Developmental trajectories of part-based and configural object recognition in adolescence

Martin Jüttner[1], Elley Wakui[2], Dean Petters[1], Surinder Kaur[1] and Jules Davidoff[2]

1. Psychology, School of Life and Health Sciences, Aston University, Birmingham, UK

2. Department of Psychology, Goldsmiths College, London, UK

Address for correspondence: Martin Jüttner, Psychology, School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, UK; email: m.juttner@aston.ac.uk

Abstract

Three experiments assessed the development of children's part and configural (part-relational) processing in object recognition during adolescence. In total 312 school children aged 7-16 and 80 adults were tested in 3-AFC tasks to judge the correct appearance of upright and inverted presented familiar animals, artifacts, and newly learned multi-part objects, which had been manipulated either in terms of individual parts or part relations. Manipulation of part relations was constrained to either metric (animals, artifacts, and multi-part objects) or categorical (multi-part objects only) changes. For animals and artifacts, even the youngest children were close to adult levels for the correct recognition of an individual part change. By contrast, it was not until 11-12 years that they achieved similar levels of performance with regard to altered metric part relations. For the newly-learned multi-part objects, performance was equivalent throughout the tested age range for upright presented stimuli in case of categorical part-specific and part-relational changes. In case of metric manipulations, the results confirmed the data pattern observed for animals and artifacts. Together the results provide converging evidence, with studies of face recognition, for a surprisingly late consolidation of configural-metric relative to part-based object recognition.

Keywords: development, object recognition, face recognition, configural, relational, part, geon, recognition-by-components

Introduction

In this paper, we shall show that children do not always recognise objects in the same way as adults. Similar claims have been made for face perception where despite indications of remarkable face recognition skills in young infants (Pellicano & Rhodes, 2003; McKone & Boyer, 2006; de Heering, Houthuys, & Rossion, 2007) there is substantial evidence that these skills continue to significantly improve during adolescence and deep into the second decade of life (e.g., Ellis, 1975; Carey & Diamond, 1977; Carey, Diamond, & Woods, 1980; Mondloch, Le Grand, & Maurer, 2002). While there is still some dispute over the developmental trajectory for face recognition (e.g., Tanaka, Kay, Grinell, Stansfield, & Szechter, 1998) that for object recognition is even more uncertain. Recent neuroimaging studies (Golarai, Ghahremani, Whitfield-Gabrieli, Reiss, Eberhardt, Gabrieli, & Grill-Spector, 2007; Scherf, Behrmann, Humphreys, & Luna, 2007; see also Aylward, Park, Field, Parsons, Richards, Cramer, & Meltzhoff, 2005; Gathers, Bhatt, Corbly, Farley, & Joseph, 2004) conclude that for objects the trajectory is steeper than for face recognition and approaches adult levels of performance at an age as early as 7 years. However, behavioural studies give a substantially different picture with a retarded trajectory much like that which has been concluded for face recognition (Davidoff & Roberson, 2002; Rentschler, Jüttner, Osman, Müller, & Caelli, 2004; Jüttner, Müller, & Rentschler, 2006). These studies highlight the importance of the processing of relations that may hold between the parts or components constituting an object. In our paper we will refer to such part-relational processing as 'configural processing', in order to contrast it with the processing of part-specific, i.e. featural,

information (cf. Maurer, Le Grand, & Mondloch, 2002) and to highlight any emerging

parallels between the development of object and face recognition[1].

Davidoff and Roberson (2002) asked about the correct appearance of animals that

children as young as 5 years could name successfully. Participants were presented in each

trial with three variations of the same animal, of which only one was completely correct.

The incorrect alternatives could either involve part changes, derived by substituting one

part of an animal with that from another, or configural changes, here defined by either an

alteration of the size-ratio of parts or the aspect ratio of the animal shape. Calibration of

the part-changed and configural-changed stimuli ensured that adults found them of equal

difficulty. The results showed that it was not until 11 years that children were at adult

levels for the correct recognition of a part change and not until 15-16 years for the

recognition of a configural change.

Rentschler et al. (2004) and Jüttner et al. (2006) observed a similar retardation of

configural processing in the rather different task of cross-modal object learning. They

investigated spatial generalization skills in children aged 8-16 using a set of three novel,

molecule-like objects, each of which consisted of four spheres and could only be

distinguished in terms of the spatial configuration of these components. Participants were

first trained to correctly classify a set of learning views of the three objects and

subsequently tested for their spatial generalization ability by having to classify a set of

---

[1] As Maurer et al. note in their review the term 'configural''' has received different interpretations,
particularly in the face recognition literature, but there is no consensus about its use. We therefore follow
their broad convention to use "the term 'configural processing' to refer to any phenomenon that involves
perceiving relations between the features of a stimulus" (Maurer et al., 2002, p. 255).

novel views. Both learning and generalization performance were found to significantly

improve throughout the tested age range. In addition, the results indicated a

developmental dissociation of view-dependent and view-independent object

representations, with younger children recognizing the objects only from previously

learned perspectives whereas older children generalized to novel viewpoints as well. In

our studies we hope to help clarify the developmental trajectory for object recognition by

examining it in a framework that explicitly considers object parts and their relations

(Biederman, 1987; Hummel & Stankiewicz, 1996; Hummel, 2001).

Biederman's (1987) Recognition-by-components (RBC) model has motivated several of

the recent studies on object processing by infants and children. According to this model,

complex objects are encoded as spatial arrangements, or configurations, of basic parts

that come from a restricted set of elementary shapes, so-called *geons*. Geons are defined

by categorical contour properties (like parallel vs nonparallel or straight vs curved) - *non-

accidental properties* (NAPs) - that are largely invariant to changes in viewpoint.

Similarly, the spatial configuration of geons is encoded in terms of certain categorical

relations between geons (like 'on top of' or 'besides'). Furthermore, Biederman (cf. also

Cooper, Biederman & Hummel, 1992) contrasts shape differences in terms of NAPs with

those arising from continuous, or *metric*, variations of part and part-relational (i.e.,

configural[2]) properties (for example, the degree of non-parallelism within the contours of

a given object part, or the precise distance between two parts). With respect to visual

recognition, Biederman (2000) suggested that metric comparisons recruit different

---

[2] Biederman only speaks of part relations but according to our adopted convention the terms 'part-relational' and 'configural' can be regarded as synonyms.

mechanisms of recognition than non-accidental comparisons. Likewise, later dual-route variants of the RBC model propose that analytic, structural object descriptions are augmented by non-analytic, view-based object representations (Hummel & Stankiewicz, 1996; Hummel, 2001). It has been suggested that the relative preponderance of analytical and non-analytical representations may depend on attentional allocation (Hummel, 2001; cf. also Thoma, Hummel, & Davidoff, 2004) and further be modulated by object category (Cooper & Brooks, 2004) or rigidity (Laeng, Carlesimo, Caltagirone, Capasso, & Miceli, 2002).

Within the context of the RBC framework, developmental differences for part and configural changes could indicate different trajectories for part and part-relational features, both of which may further dissociate into different paths for processing non-accidental and metric attributes. Previous developmental work considering the role of structural object descriptions has primarily focussed on the status of individual parts. There is ample evidence that parts receive particular attention in the analysis and detection of shape similarity (e.g., Tversky & Hemenway, 1984; Schyns & Murphy, 1994; Saiki & Hummel, 1996; Rakison & Cohen, 1999). The preference for using part information appears to emerge very early in life; toddlers and even infants have been shown to attend selectively to parts when categorizing or matching objects (Madole & Cohen, 1995; Smith, Jones, & Landau, 1996; Rakison & Butterworth, 1998, Abecassis, Sera, Yonas, & Schwade, 2001; Haaf, Fulkerson, Jablonski, Hupp, Shull, & Pescara-Kovach, 2003). Unlike for parts, very little developmental work has considered the processing of object part relations within an RBC framework. In the one known study

Mash (2006) examined similarity judgements of novel object images differing by a metric part and a part-relational property in children aged 5 yrs and 8 yrs, as well as in adults. Young children were found to show a strong bias for classifying objects on the basis of part specific information only. With increasing age participants came to select both part-specific and part-relational information in their classification judgements. These results could be interpreted as evidence for a retarded processing of part-relational relative to part-specific processing. However, it remains unclear whether the observed developmental differences are confined to tasks involving a perceptual online classification (i.e., of simultaneously available objects) as opposed to those of recognition (i.e., the matching of a sensory percept to a stored object representation).

Given this background the current experiments had two principal aims. The first aim was to systematically assess and compare developmental trajectories of part-based and configural object recognition across adolescence, extending them into an age range (7-16 years) well beyond that considered in relevant previous research with its focus on the pre-school child. The second aim was to employ both naturalistic and novel, multi-part geon stimuli in converging attempts to test specific predictions of geon-based object recognition theories within a developmental context, and to clarify the basis of the retarded configural processing reported in earlier work (Davidoff & Roberson, 2002; Rentschler et al., 2004; Jüttner et al., 2006). In Experiment 1 we asked whether the observed delays in configural processing for non-rigid objects (animals; Davidoff & Roberson, 2002) would generalise to rigid objects (artefacts). In Experiments 2 and 3 we used a specifically designed set of novel objects to compare developmental changes in

the processing of non-accidental (Experiment 2) and metric (Experiment 3) properties for individual parts and part relations, thus further confining the origin for the developmental differences between part-based and configural object recognition.

