2006

The configural advantage in object change detection persists across depth rotation

Simone K. Favelle
University of Wollongong, skeane@uow.edu.au

Stephen Palmisano
University of Wollongong, stephenp@uow.edu.au

Darren Burke
Macquarie University

William G. Hayward
University of Hong Kong

The configural advantage in object change detection persists across depth rotation

Abstract

Although traditionally there has been a debate over whether object recognition involves 3-D structural descriptions or 2-D views, most current approaches to object recognition include the representation of object structure in some form. An advantage for the processing of structural or configural information in objects has been recently demonstrated using a change detection task (Keane, Hayward, & Burke, 2003). We report two experiments that extend this finding and show that configural information dominates change detection performance regardless of an object's orientation. Experiment 1 demonstrated the advantage that configural information has over shape and part arrangement information in change detection across four different object rotations in depth. Experiment 2 showed that this advantage occurs for both categorical and coordinate configural changes. These results are consistent with the hypothesis that configural information is a critical feature of object representations and that this information is utilized effectively in object recognition across changes in viewpoint.

Keywords

configural, advantage, object, change, detection, persists, across, depth, rotation

Disciplines

Arts and Humanities | Life Sciences | Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Favelle, S. K., Hayward, W., Burke, D. & Palmisano, S. A. (2006). The configural advantage in object change detection persists across depth rotation. Attention, Perception and Psychophysics, 68 (8), 1254-1263.

The configural advantage in object change detection persists across depth rotation

Simone Favelle[1], William G. Hayward[2], Darren Burke[3], Stephen Palmisano[1].

Short title: The configural advantage across rotation

Correspondence to:

[1] Department of Psychology

University of Wollongong

Northfields Ave, Wollongong 2522

NSW, Australia

Telephone: +61 02 4221 3741

Fax: +61 02 4221 4163

Email: simone_favelle@uow.edu.au

[2] Department of Psychology, University of Hong Kong, Pokfulam Road., Hong Kong

[3] Department of Psychology, Macquarie University, Sydney, Australia

Abstract

While traditionally there has been debate over whether object recognition involves 3D structural descriptions or 2D views, most current approaches to object recognition include the representation of object structure in some form. An advantage for the processing of structural or configural information in objects has been recently demonstrated using a change detection task (Keane, Hayward, & Burke, 2003). We report two experiments that extend this finding and show that configural information dominates change detection performance regardless of an object's orientation. Experiment 1 demonstrated the advantage of configural over shape and part arrangement information in change detection across four different object rotations in depth. Experiment 2 showed that this advantage occurs for both categorical and coordinate configural changes. These results are consistent with the idea that structural or configural information is a critical feature of object representations and that this information is utilised effectively in object recognition across viewpoint change.

INTRODUCTION

As we move around the world and the world moves around us, the distance, position, view, shape, and shading of objects continually change. Despite potentially substantial variations in the retinal images produced by objects, our ability to recognise them is not disrupted. What information in the visual environment is being exploited to allow us to recognise objects as the same or discriminate between different ones? For the most part, early theorizing about visual object recognition and the problem of recognising 3D objects (despite 2D retinal image input) fell into one of two classes of models. One class postulated the existence of viewpoint invariant representations of objects that included 3D information about parts and spatial relations. The other class argued that representations were based on 2D images and were viewpoint dependent. As disparate as these two classes may appear, recent research in this area has begun to investigate the co-operation of these models (Foster & Gilson, 2002).

One particular object property that has been posited as a key factor in the co-operation between viewpoint dependent and viewpoint invariant theories is structure, or spatial configuration (Edelman & Intrator, 2000; Foster & Gilson, 2002; Hummel & Stankiewicz, 1998; Tarr & Bülthoff, 1998). Tarr & Bülthoff (1998) were among the first researchers to suggest that structural information might be encoded along with multiple 2D viewpoint dependent representations of objects. Subsequently, models of object representation and recognition have begun to openly incorporate structure and spatial relations with 2D view information. For example, Edelman and Intrator (2000, 2001, 2003) proposed the Chorus of Fragments (CoF) theory of object representation.

CoF proposes that structure is central to object representation and that the constituent parts of an object are coarsely coded 2D shape fragments, not volumetric primitives. A recent model of object recognition proposed by Foster and Gilson (2002) also effectively combines both structural and view-based information. Foster and Gilson's model is based on empirical data that shows sensitivity to both structural and 2D view information. They found that although differences to object structure were discriminated more readily than metric property differences, they displayed very similar viewpoint dependence. Based on their results, Foster and Gilson argue for a simple additive model of object recognition that includes independent terms for object structure and 2D view information.

