the one used in Experiment 1. There were two types of distractor instead of one. The distractors were always from the same category, as defined by English adult naming. For example, a blue-green search meant that the target was blue (B) and the distractors were either green1 (G1) or green2 (G2). In addition, target-distractor distance was about 20 ΔE (see Nagy & Sanchez, 1990), a distance comparable to the red-pink set in Experiment 1. The two types of distractor and the relatively small perceptual distances meant that the children were likely to make errors, which would provide an additional measure of search performance for our analysis. Making the task more difficult also helped avoiding ceiling performance, allowing category effects to emerge.

A final question related to search efficiency, as indicated by the effects of set size on the search times. The number of targets was fixed, but the number of distractors was varied. Thus, the target to distractor ratio was manipulated. We had two different conditions with a 1:4 and 1:7 target:distractor ratio, respectively. This provided, not only an additional measure of search efficiency, but also possibly some insight into how the children performed the task. If the search time was not affected by the number of distractors, the search slopes would be flat. Flat slopes, in turn, should mean very efficient search (see Wolfe, 1998), and would possibly suggest that the task was stimulus-driven (Wolfe, 2003). If, conversely, the number of distractors did affect performance, this should indicate top-down control of the search (e.g. visual comparisons).

We therefore addressed the following questions. First, would the number of distractors in the search array affect the two language groups differently, implying that the two groups of children performed the task in different ways? Second, would the effect of set size vary depending on the search condition? More specifically, we wanted to see whether cross-category search would be more efficient, with flatter slopes, than within category search.

In short, Experiment 2 addressed the same main question as Experiment 1, namely the role of linguistic boundaries in the categorical perception of colour. Of interest were the potential cross-language differences in category effects. Search for a target that crosses a name-name boundary with the distractors should be faster, more accurate and overall more efficient than search where the target and distractors had the same name. If we were successful in finding stimuli that crossed a linguistic boundary in the Himba, this group should be able to show this category advantage around the name-name boundary. This would further strengthen the suggested relationship between language and colour perception.

Method

Participants

Thirty-two 6- and 7-year-old English children, 12 boys and 20 girls, were recruited from local primary schools in Surrey. Thirty-two Himba children, 14 boys and 18 girls, with estimated ages of 6 and 7, were recruited from seven different Himba villages in the Kaokoveld region of Northern Namibia. The children received little or no schooling: 40 % of the children had never been to school; around 10% had only just started school; and the remaining 50% had been to school for 18 months on average (range: 12 to 30 months).

Stimuli and Design

The search arrays took the same form as in Experiment 1, except two types of distractor were used and the particular stimuli differed. There were 12 stimuli for the experimental arrays, 6 for each boundary condition: *blue1-blue2-blue3-green1-green2-green3* (blue-green set) and *lightgreen3-lightgreen2-lightgreen1-yellow3-yellow2-yellow1* (light green-yellow set), as defined by adult English naming. Table 5 shows the perceptual distances, ΔE, between the stimuli. As can be seen, within each set the perceptual distance between adjacent colours was approximately 20 ΔE. The target was always adjacent to the category boundary and the distractors were either the two remaining colours in the category for within-category conditions (e.g., *green1* amongst *green2* and *green3*) or in the cross-category conditions, the two colours nearest to the target across the boundary served as distractors (e.g., *green1* amongst *blue1* and *blue2*); see Figure 1. There was also a practice condition consisting of yellow targets among green and red distractors. In the practice condition, the perceptual distances between the colours were over 100 ΔE units. The yellow and green stimuli in this condition were very different, and much more discriminable, than the stimuli used in the *lightgreen-yellow* conditions.

INSERT TABLE 5 HERE

There were two set sizes: 30 and 48 stimuli. In both set sizes, there were 6 instances of the target. For size 48, there were 42 distractors (21 instances for each type of distractor); for size 30, there were 24 distractors (12 instances for each type of distractors). Note that the number of targets remained fixed: what changed was the target/distractor ratio. For set size 30, stimulus locations were randomly selected in the grid, leaving the remaining space blank. For set size 48, stimuli were present in all locations in the grid. The location of targets and distractors in the matrix was random.

There were 8 search arrays made up from each combination of 2 colour-boundaries (blue-green or yellow-green), 2 category conditions (within/cross-) and 2 display sizes (30 or 48).

Two-cm^2 coloured squares of each stimulus were used for individual stimulus naming.

Procedure

Each participant did the practice condition followed by 8 experimental arrays, and then the naming task. Half of the children searched for a blue target among (a) blue distractors (within category; B1 among B2 and B3; see Figure 1) and (b) green distractors (cross-category; B1 among G1 and G2; see Figure 1). This group of children also searched for a light green target among (a) light green distractors (within category; LG1 among LG2 and LG3) and (b) yellow distractors (cross-category; LG1 among Y1 and Y2). The other half of the children searched for a green target among (a) green distractors (within category; G1 among G2 and G3) and (b) blue distractors (cross-category (G1 among B1 and B2), and for a yellow target among (a) yellow distractors (within category; Y1 among Y2 and Y3) and (b) light green distractors (cross-category; Y1 among LG1 and LG2). Thus, within a set, the target was always the same for cross- and within-category search.

The order of the display size condition was counterbalanced, and within each display size, the order of the 4 arrays was randomised.

The testing procedure was the same as in Experiment 1.

Results

Naming

Table 6 shows the name frequencies for each stimulus in each language group. As can be seen, the English group named the stimuli with high agreement. However, the pattern was less clear for the Himba. The majority term for the blu*e* stimuli was *otjimburou* (56 to 72%); however, naming of the green colours varied: the dominant term was *otjimburou* (48%) for the boundary green colour, but the dominant term for the other two greens was *otjingirine* (44% in both cases; this is a loan term from English *green*); this term competed with *otjimburou* (32-40%). Naming of the green-yellow set was not according to prediction in the Himba: all the stimuli in that group had the dominant term *otjindumbu*, except where *otjindumbu* (32%) competed with *otjingirine* (36%). Use of other terms and *I don't know* responses were relatively low. It appears that there was no clear name boundary for the blue-green region (as intended), but nor was there one for the green-yellow set. This was supported by inspection of individual naming patterns. Only two children had a linguistic boundary for blue and green; only one child had the boundary for the green and yellow stimuli. In short, for all stimuli used, English children had name-name boundaries for both *Blue-Green* and *Lightgreen-Yellow*, while Himba children did not. Naming did not appear to vary systematically with the amount of schooling.