Experiment 1

In our first experiment we assessed the generality of the tardy development of configural object recognition which we had previously observed for animals (Davidoff & Roberson, 2002). It has been suggested that non-rigid objects (like animals) may require a different representational format than rigid objects (Laeng et al., 2002; Cooper & Brooks, 2004). For example, Laeng et al. argued that recognition of rigid objects relies on transformations of a holistic object representation without the need for structural descriptions. In contrast, non-rigid objects may require structural descriptions specifying the relationships between parts in order to permit recognition from non-canonical perspectives. Thus, our primary concern in Experiment 1 was to test whether the developmental changes found for animal recognition extend to a different type of object with a potentially different type of mental representation.

We compared two types of stimuli: A set of animals (drawn from the original stimulus set of Davidoff & Roberson, 2002) and a set of rigid artefacts. We chose artefacts that could be defined to have a base (similar to animals) in order to avoid potential confounding effects arising from the recognition of non-base objects (see Davidoff & Warrington,

1999; Verfaille & Boutsen, 1995). For both sets of objects, we examined the developmental trends associated with the identification of the correct picture of an animal/artefact from a set of three (Davidoff & Warrington, 1999): the original image and two distracters. Two types of distracters were compared; the first involved a change of an individual object part (Part Change condition) while the second involved a change of the overall proportions of the object (Configural Change condition). The two types of distracter stimuli were calibrated such that adults found them of equal difficulty. The calibrated stimuli were then used to assess recognition performance in school children aged 7-16 years.

As a further manipulation we presented the stimuli upside down. Impairment of performance by inversion has been used as one indicator for configural processing, in particular in face recognition (Carey & Diamond, 1977). Even though the disruption may apply more to objects that have - like faces - an internal part structure (Yin, 1969) it has been shown that inversion has an detrimental effect on recognition for many types of stimuli including those without internal features (e.g., de Gelder et al., 1998; Bruyer & Crispeels, 1992; McLaren, 1997).

Method

Participants

8

Six age groups took part in the experiment, each consisting of 32 participants: The groups were adult volunteers (17 females and 15 males; mean age 20 years 7 months), 7- to 8-year-olds (15 females and 17 males; mean age 7 years 2 months), 9- to 10-year-olds (16 females and 16 males; mean age 9 years 6 months), 11- to 12-year-olds (18 females and 14 males; mean age 11 years 6 months), 13- to 14-year-olds (15 females and 17 males; mean age 13 years 3 months), and 15- to 16-year-olds (17 females and 15 males; mean age 15 years 2 months). The children were drawn from state schools in Birmingham and London, UK. The adults were recruited among undergraduate students at Goldsmiths College, London. They received course credit for participation.

Stimuli and Procedure

A subset of 28 animal stimuli was drawn from the original set in Davidoff and Roberson (2002). The stimuli were derived from digitized high-resolution photographs taken by professional photographers. For each picture in this set, two manipulated versions were created using Adobe Photoshop. For 14 of the items, the manipulation involved a part change, for the other 14 they involved a configural change (Figure 1A, see Appendix A for a stimulus list). Part changes implied the substitution of the correct part with that taken from another animal (e.g., tail or feet). Configural changes involved the whole animal and altered its proportions. This was done either directly, by changing the relative size of body parts (in 7 out of 14 stimuli), or indirectly, by stretching the animal shape horizontally (5) or vertically (2). The items were chosen such that recognition accuracy of the 32 adult controls were matched in the Part change and Configural Change condition $(t(31) = .11, p = .91$; see Appendix C for further details on the calibration procedure).

For the artefact stimuli, a set of 28 images was drawn from a public internet database (Google–Images). The images showed easily recognisable everyday objects (e.g., a kettle, airplane) with a base. For each item two manipulated versions were created using the same procedure as with the animal stimuli. For 14 items this involved a part change, for the other 14 items it involved a configural change (Figure 1B, see also Appendix B). The two sets of part and configurally manipulated items were calibrated for equal difficulty in the adult control subjects [$t(31) = 1.36$, $p = .19$; see Appendix C for details on the calibration procedure].

Participants were tested using a one-in-three (3-AFC) selection task (Davidoff & Warrington, 1999; Davidoff & Roberson, 2002). In each trial, three images labelled A,B,C were presented on the screen, one original and two distracter stimuli (involving either a part or a configural change). The three images were presented vertically on a 15.4 inch monitor with a screen size (height x width) of 20.4 x 33 cm, subtending a visual angle of 23.3 x 37.8 deg at the viewing distance of 50 cm. The pictures appeared at a mean size of 4.6 x 7.1 deg with a centre-to-centre distance between images of 6.9 deg. For stimuli involving part changes the manipulated part had a mean size of 2.5 deg of visual angle. The observer had to choose the 'correct' depiction of the animal or artefact

10

by pressing the appropriately marked button (A, B, C) on the keyboard. The correct version appeared in each of the three possible locations equally often.

In case of configural manipulations, for half the stimuli the two distracters involved changes of proportions in opposite directions (for example, by enlarging the head of an animal relative to its trunk in one version and reducing it in the other), for the other half the changes were in the same direction. Thus the 'correct' image could be either in the middle of the 'stretching spectrum' or at one of its endpoints, with equal probability.

To maintain children's interest, the task was introduced as a game, where the experimenter had played with pictures on the computer and altered pictures of animals or objects in some way. On each screen they would see three pictures of animals or objects of which only one would be the correct image. Their task would be to identify the correct item (e.g., "Spot the one that the animal really looks like") as the top (A), middle (B) or bottom (C) image by pressing the corresponding response key. The task instructions were shown on the screen and also read out aloud by the experimenter.

There were no instructions regarding response speed. The stimulus remained on the screen until the participant had responded. Response time and accuracy were measured as dependent variables. Participants were also asked to name the animal/object; none ever failed to do that. Stimulus presentation and response collection were controlled by an Eprime 1.1 (Psychology Software Tools, Inc.) script running on a laptop computer.

Participants were tested individually on both the animals and artefact images in separate sessions and counterbalanced order. Each session was divided into blocks of 14 trials involving either a part or a configural change, and either an upright or an inverted stimulus presentation. Participants were instructed not to attempt to rotate their head to see the rotated pictures. The order of the resulting four presentation conditions (part–upright, part–inverted, configuration–upright, configuration–inverted) was counter-balanced across subjects. To illustrate the task in the four conditions, two practice trials with feedback preceded each block. Animals and artefacts appearing as practice items (cf. Appendices A and B) were not used in the main experiment.

Results

Performance was analysed in terms of the accuracy and the latency preceding a correct response. To check whether the accuracy data had been affected by any extreme values for individual stimuli, the distributions of scores were checked for outliers. No outliers (defined by the group mean ±3 standard deviations) were observed in any of the age groups. Regarding response times, our instructions did not promote fast responses, therefore some participants showed particularly long latencies. Three participants with latencies classified as outliers were removed from the data set prior to the statistical analysis (one participant in each of the age groups 9 – 10 yrs, 13 – 14 yrs, and 15 – 16 yrs).

Accuracy

Figure 2A shows means and standard errors of the recognition accuracy (combined across animals and artefacts) for each age group and for each of the two manipulation conditions (Part vs Configuration) and orientations (Upright vs. Inverted). The accuracy data were analysed in a 6 (Age: Adults vs. 15 - 16 yrs vs. 13 - 14 yrs vs. 11 - 12 yrs vs. 9 - 10 yrs vs. 7 - 8 yrs) x 2 (Stimulus type: Animals vs. Artefacts) x 2 (Manipulation: Part vs. Configuration) x 2 (Orientation: Upright vs. Inverted) mixed ANOVA with Age as the between factor. The analysis yielded significant main effects for Age [$F(5,183) = 25.29$, $p < .001$], Manipulation [$F(1,183) = 183.23$, $p < .001$) and Orientation [$F(1,183) = 32.82$, $p < .001$) but not for Stimulus Type ($p = .17$). In addition, there was a significant interaction between Age x Manipulation [$F(1,5) = 5.35$, $p < .001$]. With regard to performance in the Part Change condition, a contrast analysis with the adult data as reference category showed that 7 to 8-year-olds and 9 to 10-year-olds performed worse than adults ($ps < .02$) whereas there were no significant differences for older children ($ps > .8$). The effect of Age was more pronounced in the Configural Change condition. Here performance of the four youngest age groups (7 - 8 yrs, 9 - 10 yrs, 11 - 12 yrs, 13 - 14 yrs) was significantly lower than that of adults ($ps < .02$) – an effect that was still approaching significance in the 15- to 16-year olds ($p = .06$).

===========================

Insert Figure 2 about here

===========================

13

The dissociation between the developmental trajectories in the Part Change and Configural Change conditions raises the question of the impact of stimulus inversion. The initial ANOVA had shown no significant interaction of Orientation with Manipulation and Age [$F(5,182) = .78$, $p = .38$]. However, given the low recognition accuracy of young children for configural changes in the Upright condition, performance for inverted stimuli may have been affected by floor effects (note the chance level at .33 indicated by the dashed line in Figure 2A). For a more sensitive assessment of inversion effects, we conducted a second analysis that was confined to the asymptotic age range where, in both the Part Change and Configural Change condition, performance for upright stimuli no longer significantly differed (ps > .2) from adult observers. This ensured a common baseline for the comparison because adult performance in the two upright manipulation conditions had been matched in the initial calibration. For these age ranges, comprising the groups 11 - 12 yrs, 13 - 14 yrs, 15 - 16 yrs and adults, the ANOVA gave a significant interaction between orientation and manipulation ($F(1,122) = 3.88$, $p < .05$), indicating that inversion had a more detrimental effect on recognition in the Configural Change condition than in the Part Change condition. For all other factors except Age the same main effects proved significant as in the main analysis [Manipulation: $F(1,122) = 83.39$, $p<.001$; Orientation: $F(1,122) = 22.66$, $p<.001$]. At least for the plateau region of the trajectories shown in Figure 2a, performance differences between Part Change and Configural Change conditions show a differential impact of stimulus inversion.