Object structure or configuration has been widely investigated in relatively simple objects and has been generally shown to be an important type of information in visual processing. For example, the dominance of configural information over component properties has been demonstrated in the perceptual organisation of 2D objects (see Kimchi, 2003 for a review). An advantage for configural information has also been demonstrated in change detection studies with novel 3D objects (Keane, Hayward, & Burke, 2003; Favelle, Palmisano, Burke, & Hayward, in press). Keane, et al. (2003) investigated change detection in novel objects for: (i) the configuration of object parts, (ii) the shape of those parts, and (iii) the relative arrangement of object parts. Using a one-shot change detection task, Keane et al. found that observers detected changes to the configuration of object parts quicker and more accurately than changes to the shape of the parts or a switching of parts. This result suggests that information regarding the global configuration of parts is better encoded than more local details, such as part shape. Further, categorical configural changes were found to be easier to detect than coordinate (metric based) configural changes. Thus, it is the configural nature of the change and not the size of the change per se that is responsible for this “configural advantage” in which changes to the configuration of object parts are more quickly and more accurately detected than changes to the shape of those parts or a switching of parts.

The abovementioned configural effects were found in conditions where viewpoint was held constant. However, objects are rarely seen from a single viewpoint. Our visual experience typically includes numerous different views of objects in the environment. The aim of the current paper is to investigate the “configural advantage” finding further. In particular, we wish to investigate whether these differences in configuration, switch and shape change detection remain when viewpoint changes and consequently whether this “configural advantage” can be used to add to what we know about object perception and models that combine configural or structural information with 2D view information.

There is evidence that spatial configuration may not be critical for object recognition across changes in viewpoint. Johnston and Hayes (2000) examined recognition for objects that were discriminable based on either morphological (shape) information about parts or spatial configuration of parts across various viewpoint manipulations (stimuli could be rotated either in depth or in the picture plane). In four experiments, using both sequential matching and learning/recognition tasks, performance for objects that were discriminable based on the spatial configuration of parts was poorer than for objects that were discriminable by the shape of their parts, particularly for delayed recognition. Further, they found a larger viewpoint effect for spatial configuration changes than part shape changes; in effect this meant that the differences between the conditions became larger as the objects were rotated from the studied view. Johnston and Hayes argue that the reason for this poor performance is that 2D projections of relative positions of parts are highly viewpoint dependent and do not allow for accurate discrimination between objects under depth rotation.

The shape and configuration change conditions used by Keane et al. (2003) correspond approximately to Johnston and Hayes' (2000) morphological and spatial relations object sets. Both have one condition where part shape is manipulated while overall spatial relations among parts are held constant, and another condition in which part identity or shape remains constant with a manipulation of spatial configuration. However, the results of the two studies are very different. Whereas Keane et al. (2003) found better performance when spatial relations were manipulated, Johnston and Hayes (2000) report better performance for manipulations of object shape.

One obvious difference between the studies is that Johnston and Hayes (2000) included a viewpoint manipulation, whereas Keane et al. (2003) always showed objects from a single viewpoint. Thus, Keane et al.'s finding that spatial configuration changes were detected most easily may be due to the fact that they only examined one viewpoint. A possible integration of the two sets of results would be the prediction that part shape should become more useful to the observer as the viewpoint difference between the two stimuli increased. However, if object representation and recognition is accurately modelled by the combination of structural and view based information, then the configural effects found at zero degrees rotation should persist across viewpoint.

Another difference between the Keane et al. (2003) and Johnston and Hayes (2002) studies is in the experimental tasks used. Johnston and Hayes used both sequential matching and delayed learning tasks to tap in to short- and longer-term object representations, respectively. They found the pattern of performance to be fairly similar between these tasks. The experimental task employed by Keane et al. (and in the current paper) is a one-shot change detection task with a same-different response. In essence, it is a sequential matching task (typically used to investigate short-term object recognition, e.g., Hayward & Tarr, 1997; Johnston & Hayes, 2002; Tarr, Bülthoff, Zabinski & Blanz, 1997), but the analysis pertains primarily to the "different" trials rather than the "same" trials. The rationale behind the use of change detection in this case is that by discovering the types of changes to objects that are easily detected, we can identify object properties that are exploited and represented by the visual system.