INSERT TABLE 6 HERE

Search

Accuracy was not at ceiling, therefore search time and accuracy were analysed. Accuracy was expressed as A', a non-parametric equivalent of d' (Donaldson, 1993; Pollack & Norman, 1964), which combines hits and false positives into a bias free index of accuracy. Seven Himba children had search times more than 1.96 standard deviations above the mean. In the analysis reported below, these children were excluded from the sample as outliers. However, it should be noted that the directions of the means were the same both when they were included and when they were removed from the analysis. This left 25 Himba children and 32 English children. The means and standard errors in the search times of the Himba and English samples are shown in Figures 6a and 6b. Figures 7a and 7b show the mean accuracy (+/- 1 S. E.) for each condition for the Himba and English samples.

INSERT FIGURE 6 (6a AND 6b) AND FIGURE 7 (7a AND 7b) HERE

The two groups did not meet the homogeneity of variance assumption (Levene’s test $p < 0.05$ for most conditions), even after the data were transformed; therefore separate analysis was required for each language group. Three-factor (display size (30, 48) by category (within, cross-) by colour region (blue-green, green-yellow) repeated measures ANOVAS on time and accuracy were used separately for each language group.

Main effects.

Category. In the Himba, the main effect of category was not significant for search time ($p > 0.9$, nor for accuracy ($p > 0.3$); also see Figures 6a and 7a. Within-category and cross-category performance were not significantly different (means: within-category time = 38.45, SD = 13.1, cross-category time = 38.29, SD = 12.4; within-category accuracy = 0.89, SD = 0.01, cross-category accuracy = 0.88, SD = 0.01).

In the English children, however, the effect of category was significant, F (1, 31) = 74.65, $p < 0.001$ for time, and F (1, 31) = 15.94, $p < 0.001$ for accuracy. Thus, the English children were both faster and more accurate in the cross-category conditions than in within-category conditions (means: within-category time = 33.05, SD = 6.96; cross-category time = 25.09, SD = 5.55; within-category accuracy = 0.91, SD = 0.06; cross-category accuracy = 0.96, SD = 0.06; see Figures 6b and 7b).

Colour set. The effect of colour set was significant in both language groups. Main effects of colour set on time: F (1, 24) = 19.46, $p < 0.001$ for the Himba; F (1, 31) = 24.27, $p < 0.001$ for the English. Main effects of colour set on accuracy: F (1, 24) = 9.18, $p < 0.01$ for the Himba; F (1, 31) = 7.08, $p < 0.05$ for the English. Inspection of the means (also see graphs 6a and 6b, 7a and 7b) suggests that both language groups were slower and less accurate in the blue-green set than in the green-yellow set (Himba means: blue-green time = 43.52, SD = 15.05; green-yellow time = 33.22, SD = 10.8; blue-green accuracy = 0.84, SD = 0.15; green-yellow accuracy = 0.93, SD = 0.01. English means: blue-green time = 34.80, SD = 10.58; green-yellow time = 23.33, SD = 6.34; blue-green accuracy = 0.91, SD = 0.06; green-yellow accuracy = 0.96, SD = 0.06).

Set size. There was no significant main effect of set size for the Himba. In the English group the effect was significant, F (1, 31) = 30.45, $p < 0.001$ for time; F (1, 31) = 5.13, $p < 0.05$ for accuracy. From Figures 7a and 7b it can be seen that, in the Himba, search time seems unaffected by the number of distractors, yielding flat search slopes (means: small set size time = 37.11, SD = 11.95; large set size time = 39.62, SD = 13.9; small set size accuracy = 0.89, SD = 0.1; large set size accuracy = 0.89, SD = 0.1).

English children, conversely, were slower when there were more distractors in the display (means: small set size time = 25.57, SD = 6.57; large set size time = 32.56, SD = 6.91; small set size accuracy = 0.92, SD = 0.06; large set size accuracy = 0.95, SD = 0.06).

Interactions

No interactions were significant for time or accuracy for the Himba. In the English group there were two significant interactions related to search time: category and colour set, F (1, 31) = 20.15, $p < 0.001$; colour set and set size, F (1, 31) = 9.42, $p < 0.01$.

English search time: category by colour interaction. The category effect was larger for the yellow-green set than for the blue green set (means of within-cross-category difference, averaged for both set sizes: blue-green = 3.37, green-yellow = 12.55; also see Figure 6b). This was confirmed by comparing the sizes of the category effect for each colour region in the English group, t (31) = 4.49, $p < 0.001$.

Search time: colour by set size interaction. Inspection of Figure 6b suggests that search slopes were relatively flat in the green-yellow condition, but steeper in the blue-green condition. To test this, slopes for blue-green and green-yellow were calculated (mean slopes, averaged for cross- and within-category conditions: blue-green = 5.19, green-yellow = 1.80). It was confirmed that slopes for green-yellow were significantly smaller than slopes for blue-green, t (31) = 3.07, $p < 0.01$.

As can be seen, overall the analyses of accuracy and search times were very similar; accuracy results supported the patterns found for search times. The findings cannot be attributed to speed-accuracy trade-off. When the analysis was repeated on search time, with accuracy as a covariate, the same patterns of results were found.

It should be noted that the pattern of results in the Himba did not vary systematically with the amount of schooling.

Discussion of Experiment 2

A common finding in both language groups was that search was faster and more accurate in the green-yellow than the blue-green set. However, there was a number of results specific to the language groups; these are outlined here.

First, the English group showed a significant category advantage, but the Himba did not. Cross-category search was faster and more accurate than within category search, but this advantage was only present in the English group. The category advantage in the English group was significantly larger in the green- yellow set than in the blue-green set.
Second, the English children were slower when the array size was larger, while the performance of the Himba did not appear to depend on set size. In the English group, search slopes were considerably flatter (more efficient search) in the green-yellow set than the blue-green set. Third, in the Himba, these patterns of results did not depend on naming, or time spent in school. In the English group, the presence of the category advantage was consistent with the presence of a linguistic boundary.

One of the aims of Experiment 2 was to find a name-name boundary in the Himba. In this way, a within- and a cross-category condition could be compared directly within the same language group. We used blue-green stimuli, for which the English had a linguistic boundary but the Himba did not, and green-yellow stimuli, which were intended to have a name-name boundary in both the English and the Himba. However, this was not the case for the Himba. All the stimuli in the LightGreen-Yellow set, with the exception of *LightGreen3*, were given the single name *otjindumbu*. As already noted, for these stimuli to be within a required range of perceptual distance, the greens had to be light. Thus the mburou-ndumbu boundary in the Himba was missed.

Besides missing a name-name boundary in the Himba, which would strengthen our conclusions, the findings of Experiment 2 are, at first glance, straightforward. In the English group, there was a category effect for both blue and green and green and yellow. The category effect was absent in the Himba. This agrees with the naming patterns: the language group that had a linguistic boundary for these colours performed the search categorically; the group that grouped the colours under a single name did not.