The observed differences between part and configurally manipulated stimuli in young children could not be attributed to extreme values for individual stimuli. Across stimulus

14

sets, the distributions of accuracy scores were approximately symmetric (skewness < .5) for part changes and configural changes in the 9 to 10-year-olds. For 7 to 8-year-olds they were approximately symmetric for part changes and marginally negatively skewed (-.71) for configural changes, resulting in a median (.44) lower than the mean (.48). The latter can be attributed to the mitigating effect of some easier stimuli that had been deliberately included to keep up motivation even in the youngest participants. Our data therefore provides a conservative estimate of the performance difference between Part Change and Configural Change condition in young children.

Latency

Figure 2B shows means and standard errors of the latencies (combined across animals and artefacts) for each age group and for each of the two manipulation conditions (Part vs Configuration) and orientations (Upright vs Inverted). Speeded responses were not requested of participants but to dispel any issues of speed-accuracy trade-offs, the latencies were analysed, in analogy to the accuracies, in a 6 (Age) x 2 (Stimulus type) x 2 (Manipulation) x 2 (Orientation) mixed ANOVA with Age as the between factor. As for accuracy, the analysis yielded significant main effects for Age [$F(5,173) = 16.76$, $p <$ .001], Manipulation [$F(1,173) = 62.24$, $p < .001$) and Orientation [$F(1,173) = 45.31$, $p <$ .001) but not for Stimulus Type ($p = .26$). In addition, there was again a significant interaction between Age and Manipulation [$F(5,173) = 2.48$, $p < .05$]. Considering the differences between the two manipulation conditions at each age range by pairwise comparisons, part changes were recognised more slowly than configural changes in all age groups ($ps < .02$; paired t-test) except for 11- to 12-year-olds ($p = .21$).

Given the possibility of speed-accuracy trade-offs in the data of the 7- to 8-year-olds and

9- to 10-year-olds in conjunction with their distinct advantage for recognizing part

changes (cf. Figure 2A) we conducted a further test where we re-analysed the data of the

two youngest age groups by focussing on the participants with the fastest response times

in the Part Change condition. As a cut-off point we chose the largest percentile giving

non-significant latency differences between the Part Change and the Configuration

Change condition (7-8 years: 50th percentile, 9-10 years: 30th percentile, ps > 0.2, paired-

t-test). Table 1 summarises mean accuracy and latency data for these subsamples. In both

age groups, the differences in the recognition rate between the two manipulation

conditions in the Upright condition continued to be significant (ps <. 02) and the

accuracies appear only marginally different from the data in the original population (cf.

Figure 2A). We conclude that the higher recognition rate for part changes in young

children is not the result of a speed-accuracy trade-off.

============================

Insert Table 1 about here

============================

Discussion

The results of Experiment 1 re-confirm and substantially expand the evidence provided

by related previous work (Davidoff & Roberson, 2002) for a protracted development of

object representations that are sensitive to configural manipulations. The retardation is most conspicuous in the distinctly reduced recognition performance in 7- to 8-year-olds and 9- to 10-year-olds when the (correct) target item was presented alongside distracters involving a manipulation of part relations. Indeed, configural object recognition appears not fully consolidated even for older children in the age range 11 to 16. Children in this age range recognized configural changes in the upright condition equally to adults but were more affected by stimulus inversion. By contrast, the developmental trajectory for the recognition of part changes is distinctly different. For upright stimuli, even 7- to 8-year-old children were performing close to adults, and did not significantly differ from the latter beyond the age of 10.

Stimulus type had no significant effect on recognition accuracy or response latency in Experiment 1, nor did it interact with any other experimental variable. This result provides an important extension of the earlier findings of Davidoff and Roberson (2002), and suggests that the protracted development of configural processing skills affects object recognition in general, rather than being limited to objects of a certain class or domain (cf. Cooper & Brooks, 2004), or predisposed by object properties like rigidity (Laeng et al., 2002).

The latencies show that young children responded faster in the configural change than the part change condition. However, further analysis found no evidence that the reduced accuracy in the former condition could be attributed to a speed-accuracy trade-off. A possible explanation of the faster responses to configural changes might be that the

17

stimulus alternatives in that condition were visually more distinctive than those for part changes – a result of our calibration procedure that aimed at equating recognition performance rather than salience (cf. the examples shown in Fig. 1). Indeed, Davidoff and Roberson (2002, Experiment 2) demonstrated for their animal stimuli (also used in the present study) that, for all age groups, the part-changed versions were less discriminable than the configurally-changed versions in a same-different decision task. The difficulty of young children in the present study to recognize configural changes despite their saliency advantage is therefore even more surprising. It implies that our recognition paradigm, requiring observers to relate visual stimuli to stored memory representations, provides a conservative estimate for the delay in the developmental trajectory for configural relative to part-based object processing.

As outlined above, structural approaches like the RBC model (e.g., Biedermann, 1987) and its derivatives (e.g., Hummel, 2001) distinguish between non-accidental and metric properties, both with regard to part-specific and part-relational attributes. Owing to the complexity of real world stimuli, the part changes employed in Experiment 1 inevitably involved changes at both the NAP *and* the metric level though the configural changes were entirely at the metric level. It is therefore possible that children's difficulty was confined to dealing with configural metric changes. Thus, our real world stimuli do not permit us to further constrain the origin of the observed difficulty of children when dealing with configural information. In Experiment 2, we used a set of novel objects for which the different levels of attributes could be precisely controlled and systematically varied when generating distracters in the two manipulation conditions.

18

Experiment 2

In Experiment 2 we asked further questions concerning the child's ability to recognise configural changes. In the context of the RBC model, this reduced ability could be a generic feature applying to all configural processing, i.e., for both NAP and metric relational changes. Alternatively, it could apply to metric changes only as in Experiment 1, in which case there should be no developmental retardation for detecting NAP changes. Thus, we assessed the trajectory for recognizing configural NAP manipulations and contrasted this manipulation with that of NAP part changes. Configural changes at NAP level would render natural objects grotesque and too distinctive to calibrate them with part changes for equal difficulty. In order to overcome this difficulty we employed a set of novel stimuli in which NAP manipulations could be carefully controlled. We focussed on changes of the attribute 'relative size' - in line with the type of relational manipulation employed in Experiment 1. For this attribute, a NAP configural change would be a categorical change from 'larger' to 'smaller' for the relative size of two parts. As in Experiment 1 we also included a condition in which the test stimuli were presented upside down.

Method

Participants

Six age groups took part in the experiment, each consisting of 24 participants : The groups were adult volunteers (13 females and 11 males; mean age 19 years 8 months), 7- to 8-year-olds (12 females and 12 males; mean age 7 years 4 months), 9- to 10-year-olds (10 females and 14 males; mean age 9 years 4 months), 11- to 12-year-olds (13 females and 11 males; mean age 11 years 5 months), 13- to 14-year-olds (12 females and 12 males; mean age 13 years 2 months), and 15- to 16-year-olds (11 females and 13 males; mean age 15 years 3 months). The children were drawn from state schools in Birmingham and London, UK. The adults were recruited among undergraduate Psychology students at Aston University. They received course credit for participation.

Materials

A set of six compound objects was created, each consisting of three parts (Figure 3). The parts were taken from a reservoir of three-dimensional shape primitives (geons) with unique combinations of non-accidental contour properties. We constrained these properties to a subset of the attributes suggested by Biederman (1987) and Hummel (2001), characterizing the type cross section (straight vs curved), the shape of the main axis (straight vs curved), and the surface along the main axis (parallel versus expanding vs convex vs concave). For example, the NAP signature of a cube would be a straight cross-section, a straight axis and parallel surfaces; the signature of a cone would be a curved cross section, a straight axis and expanding surfaces.

==============================

20

Insert Figure 3 about here

==============================

Within each object, parts were uniquely arranged in configurations that could be characterized by the relational NAP properties *relative position* 'on top of' vs. 'beside') and *relative size* ('larger' vs. 'equal' vs. 'smaller'). For example, object 3 could be described as consisting of a curved cylinder beside a smaller truncated cone, with the latter sitting on top of an equally-sized cube.

Within the learning set, objects 1 and 2 and objects 3 - 6 formed two subsets, referred to as *facilitator* objects and *probe* objects, respectively. Objects 3 - 6 consisted of the same three parts (either two bigger and one smaller, or two smaller and one bigger), employed the same spatial structure (involving one 'beside' plus one 'on top of' relation) and were characterized – as a result of the combination of part sizes - by two different relative size relations. Thus, these objects could not be identified on the basis of a single (diagnostic) part but required consideration of their overall shape, i.e., the spatial configuration formed by all their geon components. By contrast, objects 1 and 2 consisted of a different set of geons arranged in a distinctive horizontal or vertical configuration. During the learning phase of the experiment (cf. section *procedure*), the inclusion of the (relatively easy discriminable) facilitator objects served the purpose of maintaining motivation in children during the supervised learning procedure. During the recognition test (cf. *procedure*), facilitator objects were used in practice trials, whereas experimental trials only included the four probe objects.