The two experiments in this paper investigate whether detection performance for object property changes are differentially affected by changes to an object's orientation in depth. If the configural advantage found previously for object change detection (Keane et al., 2003) generalizes to new viewpoints that would suggest that configural information is readily utilised across the different viewpoints and lend support to theories of object representation and recognition that propose a key role for object structure. In Experiment 1 we investigate the detection of three types of property changes to objects across rotations in depth. In Experiment 2 we performed a more systematic investigation of the use of configural information across viewpoints by using separate changes to categorical and coordinate relations.

EXPERIMENT 1

The object properties investigated in Experiment 1 were the same as in Keane et al. (2003). The changes to be detected were: (i) configural, where a part would move location on the object body resulting in a new part configuration, (ii) shape based, where one part would change shape, or (iii) a switch, where two parts would switch locations (see Figure 1a). Change detection performance was measured across changes in object orientation in depth, from 0 to 45 degrees. If changes to configural information are detected quicker and more accurately than shape or switch changes regardless of an object’s orientation, then this would suggest that object structure is important information which is retained across viewpoint change.

Method

Subjects

A total of 30 undergraduate students participated and were tested individually. Subjects received course credit for participating.

Materials

Stimuli were rendered images of 3D novel objects. Each object was composed of a main body with three appendage parts. The appendages attached to the body at three of six possible positions. Each of the six "standard" objects had a change in part configuration, a change in part shape and a switching of parts made to each of their "limb" parts. Objects were shown rotated by 0, 15, 30 and 45 degrees around the vertical axis giving a total of 240 different object exemplars used in the current experiment (see Figure 1b). The mask used in this experiment was 400 x 300 pixels in size and consisted of elements from a variety of object images. The experiment was controlled by RSVP software (www.tarrlab.org) on Macintosh computers with 17" Macintosh CRT (640 x 480 pixels).

Insert Figure 1 about here

Procedure

The experiment consisted of 360 randomly ordered trials. Each trial began with a fixation cross appearing for 500 ms at the centre of the screen, followed by the first object for 2 s, a mask appearing on the screen for 1500 ms, and finally the second object which remained on the screen until the subject responded. Responses for each trial timed out after 5 s. The next trial began 1 s after the subject made a response or the trial timed out. Each stimulus was jittered by 25 pixels, that is, randomly placed at a position 25 pixels in any direction from the centre of the screen.

Half of the trials were "same" trials and half were "different". The different trials were split equally into the three change conditions. The second object was presented at a viewpoint of 0, 15, 30 or 45 degrees rotated in depth. Further, the second object was either identical to the first or different in one of three ways: (i) part shape; (ii) a switching of parts; or (iii) spatial configuration. Participants were instructed to indicate, regardless of differences in rotation, whether the two objects presented to them were the "same" or "different" by pressing corresponding keys on a keyboard.

Results and Discussion

Accuracy (proportion correct), sensitivity (d') and reaction time (RT) data were collected. The d' data measured the observers' sensitivity to the different types of change independently of their response bias (MacMillan & Creelman, 1991). Each observer's "same-different" responses were converted into hit rates (H) and false alarm rates (F), respectively. These rates, expressed as probabilities ranging between 0.0 and 1.0, were used to calculate d' as follows:

$$d' = 2z[0.5\{1+[2p(c)_{max}-1]^{0.5}\}],$$

where $p(c)_{max} = \Phi\{[z(H) - z(F)]/2\}$ (MacMillan & Creelman, 1991; chapter 6). RT data analysis was conducted using accurate responses. Accuracy, d' and RT data were analysed using a 3x4 repeated measures ANOVA including the within subjects factors of change type (configuration, shape, or switch) and orientation (0, 15, 30, or 45 degrees).