In both language groups, performance was better in the green-yellow set than in the blue-green set, even though the perceptual distances in CIE were equated. This suggests a difference in discriminability between the two sets. Since we were interested in the cross-language comparisons, and this difference in discriminability affected both groups in the same way, this finding does not alter our conclusions. However, it is less clear why the category effect in the English group was more pronounced in the green-yellow set. If naming was the only factor determining CP, then we should not expect differences in the size of the category effect between colour sets. Instead, it seems that some characteristic of the cross-category green-yellow stimuli enhanced detection of the target in the English group, and thus enhanced the category effect.

On the basis of the search slopes, there is evidence for top-down search in the English group. The significant effects of set size in this group suggest that the search was performed in a way that was, at least partly, dependent on the number of distractors. The children may have matched the stimuli of the array with the target template, a process that may or may not involve direct labelling. It is possible, though not directly testable here, that the green-yellow cross-category stimuli benefited more from this top-down guidance: a green target among yellow distractors , for example, was more salient on the basis of two features: ‘green’ and ‘dark’, as opposed to one feature: ‘blue’ in the blue-greens set.

Overall, Experiment 2 reported the presence of a category effect in a group with name-name boundary (English group) and the absence of the category effect in a group with no linguistic boundary (Himba). This finding supports and replicates the results of Experiment 1 with a different language group. This pattern of results could be interpreted as perceptual warping around a linguistic boundary. However, the effects of set size in the English group and the absence of these effects in the Himba group also suggested that the two language groups performed the task in different ways. It could be argued, therefore, that the category effects shown in the English group are a result of top-down guidance, where labelling of the search stimuli may play a role.

GENERAL DISCUSSION

The experiments reported here investigated the category effect in visual search in children. They further examined the nature of the category effect, namely the involvement of native language in performance on a perceptual task. If language shapes our perceptual category boundaries, then whether we perceive stimuli categorically or not depends on the linguistic boundaries we possess for these stimuli. This would be supported by cross-language differences in category effects, accompanied by analogous differences in linguistic categorisation.

The main finding of both experiments reported here was the presence of a category effect in visual search. When the target and distractors belonged in different categories, search performance was considerably better. There is little evidence in the literature for category effects in colour search in adults (Daoutis et al., 2005; Franklin et al., 2004; Kawai et al., 1995); here we demonstrated them in children.

Less clear was the locus of the category effect, and the way in which language is related to category search. Experiment 1 demonstrated a category effect in English children, who have the equivalent linguistic boundaries, but not in Kwanyama children, who do not. However, a group of Kwanyama children with a name-‘I don’t know’ boundary did not show a category effect. Thus, if the category effect is related to language, name-name boundaries may be necessary. Although one of the aims of Experiment 2 was to include a name-name boundary for both language groups, it was not possible to find a boundary for the Himba. Possible correspondences between naming patterns and category effects in search remained, therefore, partly inconclusive. The absence of linguistic boundaries (blue-green and green-yellow) in the Himba group was accompanied by absence of a category effect, whereas the presence of these linguistic boundaries in the English group was accompanied by the presence of a category effect. This pattern was also apparent in Kwanyama children in Experiment 1. However, relationships between naming and perceptual performance would be further strengthened if we showed category effects with a name-name boundary in the Himba group.

Effects of set size provided some insight into the nature of the category effect. Language may sharpen our perceptual boundaries, but it can also be involved in top-down guided search. As already mentioned, we chose visual search to reduce the possibility of using linguistic strategies. The significant effects of set size in the English group, however, suggested that the children performed a guided search that could involve the use of language. This involvement of language could be in the form of explicit labelling: the stimuli should be beneficial in cross-category search only, where the target is nominally distinct from the distractors. Alternatively, and perhaps more in line with the search times reported here, the children may have performed the task by comparing the search items to the target. As the target was always present at the top of the array during search, it could serve a simultaneous cue (see Laarni, 2001).

In short, here we demonstrated that children can be aided by the category status of stimuli when performing visual search: when the target and distractors belong in different colour categories, search is more efficient. Moreover, this categorical advantage disappears if the target and distractors are not separated by a linguistic boundary. Naming patterns seem to agree with performance but, since there is evidence for top-down guidance, a labelling strategy cannot be dismissed.

The present paper provided some important findings on children's visual search, and some insight into the nature of the category effect. Even in a simple search task we found evidence suggesting a relationship between colour language and perception. We also presented some evidence of how the children might perform the task. Further work should aim at replicating these findings by looking at name-name boundaries in both languages being compared. It is also important to examine more closely the contribution of top-down guidance in children's visual search. Strong claims of perceptual warping induced by language would only be justified by showing a clear correspondence between naming patterns and performance on a task where top-down involvement is minimal. The present study was a first step towards addressing this issue.

Author note

Christine A. Daoutis, Anna Franklin, Amy Riddett, Ally Clifford, Ian, R.L. Davies, Department of Psychology, University of Surrey, Surrey, England.

Experiment 1 appeared in Anna Franklin's doctoral thesis (Franklin, 2003) and was also presented at the European Conference on Visual Perception in 2003 (Franklin, Riddett, Clifford & Davies, 2003). Experiment 1 was supported by Economic and Social Research Council Postgraduate Studentship R42200124191 to Anna Franklin and Experiment 2 was supported by Economic and Social Research Council Grant R000223894 to Jules Davidoff, Debi Roberson, Ian Davies and Michelle de Haan. We are grateful to Nigel Woodger for the production of the stimulus arrays and to Marieta Bester and the student nurses at the University of Windhoek for help with data collection in Namibia.

Amy Riddett and Ally Clifford are now at the University of Bath.

Correspondence concerning this article should be addressed to Christine Daoutis, Department of Psychology, University of Surrey, Guildford, Surrey, GU2 7XH, UK. Email: c.daoutis@surrey.ac.uk.

References

Berlin, B., & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.

Bornstein, M. H., & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research, 47,* 1-17.

Bornstein, M. H., Kessen, W., & Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal-of-Experimental-Psychology: Human Perception and Performance, 2,* 115-129.

Boyles, S. K. (2001). Children's colour naming and a test of the linguistic relativity hypothesis. *Unpublished PhD thesis*. University of Surrey.

Daoutis, C. A., Pilling, M., & Davies, I. R. L. (2005). Categorical effects in visual search for colour. *Manuscript submitted for publication.*

Davies, I. R. L. (1998). A study of colour grouping in three languages: A test of the linguistic relativity hypothesis. *British Journal of Psychology, 89*, 433-452.