21

For each object in the learning set, two manipulated distracter versions were created. In analogy to Experiment 1 the manipulations either involved a part change or a configural change (Figure 4). Part changes consisted in the substitution of the original part with that taken from another object. More specifically, in case of probe objects, part substitutions involved geons from the facilitator objects to ensure that altered parts had no novelty advantage and had received a similar amount of exposure during the acquisition phase of the experiment. Configural changes were confined to systematic manipulations of size relations between object parts. For each configural distracter, a given size relation was altered into one of the two remaining alternative values (for example, the relation 'smaller' between two parts in the original object would become 'larger' in one distracter and 'equal' in the second). Part and configural manipulations were calibrated across the set of probe objects for equal difficulty [t(23) = .17, p = .97] in adult observers using a similar procedure as described in Experiment 1.

===========================

Insert Figure 4 about here

===========================

Objects and distracters were designed as virtual 3D models using a graphics design software package (POV-Ray version 3.63, Persistence of Vision Raytracer Pty. Ltd.). For each object the rendered 3D model was converted into a grey-level image (resolution 300 x 300 dpi), using a fixed light source and a perspective preserving the visibility of all

22

object components. The object images were shown to the participants at a mean size (height x width) of 15.6 x 10.8 deg of visual angle during the familiarization and learning phase of the experiment. During the recognition test, in each trial the three images (the original and the two distracters) were presented horizontally, each image appearing at a mean size of 10.3 x 7.1 deg of visual angle and spaced at a centre-to-centre distance of 11.5 deg. Viewing distance was 50 cm throughout the experiment. Stimulus presentation and response collection were controlled by an Eprime 1.1 (Psychology Software Tools, Inc.) script running on a laptop computer.

Procedure

The experiment consisted of three parts: familiarization, learning and recognition test. Given the novelty of the objects the first two parts served to train participants to associate each object with a label (represented by the object number) before their object knowledge was assessed, in the final part, in analogy to Experiment 1. The three parts were introduced to children as a series of computer games of increasing difficulty to maintain their interest and motivation.

Familiarization. Participants were first introduced to the objects in a so-called "Add-Me-Up" task to motivate the children. Here in each trial the observer was shown two objects on the computer screen, separated by the symbol +. The task was to respond by typing in the sum of the numbers used to label the objects. The two objects remained on the screen until a response had been made. No memorization was required as this stage, as subjects

were encouraged to use a printed handout showing all six objects and their labels. Feedback was given on each trial about the correctness of the answer.

Learning. Here participants were systematically trained to associate each object with its label, employing a modified version of a previously used supervised-learning paradigm (Rentschler *et al.*, 2004; Jüttner *et al.*, 2006). The training procedure was partitioned into learning cycles, each consisting of a learning phase and a test phase. During the learning phase each object of the current learning set was presented once for 250 msec and in random order, followed by the corresponding object label displayed for 1s. During the test phase, each object of the set was presented twice and assigned to its label by the observer. Upon completion of the test phase, participants received feedback concerning their percent correct value of their responses. The series of learning cycles continued until the observer had reached a criterion of 90% in the recognition test. For the current study, this standard paradigm was modified by using an expanding learning set. The learning started with a set of (randomly chosen) two objects. Once these objects had been learned the learning set was expanded by a third object (randomly chosen from the remaining four) and the subject re-trained to criterion. In this way, the learning set was gradually expanded until all six objects had been included and successfully learned to criterion. The gradual expansion of the learning set from 2 to 6 implied a minimum number of five learning cycles to be performed by each participant.

Recognition test. In the final part of the experiment participants were tested on the previously learned objects using the one-in-three selection task (cf. Experiment 1). In

each trial, three images labelled A, B, C were presented on the screen, one original and two distracter stimuli (both involving either a part or a configural change). The observer had to choose the 'correct' depiction of the object by pressing the appropriately marked button (A, B, C) on the keyboard. The stimulus remained on the screen until the participant had responded. Response time and accuracy were measured as dependent variables. The recognition test was divided into blocks involving either a part or a configural change, and either an upright or an inverted stimulus presentation. Each block was preceded by two practice trials involving the facilitator objects (objects 1 and 2, cf. section *materials*), whereas experimental trials only involved the probe objects (objects 3 - 6) from the learning set. Each object was shown once in each block. The order of the four presentation conditions (part–upright, part–inverted, configuration–upright, configuration–inverted) was counterbalanced across subjects. Participants were instructed not to attempt to rotate their head to see the rotated pictures.

Results

During the learning part of the experiment, children and adults acquired the set of six objects with relative ease. Seven participants (two within the age groups 7 – 8 yrs and 13-14 yrs, one within age groups 9 – 10 yrs, 11 - 12 yrs and 15 – 16 yrs) did not complete the learning procedure and had to be replaced. On average, participants required 5.49 (SD .78) learning cycles to reach the target criterion of 90% correct responses - marginally longer than the minimum of 5 cycles implied by the expanding learning set. There was a

weak trend of young children using more cycles than older ones [F(4,143) = 1.91, p = .09; one-way ANOVA].

Performance in the recognition test was analysed in terms of the accuracy and the latency preceding a correct response. As in Experiment 1 some participants showed particularly long latencies. Five participants with latencies classified as outliers (defined by the group mean ±3 standard deviations) were removed from the data set prior to the statistical analysis (two participants in age group 11 – 12 yrs, and one in each of the age groups 13 – 14 yrs, 15 – 16 yrs and adults).

Accuracy

Means and standard errors of the recognition accuracy for each age group and for each of the two manipulation conditions (Part vs. Configuration) and orientations (Upright vs. Inverted) are shown in Figure 5A. The accuracy data were analysed in a 6 (Age: Adults vs. 15 - 16 yrs vs. 13 – 14 yrs vs. 11 - 12 yrs vs. 9 - 10 yrs vs. 7 - 8 yrs) x 2 (Manipulation: Part vs. Configuration) x 2 (Orientation: Upright vs. Inverted) mixed ANOVA with Age as the between factor. The analysis yielded significant main effects for Age [F(5,133) = 2.93, p < .05], Orientation [F(1,133) = 18.75, p < .001) while Manipulation failed to reach significance (F(1,133) = 2.64, p = .11). Importantly, and unlike Experiment 1, there was no interaction between Age and Manipulation. The only significant interaction was between Manipulation and Orientation [F(1,133) = 6.34, p < .05].

A separate ANOVA for the Upright condition showed no significant main effects for Age [$F(5,133) = 1.16$, $p = .33$] or Manipulation [$F(1,133) = .06$, $p = .81$], or for their interaction [$F(5,133) = 0.30$, $p = .91$]. By contrast, in a similar analysis for the Inverted condition both main effects proved significant [Age: $F(5,133) = 4.00$, $p < .01$; Manipulation: $F(1,133) = 7.62$, $p < .01$] with no significant interaction [$F(5,133) = .55$, $p = .74$]. Inversion negatively affected recognition significantly more in case of configuration than for part changes.

Latency

Response times were analysed for the correct responses of each observer. Figure 5B shows means and standard errors of the latencies for each age group, manipulation condition and orientation. In analogy to the accuracies, the latencies were analysed in a 6 (Age) x 2 (Manipulation) x 2 (Orientation) mixed ANOVA with Age as between factor. The analysis yielded a significant main effect for Age [$F(5,120) = 11.86$, $p < .001$], with older children (15 – 16 yrs) responding faster than the children in the two youngest age groups (ps $< .001$; Tukey HSD test) but not relative to adults. Orientation also proved significant [$F(1,120) = 15.78$, $p < .001$], with latencies to inverted stimuli being longer than to upright ones. All other main effects and interactions were non significant.

Comparison of Experiment 1 and Experiment 2

In order to compare the trajectories observed in Experiments 1 and 2, the accuracies were combined in a joint mixed ANOVA with Experiment as an additional between-subjects factor. The analysis gave significant main effects for Age [$F(5,316) = 16.62$, $p < .001$], Manipulation [$F(1, 316) = 47.55$, $p < .001$] and Orientation [$F(1,316) = 42.08$, $p < .001$], while Experiment was approaching significance [$F(1, 316) = 3.56$, $p = .06$]. There were significant interactions between Manipulation and Experiment [$F(1,316) = 13.47$, $p < .01$], Manipulation and Orientation [$F(1, 316) = 8.81$, $p < .01$] and also for the interaction between Manipulation x Orientation x Experiment [$F(1,316) = 3.87$, $p=.05$]. To consider the three-way interaction in more detail separate ANOVAs for the Part Change and Configural Change conditions were conducted. For part changes, only Orientation and Age were significant ($ps < .01$), and there were no significant interactions ($ps > .37$). For configural changes, Orientation, Age and Experiment produced significant main effects ($ps < .001$) and there was a strong trend in the Orientation x Experiment interaction [$F(1,316) = 3.87$, $p = .05$]. The interaction was further analysed in separate ANOVAS for upright and inverted stimuli. These showed that the factor Experiment only reached significance in the Upright condition [$F(1,316) = 17.76$, $p < .001$] but not in the Inverted condition [$F(1,316) = 3.29$, $p = .08$]. Thus, the critical developmental difference between the two experiments was in the poor recognition of upright metric configural changes (Experiment 1) and the good recognition of upright NAP configural changes (Experiment 2).