Considering proportion correct (see Figure 2), participants were more accurate at detecting a spatial configuration change (0.89) and less accurate at detecting either a switch (0.81) or a shape change (0.81). A 3x4 repeated measures ANOVA for accuracy showed significant main effects for change type, $F(2,58) = 13.05$, $p < .01$ and orientation, $F(2,58) = 11.07$, $p < .01$, but no interaction between change type and orientation, $F(6,174) = 1.71$, $p = .12$. Bonferroni-adjusted post hoc contrasts show that configuration changes were detected significantly more accurately than either shape or switch changes (both $p < .01$) and that there was no significant difference in accuracy between shape and switch changes ($p = .73$). A trend analysis was conducted for the main effect of orientation on accuracy. The linear and quadratic contrasts were significant, $F(1,29) = 25.38$, $p < .01$ and $F(1,29) = 7.47$, $p < .01$, respectively. The cubic contrast was not significant ($p = .56$). The quadratic trend was interpreted as indicating that there was very little difference in detection performance between the 0 and 15 degree conditions. As seen in Figure 2, the d' data showed the same pattern as the accuracy data. Both the main effects of change type and orientation were significant while their two-way interaction was not significant indicating that our experimental manipulations altered participants' sensitivity to change, rather then their response bias.

Insert Figure 2 about here

The RT data reflected the pattern found in both the accuracy and d' data (see Figure 3). A 3x4 repeated measures ANOVA on RT, also showed significant main effects for change type, $F(2,58) = 11.57$, $p < .01$ and orientation, $F(2,58) = 13.32$, $p < .01$, with no interaction between change type and orientation, $F(6,174) = 1.17$, $p = .32$. Bonferroni-adjusted post hoc contrasts showed that configuration changes were detected significantly quicker than either shape or switch changes (both $p < .01$) and that there was no significant difference in RT between shape and switch changes ($p = .15$). A trend analysis was conducted for the main effect of orientation on RT. The linear contrast was significant, $F(1,29) = 36.10$, $p < .01$. Neither the quadratic nor cubic contrasts were significant (both $p > .06$).

Insert Figure 3 about here

Overall, detection of change to 3D objects was found to be viewpoint dependent (Johnston & Hayes 2000; Tarr, Williams, Hayward, & Gauthier, 1998). However, contrary to Johnston and Hayes' (2000) findings, the current results showed that accuracy and sensitivity were greater and RT faster for detecting configural changes than for either shape or switch changes. There was no significant difference between the shape and switch conditions in accuracy, d' or RT. This suggests that coarse configural information was being utilised quickly and accurately across different object viewpoints in depth, and that observers were more sensitive to changes to this information than to information about the shape of individual components.

Magnitude of Change

When comparing different types of change, a concern arises regarding quantitative differences in the magnitude of the changes. This issue was addressed in a previous paper (Keane et al., 2003) investigating the detection of the same kinds of changes (configural, shape and switch) made to novel objects. In that paper, objects were always viewed at 0^{o} rotation. Keane et al (2003) employed a quantitative measure of change in which the number of pixels that changed from black to white or vice-versa in the two silhouetted objects was calculated (see also Williams & Simons, 2000). This pixel change measure approximates the extent to which both external contours and internal visual features differ between two stimuli. Keane et al. found that this measure could not account for the differences in accuracy between the three types of change, suggesting that it was indeed the type of change, rather than the quantitative nature of the change that was responsible for their findings.

The current accuracy data were subjected to a similar analysis. We employed two different measures of quantitative change (see Appendix A for the mean values for each measure in each condition). The first was the pixel change measure used previously by Keane et al (2003) and Williams and Simons (2000). The second was a measure of change in the 8-bit (256 colour) images[1]. This method gives a quantitative measure of colour change; while it is somewhat limited (since similarities in the colour palette table may not equate with psychological reality), it allows us to differentiate between subtle and larger changes in colour[2]. Two separate ANCOVAs were conducted; both with proportion correct as the dependent value and either pixel change or change in colour image as the covariate. The results of both ANCOVAs showed that neither the pixel change nor the change in colour image covariate significantly predicted accuracy ($F(1,3) = 0.16$, $p = .72$, and $F(1,3) = 0.07$, $p = .81$, respectively). Thus, while there are physical differences between configuration, switch and shape changes, these are clearly not large enough to account for the configural advantage found here and in previous studies.

EXPERIMENT 2

In Experiment 1, changes to part configuration were consistently detected quicker and more accurately than changes made to part shape and switch changes, even across object rotations in depth. Thus, information about the spatial relations between components seems to be encoded in a fashion that can be deployed not just for images depicting the object from the same viewpoint (Keane et al, 2003), but also across viewpoints. If this is indeed the case, in what form is this information encoded? Different models make distinct predictions. Structural description theories of object recognition propose the representation of object parts in terms of categorical spatial relations (e.g., one part is above or below another; Biederman, 1987; Hummel &

[1] These are in terms of the colour values, so the maximum difference would be obtained from a purely black image changing to a purely white image (this would score 100%). Any other change would score less than 100%.
[2] Given that all components had the same physical colour (purple), this measure will reflect increasing differences in colour caused by, for example, increasing rotations of an object that systematically change surface angle with respect to lighting and therefore make a surface seem brighter or darker.