Davies, I. R. L., & Corbett, G. C. (1997). A cross-cultural study of colour-grouping: evidence for weak linguistic relativity. *British Journal of Psychology, 88,* 493-517.

Davies, I. R. L., & Corbett, G. C. (1998). A cross-cultural study of color-grouping: tests of the perceptual-physiology account of color universals. *Ethos*, *26*, 338-360.

Davies, I. R. L., Corbett, G., McGurk, H., & Jerrett, D. (1994). A developmental study of the acquisition of colour terms in Setswana. *Journal of Child Language*, *21*, 693-712.

Davies, I. R. L., Daoutis, C. A., Pilling, M., & Wiggett, A. (2003, September). *Categorical perception in visual search.* Paper presented at the XIIIth European Society for Cognitive Psychology conference, Granada, Spain.

Davies, I.R.L. & Franklin, A. (2002). Categorical similarity may affect colour pop-out in infants after all. *British Journal of Developmental Psychology, 20,* 185-203.

Davies, I. R. L., Sowden, P. Jerrett, D. T., & Corbett, C. G. (1998). A cross-cultural study of English and Setswana speakers on a colour triads task: A test of the Sapir-Whorf hypothesis. *British Journal of Psychology, 89,* 1-15.

Donaldson, W. (1993). Accuracy of d' and A' as estimates of sensitivity. *Bulletin of the Psychonomic Society*, *31*, 271-274.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433-458.

Franklin, A. (2003). The origin and nature of Categorical Perception of colour. *Unpublished PhD thesis*, University of Surrey.

Franklin, A., & Davies, I. R. L. (2004). New evidence for infant colour categories. *British Journal of Developmental Psychology, 22,* 349-377.

Franklin, A., Clifford, A., Williamson, E., & Davies, I. R. L. (2005). Color term knowledge does not affect Categorical Perception of color in toddlers. *Journal of Experimental Child Psychology, 90*, 114-141.

Franklin, A., Pilling, M. & Davies, I.R.L. (2004). Category effects in visual search: Evidence from eye-movement latencies. *Perception, 33,* 147.

Franklin, A., Pilling, M. & Davies, I.R.L. The nature of infant colour categorisation: Evidence from eye-movements on a target detection task. *Manuscript submitted for publication.*

Fujisaki, H., & Kawashima, T. (1971). On the modes and hearing mechanisms of speech perception. *Annual Report of the Engineering Research Institute, Vol. 28.* Faculty of Engineering, University of Tokyo.

Harnad, S. (1987). *Categorical Perception. The Groundwork of Cognition.* Cambridge: Cambridge University Press.

Heider, E. R. (1972). Universals in colour naming and memory. *Journal of Experimental Psychology, 93,* 10-20.

Heider, E. R., & Olivier, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology, 3,* 337-354.

Hunt, R. W. G. (1987). *Measuring colour.* Chichester: Ellis Horwood.

Kawai, M., Uchikawa, K., & Ujike, H. (1995). Influence of color category on visual search. *Investigative Ophthalmology, 36*, 654.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist, 86,* 65-78.

Laarni, J. (2001). Colour-cueing in visual search. *Scandinavian Journal of Psychology*, *42*, 71-78.

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of the optical Society of America, 7,* 1209-1217.

Nagy, A.L., & Winterbottom, M. (2000). The achromatic mechanism and mechanisms tuned to chromaticity and luminance in visual search. *Journal of the Optical Society of America, A: Optics and Image Science*, *17*, 369-379.

Özgen, E., & Davies, I. R. L. (2002). Acquisition of categorical color perception: A perceptual learning approach to the linguistic relativity hypothesis. *Journal of Experimental Psychology: General, 131,* 477-493.

Pilling, M. (2001). On colour categorisation. *Unpublished PhD thesis.* University of Surrey.

Pilling, M., Wiggett, A., Özgen, E., & Davies, I. R. L. (2003). Is color "categorical perception" really perceptual? *Memory and Cognition, 31*, 538-551.

Pollack, I., & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, *1,* 125-126.

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory and Cognition, 28,* 977-986.

Roberson, D., Davies, I. R. L., & Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone-age culture. *Journal of Experimental Psychology: General, 129*, 369-398.

Rosch, E. H. (1973). Natural categories. *Cognitive Psychology, 4,* 328-350.

Rosen, S., & Howell, P. (1987). Auditory, articulatory, and learning explanations of categorical perception in speech. In Harnad S. (Ed.), *Categorical perception. The groundwork of cognition.* Cambridge: Cambridge University Press.

Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97-136.

Whorf, B. L. (1956). *Language, thought and reality: Essays by B. L. Whorf.* J. B. Caroll (Ed.). Cambridge: Cambridge University Press.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*, 33-39.

Footnotes

[1.] Munsell is a colour system standardised to produce perceptually equal steps. Munsell colours vary along three dimensions: hue, value (lightness), and chroma. Equal sized steps along a single dimension result in equal perceptual differences. Similarly, CIE (1976) is a perceptually uniform colour space, rather like the classical colour circle or colour solid if lightness is included. The CIE dimensions are L^* (lightness) u^* (roughly the red-green axis), and v^* (the blue-yellow axis). Chroma and saturation increase with distance from the origin (grey). Perceptual distances are given by the Euclidean distance in the space (ΔE).

[2.] The red-pink stimulus separation sizes were 10 ΔE units less than the blue-green and blue-purple. Therefore, if ΔE contributes to search speed, we would expect search to be slower for the red-pink set. This is the case for the English times but not the Kwanyama. It is unclear why there is no difference in search time across sets for the Kwanyama.

Table 1. Predicted patterns of naming (Experiment 1). Predicted patterns of naming (Experiment 2). The perceptual distances in CIE units (ΔE) of the stimulus pairs for each colour set are also shown. The Kwanyama colour terms have the prefix 'Oshi-', meaning 'the colour of'.

Colour set	English term	Kwanyama term
Blue-Purple	Blue	Oshitwima
(ΔE=30)	Blue	Oshitwima
	Purple	I don't know
Blue-Green	Green	Oshitwima
(ΔE=30)	Green	Oshitwima
	Blue	Oshitwima
Red-Pink	Red	Oshitilyana
(ΔE=20)	Red	Oshitilyana
	Pink	I don't know

Table 2. Terms given to each stimulus (A, B, C) in each colour set by the English and Kwanyama groups with the percentage of each sample that gave that term. Table 2a: Blue-Purple set; Table 2b: Blue-Green set; Table 2c: Red-Pink set. The expected term (and most frequent) is shown in bold (IDK = 'I don't know'). A specific reference is when an object name is offered as a colour term. For example, *olaole* which translates as 'earthworm' was offered for stimuli from the red-pink set.