Discussion

The principal aim of Experiment 2 was to compare the developmental trajectory for part and configural changes that - unlike Experiment 1- were constrained to the manipulations of non-accidental properties. The most striking finding in Experiment 2 - compared to the data in Experiment 1- is the accurate performance for all age groups with the upright configural change stimuli. We turn therefore to potential explanations of this finding. Given the particular distinctiveness of target stimulus and distracters in the Configural Change condition of Experiment 2 (cf. Figure 4) it is conceivable that performance in the upright condition may have benefited from an old-new advantage mediated by a view-based matching of the target with the (identical) stimulus image seen during the learning phase preceding the recognition test. View-based object representations have been proposed, as an alternative to structural descriptions, in various forms (e.g., Poggio & Edelman, 1990; Tarr & Bülthoff, 1998; Edelman, 1998; Riesenhuber & Poggio, 1999) but their common denominator is the idea of a low-level, non-analytic format where objects are represented as spatial arrangements of local features. However, view-based accounts fail to account for the developmental aspects in our data. Contrary to our findings, they would predict a recognition advantage for configural relative to part changes in Experiment 1 owing to the greater spatial correlation between the features of target and distracters implied by the latter. Nor would they explain the different trajectories for the two manipulation conditions for inverted stimuli. In case of stimulus inversion view-based models rely on mental rotation or similar stimulus alignment mechanisms. Such mechanisms have been shown to become effective relatively early, by

29

the age of five (Marmor, 1975; Estes, 1994); hence, they are unlikely to account for the much longer development documented in our data. Even if older children applied skills like mental rotation to inverted stimuli with different degrees of proficiency there is no evidence for latency costs in our data that would reflect the age-dependent differences in recognition performance between the Part Change and Configural Change condition in both experiments.

A different explanation can be provided in the context of the RBC model, by attributing the contrasting results in Experiments 1 and 2 to different developmental trajectories for configural processing at the levels of metric and non-accidental properties. Within this framework, the data of Experiment 2 indicate that even our youngest children had acquired structural object descriptions involving non-accidental part relations, which enabled them to recognize configural NAP changes in upright stimuli with great accuracy. Other aspects of our data follow RBC predictions. First, stimulus inversion disrupted recognition performance for configural changes more than for part changes because recognizing inverted stimuli should incur inversion costs both at the level of parts and relations for configural changes. Second, the similarity of performance levels for recognizing inverted animals/artefacts on the one hand and inverted geon objects on the other suggests that it is an analytical route involving structural descriptions that is being used for both types of stimuli. RBC generally attributes the costs of configural processing for inverted stimuli to the problem of finding and combining part-relational information that is invariant to this particular transformation. Hence the same mechanisms would be employed when dealing with metric and categorical relational

30

attributes in inverted stimuli, which would explain the statistically equal trajectories in the inverted Configural Change conditions of Experiments 1 and 2.

Experiment 3

The absence of age effects on recognition performance for categorical part-relational changes in Experiment 2 provides supporting evidence for the hypothesis that young children's difficulty with configural changes in Experiment 1 was due to the fact that these changes were confined to the processing of metric relations. However, the complex structure of the naturalistic stimuli in Experiment 1 makes it difficult to assess the part changes as metric or non-accidental, so there is a theoretical possibility that performance for part-changes in Experiment 1 was largely driven by the processing of non-accidental properties. It therefore could be that children's difficulty in Experiment 1 was not strictly confined to configural-metric changes. Rather it might have extended to metric changes in general (including those of individual parts) but was being masked by potentially non-accidental object manipulations in the Part Change condition.

To test for this possibility we conducted a control experiment that employed the same set of stimuli and the same learning procedure as in Experiment 2 but assessed recognition against distracters which now contrasted metric part changes against metric configural changes. If children's difficulty in Experiment 1 was confined to configural metric changes only we expected to observe a similar performance pattern as in Experiment 1,

31

and in particular a significant interaction of the factors Age and Manipulation. However, if children's difficulty in Experiment 1 was confined to metric changes in general (i.e., both at the level of parts and part relations) we predicted no such interaction. The control experiment only involved the oldest and the youngest age group in our study, i.e. 7-8 year olds and adults.

Method

Participants

Two age groups took part in the experiment: 24 adult volunteers (15 females and 9 males; mean age 19 years 4 months) and 32 7- to 8-year-olds (17 females and 15 males; mean age 8 years 1 month). The children were drawn from state schools in Birmingham, UK. The adults were recruited among undergraduate Psychology students at Aston University. They received course credit for participation.

Materials

The same set of learning objects was used as in Experiment 2 (cf. Figure 3).

For each object in the learning set, two manipulated distracter versions were created. The manipulations either involved a metric part change or a metric configural change (Figure 6) of the learning objects. Metric part changes were obtained by changing the aspect ratio of the original part in the distracters. Metric configural changes were confined to systematic manipulations of the relative size between object parts, which – in contrast to Experiment 2 – did not alter their categorical relation. Thus, the relation between two

32

parts in the original object (for example, 'smaller') would continue to apply to the corresponding size-changed parts of the distracter versions. Part and configural manipulations were calibrated across the set of probe objects for equal difficulty [$t(23)$ = -.15, p = .96] in adult observers. Stimulus dimensions and presentation conditions were identical to those in Experiment 2.

============================

Insert Figure 6 about here

============================

Procedure

The experimental procedure was identical to that in Experiment 2.

<u>*Results*</u>

As in Experiment 2, children and adults learned the set of six objects with relative ease. Two participants (both in the age groups 7 – 8 yrs) did not complete the learning procedure and had to be replaced. On average, adults required 5.5 learning cycles to reach the target criterion of 90% correct responses, significantly fewer than the 7.46 cycles required by the 7 to 8-year-olds ($t(54)$ = 5.63, p < .001). However, even for the children mean learning duration was only modestly longer than the minimum of 5 cycles implied by the expanding learning set.

Performance in the recognition test was analysed in terms of the accuracy and the latency preceding a correct response. There were no outliers (defined by the group mean ±3 standard deviations) with regard to latency in Experiment 3.

Accuracy

Means and standard errors of the recognition accuracy for each age group and for each of the two manipulation conditions (Part vs. Configuration) and orientations (Upright vs. Inverted) are shown in Figure 7. The accuracy data were analysed in a 2 (Age: Adults vs. 7 - 8 yrs) x 2 (Manipulation: Part vs. Configuration) x 2 (Orientation: Upright vs. Inverted) mixed ANOVA with Age as the between factor. The analysis yielded significant main effects for Age [$F(1,54) = 21.07$, $p < .001$], Orientation [$F(1,54) = 15.46$, $p < .001$) while Manipulation failed to reach significance ($F(1,54) = 1.40$, $p = .24$). Importantly, there was a significant interaction between Age, Manipulation, and Orientation [$F(1,54) = 4.03$, $p < .05$]. Posthoc comparisons revealed that the interaction was caused by performance differences for part changes and configural changes of upright stimuli in the 7 to 8-year-olds ($t(31) = 2.10$, $p < .05$; paired t-test), all other comparisons were non significant (ps > .58).

============================

Insert Figure 7 about here

============================

Latency

34

Response times were analysed for the correct responses of each observer. Table 2 summarises means and standard errors of the latencies for each age group, manipulation condition and orientation. The latencies were analysed in a 2 (Age) x 2 (Manipulation) x 2 (Orientation) mixed ANOVA with Age as between factor. The analysis yielded no significant main effects for Age [$F(1,41) = .44$, $p = .51$], Orientation [$F(1,41) = .85$, $p = .36$], and Manipulation [$F(1,41) = .33$, $p = .56$], nor any significant interactions.

===========================

Insert Table 2 about here

===========================

Discussion

The substitution of categorical by metric object manipulations in Experiment 3 led to a markedly changed performance pattern in young children compared to Experiment 2. Most importantly, it resulted in a distinct decline in performance for recognizing configural changes but not part changes, demonstrating that metric changes per se could not have accounted for the retarded development of configural object recognition in Experiment 1. The similarity between the data patterns obtained in Experiments 3 and 1 is further illustrated in Figure 7 by re-plotting the endpoints of the trajectories shown in Figure 2 (i.e., the data of the 7 to 8-year-olds and adults) for comparison. Overall, the two data sets show good correspondence despite the very different stimuli employed. Their main difference is the strong inversion effect observed for 7 to 8-year-olds for part

35

changes in Experiment 3. However, one has to remember that in this experiment part manipulations were strictly confined to metric changes, whereas in Experiment 1 they could include non-accidental changes, for which a greater robustness to inversion is to be expected (e.g. Biederman, 1987) - and indeed had been demonstrated in our Experiment 2 even for youngest children tested.

Experiment 3 provides also further evidence for cross-links between developmental dissociations for part-based and configural object recognition on the one hand and the processing of non-accidental and metric attributes on the other. In young children, these became manifest in a generic performance advantage for recognizing upright part changes, regardless whether they occurred at metric (Experiment 3) or non-accidental (Experiment 2) level. By contrast, when recognizing configural changes in objects this equivalence was broken, resulting in relative developmental retardation in the recognition of metric relative to non-accidental part relations. We will return to the implications of these dissociations in the General Discussion.

General Discussion

In three experiments children aged 7-16 and adults were tested in 3-AFC tasks to judge the correct appearance of familiar animals, artifacts and newly learned multi-part objects, which had been manipulated either in terms of individual parts or part relations. The results provide converging evidence for dissociating developmental trajectories of part-

based and part-relational object processing, with a surprisingly late consolidation of the latter when based on metric changes. For the animals and artefacts used in Experiment 1, even the youngest children were close to adult levels for the correct recognition of an individual part change. By contrast, it was not until 11 - 12 years that they achieved similar levels of performance with regard to altered metric part relations. Even for older children aged 11 - 14 years configural processing appeared not fully consolidated and more susceptible to effects of stimulus inversion. In case of the newly-learned multi-part objects, performance for part-specific and part-relational manipulations was equivalent throughout the tested age range for upright presented stimuli, if the manipulations were confined to categorical changes (Experiment 2). If they were restricted to metric changes (Experiment 3) the results replicated the data from Experiment 1, providing further evidence for a particularly late consolidation of configural-metric relative to part-based object recognition.

The present study is a first attempt to systematically trace the developmental trajectories of part-based and part-relational object processing in adolescence from the perspective of a structural theory of object recognition, the RBC model (Biederman, 1987). A key assumption of the model is the notion that objects are described in terms of certain part primitives (geons) that are connected by a restricted set of categorical relations. In the past, the RBC model has inspired considerable developmental work that mainly focussed on the preschool child. Most of this research has been concerned with the role of individual parts, and has demonstrated even in young children and toddlers the particular importance of part information for object categorization and matching (Madole & Cohen,

1995; Smith et al., 1996; Rakison & Butterworth, 1998, Abecassis et al., 2001; Haaf et al., 2003; Mash, 2006). Whether the early primacy of parts in visual processing reflects a peculiar status of geons has been more controversial. Haaf et al. (2003) found that 4-month-old infants were already able to discriminate between geon stimuli that differed on a single non-accidental property, suggesting an early emerging perceptual salience of parts defined by geons. However, Abecassis et al. (2001) demonstrated that the use of geons for the purpose of object recognition and naming may not be present before 4 years. It was only in adults that NAP changes defined boundaries for generalisation in the sense that name extensions were confined to continuous metric variations that did not transgress categorical boundaries. While Abecassis et al. did not trace the performance trajectory in their task beyond the age of 4, the results of our experiments suggest a relatively steep maturation of part-based recognition which permits by the age of 8 the detection of (NAP-defined) geon changes with a reliability close to that of adults. Unlike for parts, developmental studies considering the processing of part relations within the RBC framework have been scarce. In the present study we find a critical difference in development between metric and NAP configural changes.

The retarded development found here for metric configural changes is reminiscent of studies on faces or face-like stimuli. The well-documented, persisting difficulty of children when dealing with faces has been attributed to age-dependent differences in the processing of spatial relations between facial features (Carey & Diamond, 1977; Carey et al., 1980; Mondloch et al., 2002). Two types of relations have been distinguished in the face-recognition literature (cf. e.g., Maurer, Le Grand & Mondloch, 2002; Rhodes, Brake

& Atkinson, 1993; Rhodes, 1993): First-order relations, defining the basic categorical, spatial layout of the key features eyes, nose, and mouth in order to constitute a face; and second-order relations, specifying the precise spacing of those features and conveying facial identity[3]. The dichotomy of first- and second-order relations offers a parallel to the distinction of non-accidental and metric attributes in structural object recognition. In the following, we will therefore relate our findings to those in the face recognition literature.

We first note that the trajectory for configural changes in Experiment 1 (cf. Figure 2), which employed predominantly metric manipulations of the shapes of animals and common artefacts, is remarkably similar to the trajectory reported by Carey et al. (1980) for face recognition. As in our data, Carey et al. report a plateau in performance at an age of 10 - 11. They attribute the lack of improvement (or even slight dip) in performance for children in the age range 10 - 14 to the interference of late developing configural skills and extant featural processing. Analogously, the non-monotonic trajectory observed in our study could indicate the interaction of early maturing part-based and late developing configural (metric) object processing. The emergence of processing skills for second-order relations is accompanied by a particular sensitivity to stimulus inversion. In Experiment 1, inversion disrupted recognition performance more severely for configural than for part changes, but only in children aged 12 or beyond. Inversion effects in the same age range have been observed behaviourally (Carey, 1992; Schwarzer, 2000) and neurophysiologically (Jeffreys, 1996) for face processing.

---

[3] Note that this use of the terms 'first-order' and 'second-order' is unrelated to their use in the vision research literature to distinguish between luminance-based and texture-based stimulus-background segregation, respectively.

Children's difficulties in recognizing faces have been specifically attributed to the fact that face identification requires a categorization at the subordinate level involving a sophisticated processing of second-order spatial relations between facial features (Diamond & Carey, 1986; Freire, Lee, & Symons, 2000; Kemp et al., 1990; Mondloch et al., 2002). The implication would be that such difficulties should be confined to faces only, or at least subordinate categorizations involving particular expertise. By contrast, Experiment 1 demonstrated a protracted development of configural skills for common objects (animals and artefacts) that children could name correctly without the need of expert knowledge. Late developing configural processing therefore appears considerably more pervasive in object recognition and to affect even basic level categories, the most important taxonomic level of human knowledge organisation (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976).

In basic-level object recognition - unlike face identification - maturation effects for configural-metric processing may be obscured by the fact that objects at this level of categorisation typically differ in terms of non-accidental, categorical differences in structural descriptions (Biederman, Subramaniam, Bar, Kalocsai, & Fiser, 1999). Experiment 2 demonstrated that children's ability to deal with such categorical changes develops much earlier and follows the same trajectory for parts and part relations. Thus, emerging configural skills may be masked by simpler mechanisms if the latter are sufficient to solve a given object recognition problem. Likewise, we would suggest that the much earlier maturation of object recognition skills reported in the neuroimaging studies of Golarai et al. (2007) and Scherf et al. (2007) can be attributed to task

40

requirements which do not require a metric analysis of part relations. Golarai et al. employed an old-new recognition task involving objects, faces and scenes. As object stimuli they used photographs of abstract sculptures that distinctly differed from one another in terms of their constituent parts. In Scherf et al.'s study participants freely viewed short movie vignettes, which in case of the 'object' condition included common objects being manipulated by hand (e.g., picking up objects from a desk). Again, the objects were highly discernable in terms of their parts. In neither study had participants to relate perceptual information to precise memory information as in the 'configural change' condition of Experiment 1. Instead, they could base recognition on diagnostic part information alone, a route of processing for which our data (Experiment 2) suggests an early maturation – in line with Golarai et al.'s and Scherf et al.'s results. However, this does not imply that the underlying object representations are already adult like. Our results demonstrate that they are not and we conclude that precise memory representations for objects continue to develop, along a trajectory surprisingly similar to the one established for faces, until deep into adolescence.

Acknowledgements

References

Abecassis, M., Sera, M. D., Yonas, A., & Schwade, J. (2001). What's in a shape? Children represent shape variability differently than adults when naming objects. *Journal of Experimental Child Psychology, 78*, 213-239. doi: 10.1006/jecp.2000. 2573

Aylward, E. H., Park, J. E., Field, K. M., Parsons, A. C., Richards, T. L., Cramer, S. C., & Meltzhoff, A. N. (2005). Brain activation during face perception: evidence of a developmental change. *Journal of Cognitive Neuroscience, 17,* 308-319. doi: 10.1162/0898929053124884

Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review, 94,* 115-147. doi:10.1037//0033-295X.94.2.115

Biederman, I. (2000). Recognizing depth-rotated objects: A review of recent research and theory, *Spatial Vision, 13,* 241-254. doi:10.1163/156856800741063

Biederman, I., Subramaniam, S., Bar, M., Kalocsai, P., & Fiser, J. (1999). Subordinate-level object classification reexamined. *Psychological Research, 62,* 131-153. doi: 10.1007/s004260050047

Bruyer R., & Crispeels, G. (1992). Expertise in person recognition. *Bulletin of the Psychonomic Society, 30,* 501-504.

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of the Royal Society of London B,* 335, 95-103. doi:10.1098/rstb.1992.0012

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of

    faces. *Science, 195,* 312-314. doi:10.1126/science.831281

Carey, S., Diamond, R., & Woods, R. (1980). Development of face recognition: A

    maturational component. *Developmental Psychology, 16,* 257-269. doi:10.1037//

    0012-1649.16.4.257

Cooper, E. E., Biederman, I., & Hummel (1992). Metric invariance in object recognition:

    a review and further evidence. *Canadian Journal of Psychology, 46,* 191-214. doi:

    10.1037/h0084317

Cooper, E. E., & Brooks, B. E. (2004). Qualitative Differences in the Representation of

    Spatial Relations for Different Object Classes. *Journal of Experimental Psychology:*

    Human Perception and Performance, 30, 243-256. doi:10.1037/0096-1523.30.2.243

Davidoff, J., & Roberson, D. (2002). Development of animal recognition: A difference

    between parts and wholes. *Journal of Experimental Child Psychology, 81,* 217-234.

    doi:10.1006/jecp.2002.2659

Davidoff, J., & Warrington, E. K. (1999). The bare bones of object recognition:

    Implications of a case of object recognition impairment. *Neuropsychologia, 37,* 279-

    292. doi:10.1016/S0028-3932(98)00076-1

de Gelder, B., Bachoud-Lévi, A.-C., & Degos, J.-D. (1998). Inversion superiority in

    visual agnosia may be comon to a variety of orientation polarised objects besides

    faces. *Vision Research, 38,* 2855-2861. doi:10.1016/S0042-6989(97)00458-6

de Heering, A., Houthuys, S., & Rossion, B. (2007). Holistic face processing is mature at

    4 years of age: evidence from the composite face effect. *Journal of Experimental*

    Child Psychology, 96, 57-70. doi:10.1016/j.jecp.2006.07.001

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of

    expertise. *Journal of Experimental Psychology: General, 115,* 107-117. doi:

    10.1037//0096-3445.115.2.107

Edelman, S. (1998). Representation is representation of similarities. *Behavioral and*

    Brain Sciences, 21, 449-498. doi:10.1017/S0140525X98001253

Ellis, H. D. (1975). Recognizing faces. *British Journal of Psychology, 66,* 409-426.

    doi:10.1111/j.2044-8295.1975.tb01477.x

Estes, D. (1994). Young children's understanding of the mind: imagery, introspection,

    and some implications. *Journal of Applied Developmental Psychology, 15,* 529-548.

    doi:10.1016/0193-3973(94)90021-3

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the

    encoding of configural information: Direct evidence. *Perception, 29,* 159-170. doi:

    10.1068/p3012

Gathers, A. D., Bhatt, R., Corbly, C. R., Farley, A. B., & Joseph, J. E. (2004).