Table 2a.

Colour set	Stimulus	English Naming		Kwanyama Naming	
Blue-Purple	**A**	**Blue**	**100.0**	**Oshitwima**	**60.9**
				Specific reference	23.9
				IDK	15.2
	B	**Blue**	**100.0**	**Oshitwima**	**39.1**
				Specific reference	37.0
				IDK	23.9
	C	**Purple**	**96.9**	**IDK**	**50.0**
		Blue	3.1	Specific reference	28.3
				Oshitwima	21.7

Table 2b.

Colour set	Stimulus	English Naming		Kwanyama Naming	
Blue-Green	**A**	**Green**	**100.0**	**Oshitwima**	**69.5**
				Specific reference	17.4
				IDK	13.0
	B	**Green**	**98.4**	**Oshitwima**	**67.4**
		Blue	1.6	Specific reference	19.5
				IDK	13.0
	C	**Blue**	**98.4**	**Oshitwima**	**52.2**
		Green	1.6	IDK	28.3
				Specific reference	17.4

Table 2c.

Colour set	Stimulus	English Naming		Kwanyama Naming	
Pink-Red	**A**	**Red**	**100.0**	**Oshitilyana**	**82.6**
				IDK	13.1
				Specific reference	4.3
	B	**Red**	**92.2**	**Oshitilyana**	**47.8**
		Pink	7.8	IDK	38.0
				Specific reference	15.2
	C	**Pink**	**100.0**	**IDK**	**58.6**
				Specific reference	30.4
				Oshitilyana	10.9

Table 3. Mean search times in secs (standard deviations in brackets) for within and cross-category arrays, for each colour set and for each language group (Experiment 1).

		English				Kwanyama			
Set	Age	Within		Cross-		Within		Cross-	
Blue-Purple									
	4	35.9	(12.8)	30.7	(12.3)	27.4	(9.6)	26.0	(11.8)
	5	32.4	(5.2)	27.9	(5.3)	35.2	(11.5)	30.8	(6.3)
	6	28.8	(8.5)	24.6	(7.8)	22.5	(10.5)	20.4	(7.2)
	7	19.0	(4.3)	16.4	(3.2)	19.0	(4.9)	17.4	(5.1)
	All	29.0	(10.3)	24.9	(9.4)	25.4	(10.7)	23.2	(9.3)
Blue-Green									
	4	35.1	(13.3)	34.4	(14.6)	24.7	(5.9)	23.1	(4.4)
	5	32.9	(4.2)	28.3	(8.5)	28.1	(12.9)	32.8	(14.0)
	6	28.1	(9.2)	25.1	(6.8)	21.7	(7.0)	20.1	(5.3)
	7	20.0	(6.5)	16.8	(4.1)	20.0	(5.7)	19.2	(4.2)
	All	29.0	(10.5)	26.2	(11.2	23.3	(8.3)	23.2	(8.8)
Pink-red									
	4	43.5	(9.2)	34.3	(8.3)	24.6	(7.2)	25.1	(11.6)
	5	39.3	(12.1)	29.0	(9.5)	35.2	(9.1)	33.8	(9.3)
	6	38.5	(9.0)	30.0	(9.3)	22.9	(8.8)	22.6	(10.9)
	7	30.1	(11.1)	20.7	(5.1)	22.5	(7.8)	17.5	(3.4)
	All	37.8	(11.3)	28.5	(9.4)	25.7	(9.3)	24.1	(10.7)

Table 4. Mean search times in seconds (SD) for within and cross-category arrays, for each set (Experiment 1).

	Set type		
Category	Blue-purple	Blue-green	Pink-red
Within	28.7 (11.1)	21.8 (4.9)	26.4 (10.1)
Cross-	25.5 (7.9)	21.1 (5.3)	27.7 (11.8)

Table 5. Predicted patterns of naming (Experiment 2). The perceptual distances in CIE units (ΔE) of the stimulus pairs for each colour set are also shown. The Himba colour terms have the prefix 'Otji'.

Colour set	English term	Himba term (Otji-)
Blue-Green	Green	Mburou
(ΔE=20)	Green	Mburou
	Green	Mburou
	Blue	Mburou
	Blue	Mburou
	Blue	Mburou
LightGreen-Yellow	Green	Mburou
(ΔE=20)	Green	Mburou
	Green	Mburou
	Yellow	Ndumbu
	Yellow	Ndumbu
	Yellow	Ndumbu

Table 6. Frequencies (%) of terms given for each stimulus in each set by English and Himba children in Experiment 2 (Table 6a: G stimuli in Blue-Green set; Table 6b: B stimuli in Blue-Green set; Table 6c: LG stimuli in Light Green-Yellow set; Table 6d: Y stimuli in Light Green - Yellow set). The most frequent term is shown in bold (IDK = 'I don't know'). All Himba names begin with the prefix 'Otji-'.

Table 6a.

Colour set	**Stimulus**	**English Naming**		**Himba Naming (Otji-)**	
Blue-Green	**G3**	**Green**	**93.75**	**Ngirine**	**44**
		Blue	6.25	Mburou	32
				Ndumbu	16
				Hui	4
				Binde	4
	G2	**Green**	**100.0**	**Ngirine**	**44**
				Mburou	40
				Ndumbu	4
				Hui	4
				Binde	4
				Zoozu	4
	G1	**Green**	**100.0**	**Mburou**	**48**
				Ngirine	28
				Ndumbu	12
				Hui	4
				Binde	4
				IDK	4

Table 6b.

Colour set	Stimulus	English Naming		Himba Naming (Otji-)	
Blue-Green	**B1**	**Blue**	**100.0**	**Mburou**	**72**
				Ngirine	16
				Ndumbu	4
				Hui	4
				IDK	4
	B2	**Blue**	**100.0**	**Mburou**	**56**
				Binde	12
				Ngirine	8
				Zoozu	8
				Ndumbu	8
				Hui	4
				IDK	4
	B3	**Blue**	**96.88**	**Mburou**	**68**
		Green	3.12	Ngirine	16
				Ndumbu	4
				Hui	4
				Zoozu	4
				IDK	4

Table 6c.

Colour set	Stimulus	English Naming		Himba Naming (Otji-)	
LightGreen-Yellow	**LG3**	**Green**	**100.0**	**Ngirine**	**36**
				Ndumbu	36
				Mburou	20
				Binde	4
				IDK	4
	LG2	**Green**	**100.0**	**Ndumbu**	**60**
				Ngirine	16
				IDK	12
				Mburou	8
				Binde	4
	LG1	**Green**	**100.0**	**Ndumbu**	**88**
				Mburou	4
				Ngirine	4
				IDK	4

Table 6d.