    Developmental shifts in cortical loci for face and object recognition. *NeuroReport,*

    15, 1549-1553. doi:10.1097/01.wnr.0000133299.84901.86

Golarai, G., Ghahremani, D. G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J. L.,

    Gabrieli, J. D. E., & Grill-Spector, K. (2007). Differential development of high-level

    visual cortex correlates with category-specific recognition memory. *Nature*

    Neuroscience, 10, 512-522. doi:10.1038/nn1865

Haaf, R.A., Fulkerson, A. L., Jablonski, B. J., Hupp, J. M., Shull, S. S., & Pescara-

    Kovach, L. (2003). Object recognition and attention to object components by

preschool children and 4-month-old infants. *Journal of Experimental Child Psychology, 86,* 108-123. doi:10.1016/S0022-0965(03)00108-5

Hummel, J. E. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. *Visual Cognition,* 8, 489-517. doi:10.1080/13506280143000214

Hummel, J. E. & Stankiewicz, B. J. (1996). An architecture for rapid, hierarchical structural description. In T. Inui & J. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 93-121). Cambridge, MA: MIT Press.

Jeffreys, D. A. (1996). Evoked studies of face and object processing. *Visual Cognition, 3,* 1-38. doi:10.1080/713756729

Jüttner, M., Müller A., & Rentschler, I. (2006). A developmental dissociation of view-dependent and view-invariant object recognition in adolescence. *Behavioural Brain Research, 175,* 420-424. doi:10.1016/j.bbr.2006.09.005

Kemp, R., McManus, C., & Pigott, T. (1990). Sensitivity to the displacement of facial features in negative and inverted images. *Perception, 19,* 531-543. doi:10.1068/p190531

Laeng, B., Carlesimo, G. A., Caltagirone, C., Capasso, R., & Miceli, G. (2002). Rigid and nonrigid objects in canonical and noncanonical views: Hemisphere-specific effects on object identification. *Cognitive Neuropsychology, 19,* 697-720. doi:10.1080/02643290244000121

Madole, K. L., & Cohen, L. B. (1995). The role of object parts in infants' attention to

form-function correlations. *Developmental Psychology, 31,* 637-648. doi:10.1037//

0012-1649.31.4.637

Marmor, G. S. (1975). Development of kinetic images: When does the child first

represent movement in mental images? *Cognitive Psychology, 7,* 548-559. doi:

10.1016/0010-0285(75)90022-5

Mash, C. (2006). Multidimensional shape similarity in the development of visual object

classification. *Journal of Experimental Child Psychology, 95,* 128-152. doi:10.1016/

j.jecp.2006.04.002

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural

processing. *Trends in Cognitive Sciences, 6,* 255-260. doi:10.1016/S1364-

6613(02)01903-4

McKone, E., & Boyer, B. L. (2006). Sensitivity of 4-year-olds to featural and second-

order relational changes in face distinctiveness. *Journal of Experimental Child

Psychology, 94,* 134-162. doi:10.1016/j.jecp.2006.01.001

McLaren, I. P. L. (1997). Categorization and perceptual learning: An effect of the face

inversion effect. *Quarterly Journal of Experimental Psychology, 50A,* 257-273. doi:

10.1080/713755705

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing

develops more slowly than featural face processing. *Perception*, *31,* 553-566. doi:

10.1068/p3339

Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science, 14,* 618-622. doi:10.1046/j.0956-7976.2003.psci_1474.x

Poggio, T., & Edelman, S. (1990). A neural network that learns to recognize three-dimensional objects. *Nature, 343,* 263-266. doi:10.1038/343263a0

Rakison, D., & Butterworth, G. E. (1998). Infants' use of object parts in early categorization. *Developmental Psychology, 34,* 49-62. doi:10.1037//0012-1649.34.1.49

Rakison, D. H., & Cohen, L. B. (1999). Infants' use of functional parts in basic-like categorization. *Developmental Science*, *2,* 423-431. doi:10.1111/1467-7687.00086

Rentschler, I., Jüttner, M., Osman, E., Müller, A., & Caelli, T. (2004). Development of configural object recognition. *Behavioural and Brain Research, 104,* 107-111. doi: 10.1016/S0166-4328(03)00194-3

Rhodes, G. (1988). Configural coding, expertise and the right hemisphere advantage for face recognition. *Perception,* 17, 43-63. doi:10.1068/p170043

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47,* 25-57. doi:10.1016/0010-0277(93)90061-Y

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neurosciences*, *2,* 1019-1025. doi:10.1038/14819

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382-439. doi: 10.1016/0010-0285(76)90013-X

Saiki, J., & Hummel, J. E. (1996). Attribute conjunctions and the part configuration advantage in object category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1002-1019. doi:10.1037//0278-7393.22.4.1002

Scherf, K. S., Behrmann, M., Humphreys, K., & Luna, B. (2007). Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Developmental Science, 10,* F15-F30. doi:10.1111/j.1467-7687.2007.00595.x

Schwarzer, G. (2000). Development of face processing: the effect of face inversion. *Child Development, 71,* 391-401. doi:10.1111/1467-8624.00152

Schyns, P. G., & Murphy, G. L. (1994). The ontogeny of part representations in object concepts. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 31, pp. 305-349). San Diego: Academic Press.

Smith, L. B., Jones, S.S., & Landau, B. (1996). Naming in children: a dumb attentional mechanism? *Cognition, 60,* 143-171. doi:10.1016/0010-0277(96)00709-3

Tanaka, J. W., Kay, J. B., Grinell, E., Stansfield, B., & Szechter, L. (1998). Face recognition in young children. When the whole is more than the sum of its parts. *Visual Cognition, 5,* 479-496. doi:10.1080/713756795

Tarr, M. J., & Bülthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Psychological Bulletin & Review, 2,* 55-82. doi:10.1016/S0010-0277(98)00026-2

Thoma, V., Hummel, J. E., & Davidoff, J. (2004). Evidence for holistic representation of ignored images and analytic representation of attended images. *Journal of*

Experimental Psychology: Human Perception and Performance, 30, 257-267. doi:
10.1037/0096-1523.30.2.257

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113,* 169-193. doi:10.1037//0096-3445.113.2.169

Verfaillie, K., & Boutsen, L. (1995). A corpus of 714 full-color images of depth-rotated objects. *Perception & Psychophysics*, *57,* 925-961. doi:10.3758/BF03205454

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141-145. doi:10.1037/h0027474

Figure captions

Figure 1. Examples of stimuli used in Experiment 1. **(A)** Animal stimuli. A one-in-three selection task was used where an item was presented with two distracters that either involved a part change involving the substitution of the correct part with that of another animal (left), or a configural change obtained by altering the animal's proportions (right). Participants had to choose the correct depiction (here: the middle left and bottom right stimulus). **(B)** Artefact stimuli. As with animals, each item was shown with two distracters that either involved a part change (left) or a configural change (right). Again, participants had to choose the correct depiction (here: the bottom left and middle right stimulus). For both animals and artefacts actual stimuli used were coloured versions.

Figure 2. Mean accuracies and latencies in Experiment 1. **(A)** Mean rate of correct identifications within each age group, combined for animals and artefacts, for part- and configurally manipulated stimuli in upright and inverted orientation. **(B)** Mean latencies of correct responses, corresponding to the age groups and conditions shown in (A). Error bars are standard errors.

Figure 3. Geon objects used in Experiment 2. Note that objects 3-6 (Probe objects) consisted of the same parts (geons) and only differed in terms of the categorical relational properties 'relative size' and 'relative location'. Objects 1 and 2 (Facilitator objects) consisted of a different set of geons arranged in a distinctive horizontal or vertical configuration. During the learning phase of the experiment, participants were trained to

associate each of the six objects with its label (number). During the recognition test, facilitator objects were used during practice trials, whereas experimental trials only included the four probe objects. See main text for further details.

Figure 4. Examples of geon stimuli used in the recognition test Experiment 2. As with animals and artefacts in Experiment 1, each object of the learning set was shown with two distracters. They either involved a categorical part change (left) or a categorical configural change in relative size (right). Participants had to choose the correct depiction of the previously learnt object (here: the middle left and top right stimulus).

Figure 5. Mean accuracies and latencies in Experiment 2. **(A)** Mean rate of correct identifications within each age group for part- and configurally manipulated geon stimuli in upright and inverted orientation. **(B)** Mean latencies of correct responses, corresponding to the age groups and conditions shown in (A). Error bars are standard errors.

Figure 6. Examples of geon stimuli used in the recognition test Experiment 3. As with all other experiments, each object of the learning set was shown with two distracters that either involved a metric part change (left) or a metric configural change in relative size (right). Participants had to choose the correct depiction of the previously learnt object (here: the bottom left and top right stimulus).

Figure 7. Mean accuracies in Experiment 3. For both part- and configurally manipulated geon stimuli presented in upright and inverted orientation the mean rates of correct identifications are shown for each age group. For comparison the corresponding values from the same age groups in Experiment 1 are re-plotted from Figure 2A. Error bars are standard errors.

Table 1. Mean accuracies of two subgroups of 7- to 8-year-olds and 9- to 10-year-olds

with non-significant latency differences between Part Change and Configural Change

condition for upright stimuli in Experiment 1.

		Part Change			Configural Change	
Age	Ori.	RT [msec]	Pcorr		RT [msec]	Pcorr
7-8	up	5523	.64		5219	.44
	inv	6486	.53		5866	.44
9-10	up	4387	.67		4122	.50
	inv	5030	.59		4518	.49

Note. Stimulus orientation (Ori.): up (upright) or inv (inverted); RT: mean latency for a

correct response; Pcorr: mean recognition accuracy.

Table 2. Mean latencies of the two subgroups of 7- to 8-year-olds and adults for recognizing metric part and metric configural changes in Experiment 3

		Part Change		Configural Change	
Age	Ori.	RT [msec]	SEM	RT [msec]	SEM
7-8	up	6371	607	5379	371
	inv	5437	403	5711	510
Adults	up	5687	520	6521	386
	inv	5637	468	6361	630

Note. Stimulus orientation (Ori.): up (upright) or inv (inverted); RT: mean latency for a correct response; SEM: standard error of the mean.

Appendix A

List of Animal Stimuli for the Part Change condition and Configural Change condition

Manipulation	Part	Configuration
Practice trials	Gorilla head	Alligator
	Owl ears	Tortoise
Main experiment	Caterpillar head	Bear
	Chicken feet	Chicken
	Rhino body	Cow
	Cow body	Donkey
	Cow head	Fly
	Fish head	Frog
	Fox tail	Giraffe
	Goat legs	Horse
	Horse tail	Kangaroo
	Lion tail	Leopard
	Seal head	Ostrich
	Sheep legs	Penguin
	Swan head	Snail
	Tiger tail	Swan

Appendix B

List of Artefact Stimuli for the Part Change condition and Configural Change condition

Manipulation	Part	Configuration
Practice trials	Bicycle wheels	Saucepan
	Truck wheels	Tricycle
Main experiment	Dustbin lid	Bus
	Fire extinguisher handle	Camera
	House window	Cycle helmet
	Kettle handle	Electric kettle
	Motorcycle wheels	Helicopter
	Pillar box top	Inline skates
	Roller skate wheels	Iron
	Saxophone mouthpiece	Motorcycle
	Taxi window	Pistol
	Teapot lid	Plane
	Tractor wheels	Telephone
	Trumpet buttons	Toaster
	Washing machine drawer	Washing machine
	Wheelie bin wheels	Water tap

Appendix C

Experiment 1 – stimulus calibration

For the animal stimuli, a pilot study with 24 adult participants was run, using the same 3-AFC task as in the main experiment but employing Davidoff & Roberson's (2002) full original stimulus set of 20 part-changed and 20 configurally changed animal pictures. Two subsets of 14 part manipulated and 14 configurally manipulated animals were chosen from this initial set for which the means (Part Change vs. Configural Change) did not differ significantly, with a target accuracy of about .75. The calibration was validated with a further 32 participants (i.e., the adult control subjects in the main experiment) yielding adequately matched recognition accuracies for the two conditions [Part Change: .77; Configural Change: .76; $t(31) = .11$, $p = .91$].

For the artefact stimuli, a set of 24 part manipulated and 24 configurally manipulated stimuli was created and tested with 24 adult participants in a pilot experiment. Again two subsets of 14 part changed and 14 configurally changed stimuli were retained for which the mean accuracies did not significantly differ for the two types of changes, with a target accuracy of about .75. The calibration was validated with the same 32 adult control participants as used for the animal experiment. Again recognition accuracies for the two conditions (Part Change: .77; Configural Change: .74) proved adequately matched [$t(31) = 1.36$, $p = .19$].

(A) Part Change   Configural Change

(B)

Fig. 1

(A)

Part Change

Configural Change

■ upright
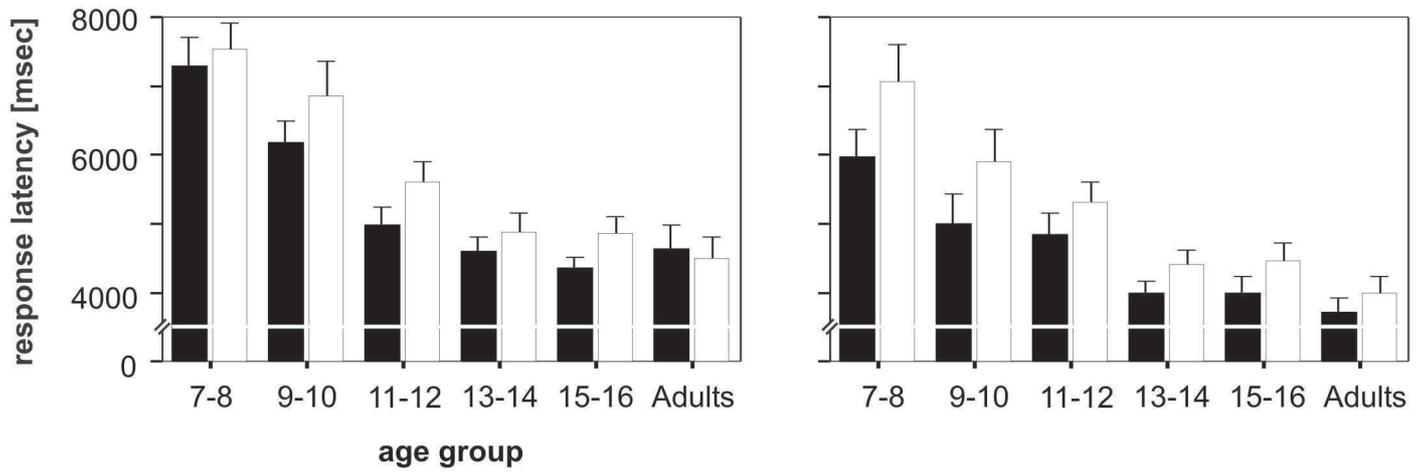□ inverted

correct responses

(B)

response latency [msec]

age group

Fig. 2

Facilitator objects   Probe objects

1   **3**   **5**

2   **4**   **6**
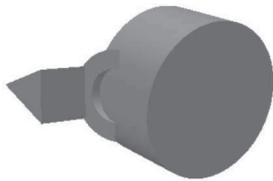
Fig. 3

Part Change   Configural Change

Fig. 4

(A)

Part Change

Configural Change

- ■ upright
- □ inverted

(B)

age group

Fig. 5

Part Change   Configural Change

Fig. 6

Part Change

Configural Change

Exp. 1   Exp. 3
● upright   ■
○ inverted   □

correct responses

1.00
.90
.80
.70
.60
.50
.40
.30

7-8   Adults

age group

Fig. 7

65

(A) Part Change   Configural Change

(B)

Fig. 1

(A)

Part Change

Configural Change

■ upright
□ inverted

correct responses

7-8 9-10 11-12 13-14 15-16 Adults

7-8 9-10 11-12 13-14 15-16 Adults

(B)

response latency [msec]

7-8 9-10 11-12 13-14 15-16 Adults

7-8 9-10 11-12 13-14 15-16 Adults

age group

Fig. 2

Facilitator objects

Probe objects

1

2

3

4

5

6

Fig. 3

Part Change

Configural Change

Fig. 4

(A)

Part Change

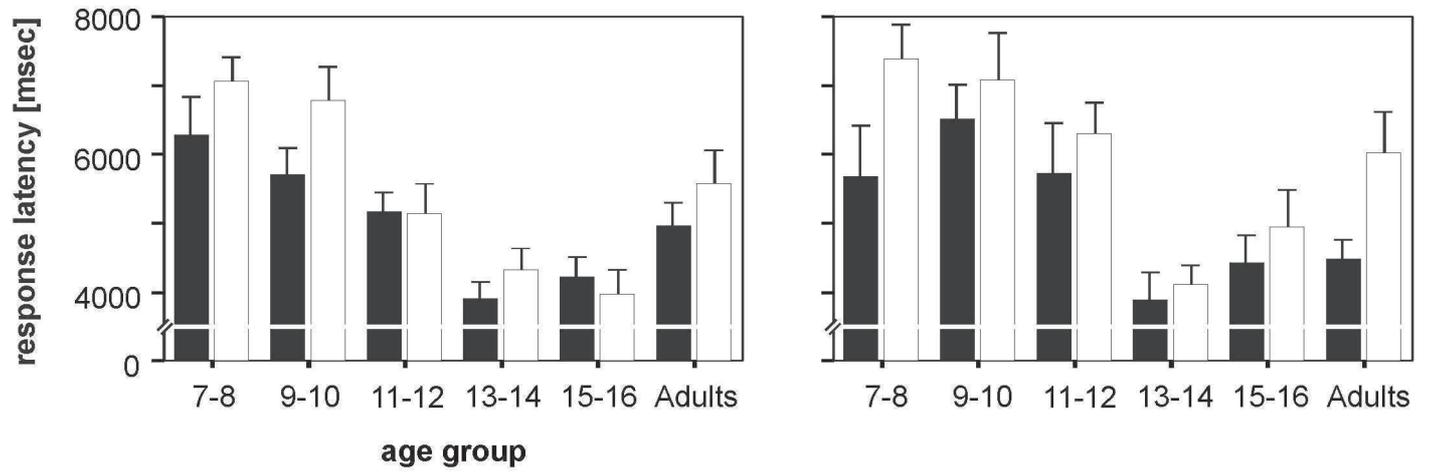Configural Change

■ upright
□ inverted

(B)

age group

Fig. 5

Part Change   Configural Change

Fig. 6

Fig. 7