Colour set	Stimulus	English Naming		Himba Naming (Otji-)	
LightGreen-Yellow	**Y1**	**Yellow**	**100.0**	Ndumbu	**92**
				Mburou	4
				Ngirine	4
	Y2	**Yellow**	**100.0**	Ndumbu	**84**
				Vapa	8
				Ngirine	4
				IDK	4
	Y3	**Yellow**	**100.0**	Ndumbu	**80**
				Vapa	16
				Ngirine	4

Figure Captions

Figure 1. Cross- and within-category stimuli used to demonstrate CP. The six stimuli (G3, G2, G1, B1, B2, B3) are equidistant in colour space. The category boundary lies between G1 and B1. Discrimination is better between G1 and B1 (cross-category pair) than between B1 and B2 (within-category pair), yet the G1-B1 and B1-B2 perceptual distances are the same.

Figure 2. Mean within and cross-category search times for English and Kwanyama (error bars are +/- 1 standard error).

Figure 3. Mean search times for each age group, for English and Kwanyama (error bars are +/- 1 standard error).

Figure 4. Mean search times for each set, for each language group, (error bars are +/- 1 standard error).

Figure 5. Mean within and cross-category search times for the English children (a) and the Kwanyama children (b), for each set, (error bars are +/- 1standard error).

Figure 6. Mean within- and cross-category search times for the green-blue and green-yellow sets, as a function of set size (30 or 48 colours), in the Himba (6a) and the English (6b) group. Error bars are +/- 1 standard error.

Figure 7. Mean within- and cross-category accuracy (A') for the green-blue and green-yellow sets, as a function of set size (30 or 48 colours), in the Himba (7a) and the English (7b) group. Error bars are +/- 1 standard error.

Figure 1

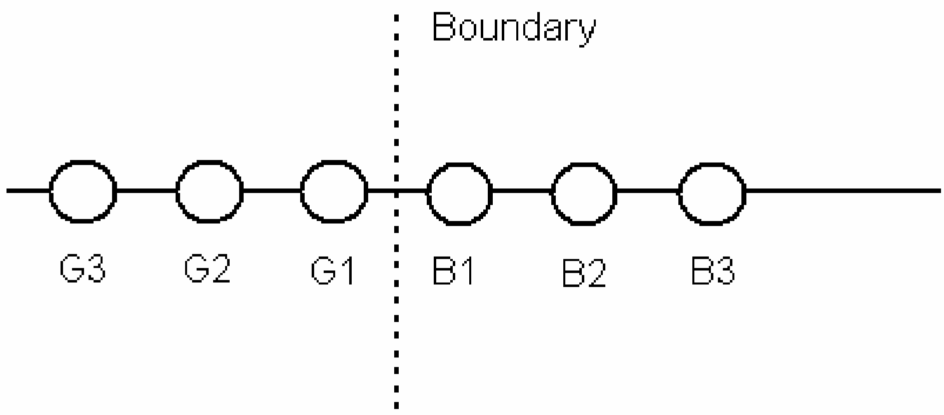

Figure 2.

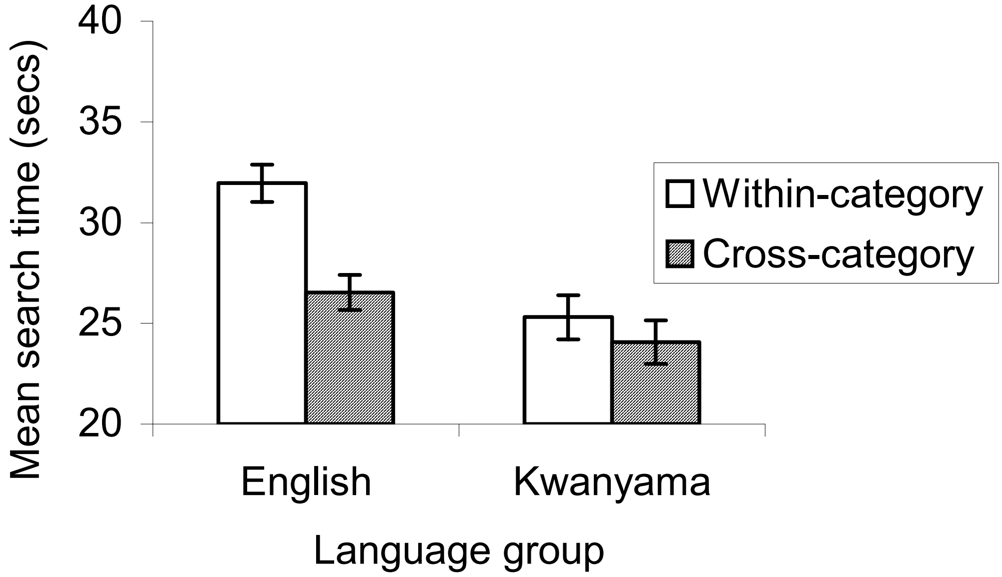
40
35
30
25
20
Mean search time (secs)
Within-category
Cross-category
English
Kwanyama
Language group

Figure 3.

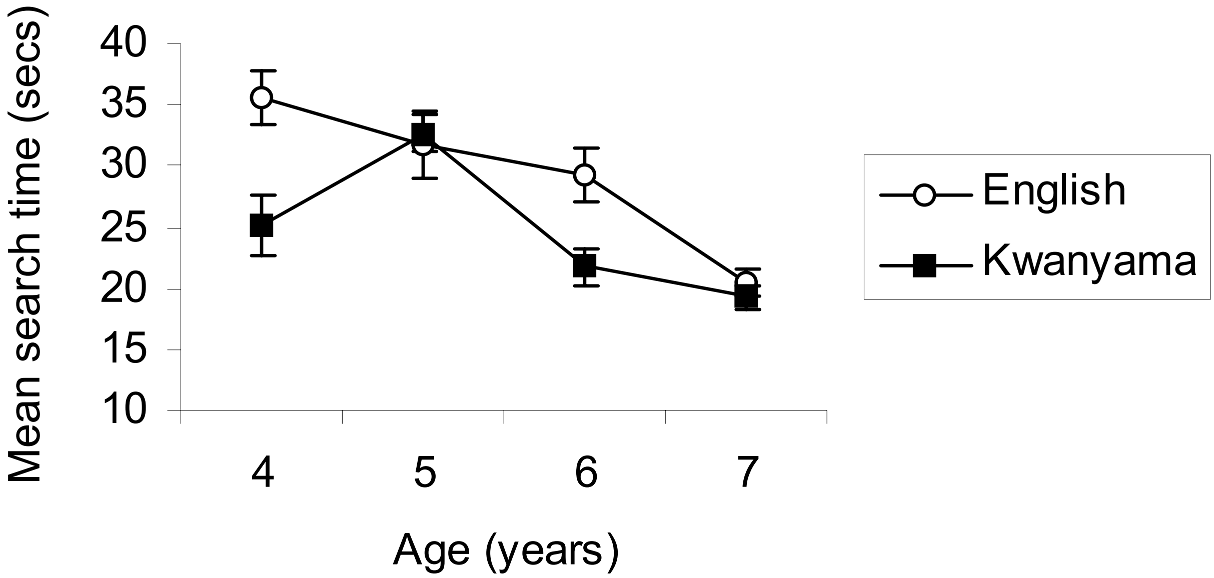
Mean search time (secs)
40
35
30
25
20
15
10
4
5
6
7
Age (years)
English
Kwanyama

Figure 4.

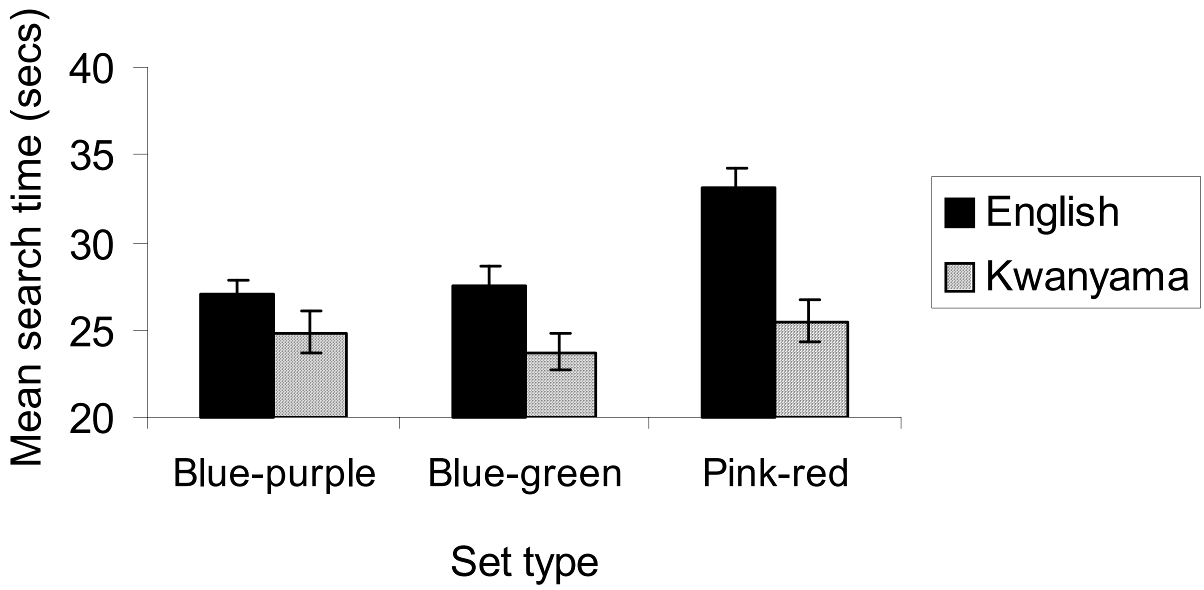
Mean search time (secs)
40
35
30
25
20
Blue-purple
Blue-green
Pink-red
Set type
English
Kwanyama

Figure 5.

Figure 5a.

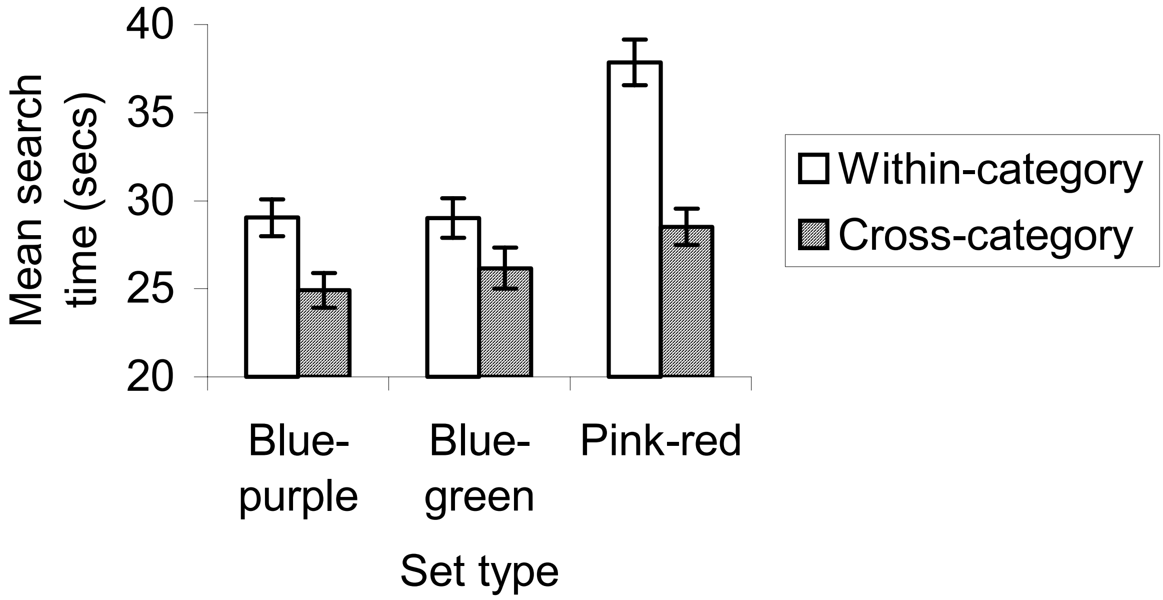

Fig. 5b.

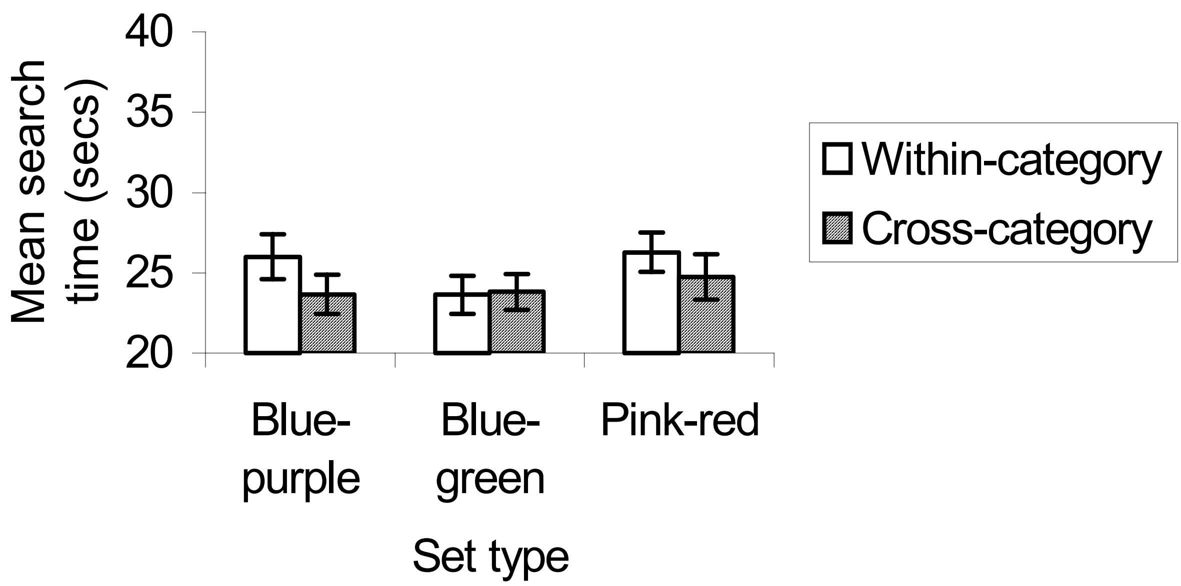

Figure 6.

Fig 6a.

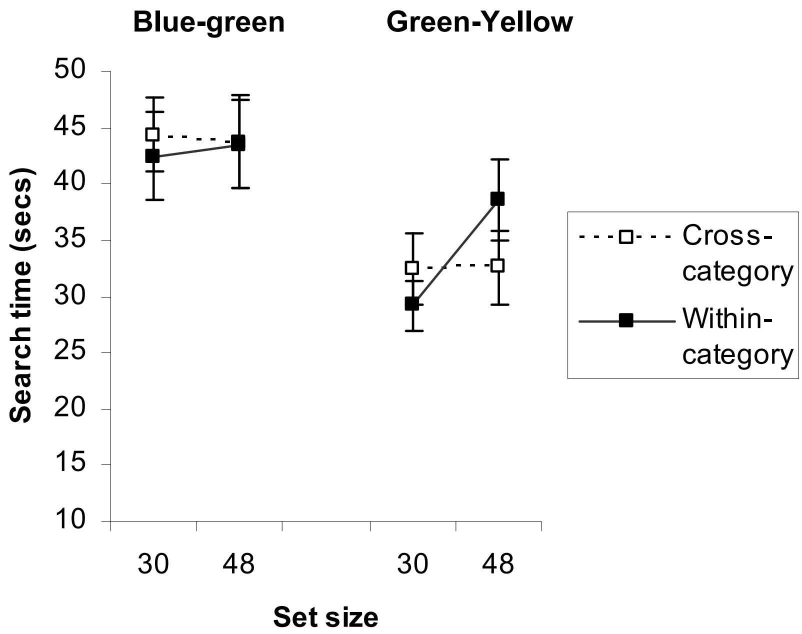

Fig 6b.

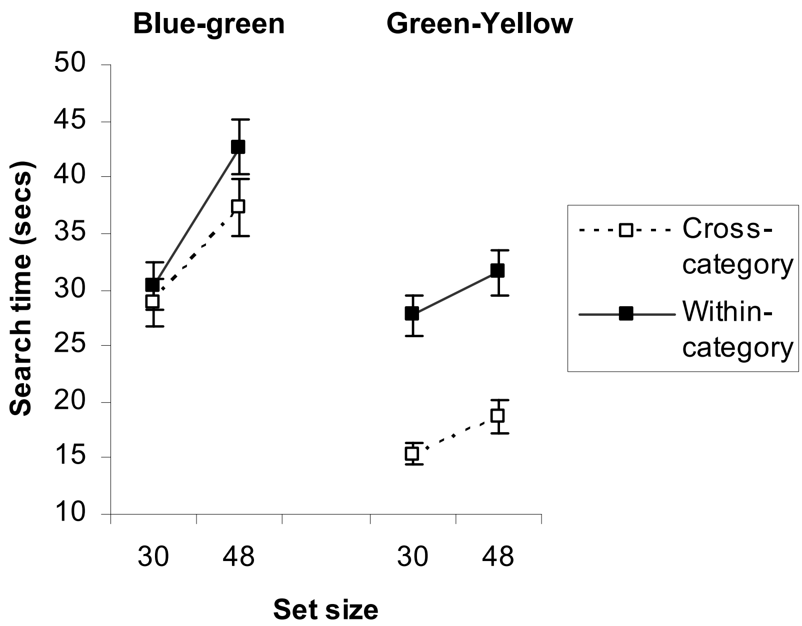

Figure 7.

Fig 7a.

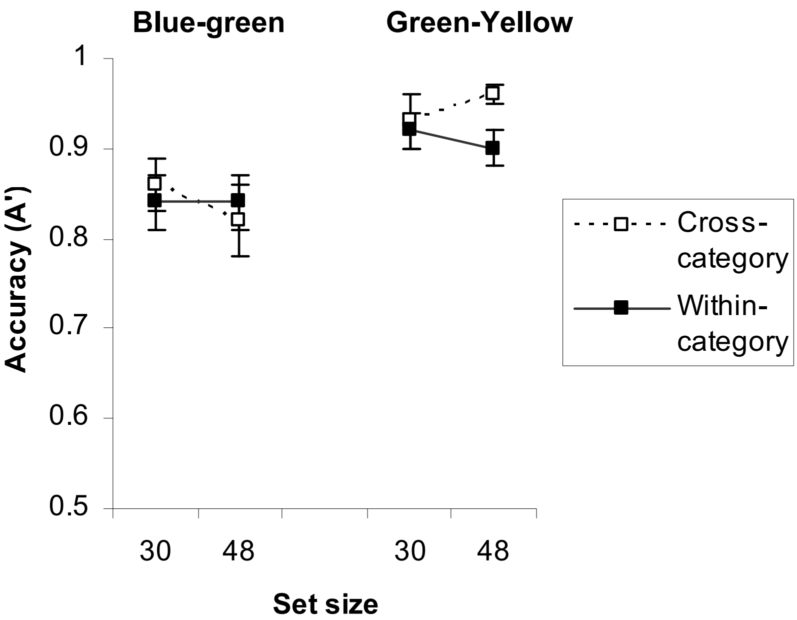

Fig 7b.