Stankiewicz, 1996). Alternatively, viewpoint dependent theories rely on coordinate spatial relations (e.g., the metric direction and distance between two points; Bülthoff, Edelman, & Tarr, 1995).

With viewpoint held constant, Keane et al. (2003) found that change detection for categorical relations was more accurate than for coordinate relations (although the RTs were not different), in accordance with structural description theories of object recognition. Given this result, it was expected that categorical spatial relations would also be more useful when the studied image had to be generalized across viewpoint. In Experiment 2, we tested the ability to detect changes to categorical and coordinate spatial relations across different object rotations in depth. Importantly, as in Keane et al. (2003), the categorical and coordinate spatial relations changes investigated in the current experiment were created such that it was only the relations between parts that distinguished the two changes - the quantitative difference between categorical and coordinate changes was identical. We expected to replicate the pattern of results found by Keane et al (2003) at zero degrees rotation; however, the main concern of this experiment was whether this pattern would remain across changes in viewpoint. In particular, we were interested in (i) whether we would replicate the general configural detection advantage over switch and shape detection, and (ii) whether categorical changes to configurations would be detected more easily than coordinate changes.

Method

Subjects

A total of 29 undergraduate students participated and were tested individually. Subjects received course credit for participating.

Materials

The stimuli used in this experiment were a subset of those used by Keane et al. (2003, Experiment 4; see Figure 3 for an example). The stimuli were rendered images of three-dimensional novel objects similar to those used in Experiment 1. Each object was composed of a main body with three appendage parts. The appendages or "limbs" were attached to the body at three of nine possible positions. There were six "base" objects, each having categorical, coordinate or shape changes made to them. For any particular object the same part was involved in all the experimental conditions. The two other appendage parts of the object were also involved in configuration and shape changes; however, these “other” changes were included so that subjects would not focus attention solely on the part involved in the categorical or metric change (which would perhaps bias results). The distance an appendage part moved was quantitatively identical for both categorical and coordinate changes, but involved different changes in relation to the two other appendage parts. The categorical change involved one part moving along the body such that it went from above a second appendage to below (or vice versa) and in line with the third appendage. The coordinate change involved one part moving along the body such that it was further above or below the two other appendages (see Figure 3).

Insert Figure 3 about here

Objects were shown at four different viewpoints in depth. Each object was rotated around the vertical axis 0º, 15º, 30º and 45º giving a total of 288 different object exemplars used in the current experiment. All objects were photorealistically rendered with the same colour and texture. The entire background screen was white. The mask used in this experiment was 425 by 312 pixels in area. The experiment was controlled by RSVP software (www.tarrlab.org) on Macintosh computers with 17" Macintosh CRT (640 x 480 pixels).

Procedure

The experiment consisted of 528 randomly ordered trials. The same procedure was used as for Experiment 1.

Results and Discussion

Three separate 3x4 repeated measures ANOVAs including the within subjects factors of change type (categorical configuration, coordinate configuration, shape) and orientation (0, 15, 30, or 45 degrees) were conducted on accuracy (proportion correct), sensitivity (d') and reaction time (RT). RT data analysis was conducted using accurate responses.

The 3x4 ANOVA for accuracy data showed significant main effects of change type, $F(2,56) = 16.45$, $p < .01$, and orientation, $F(3,84) = 7.93$, $p < .01$ but no significant interaction between factors, $F(6,168) = 1.21$, $p = .31$. Figure 5 suggests a decrease in accuracy as orientation change increased. A trend analysis conducted for the main effect of orientation on accuracy revealed that the linear and quadratic trends were significant (both $p < .05$). This pattern was likely to be due to the greater accuracy at 15 degrees rotation relative to both zero and thirty degrees rotation.

Even though the interaction between orientation and change type was not significant, Figure 4 shows that the trend at zero degrees is similar to Keane et al's (2003) results. Indeed, a planned comparison between categorical and coordinate changes at zero degrees (mean proportion correct 0.90 and 0.83, respectively) showed a significant difference ($p < .05$). The same comparison at 45 degrees (mean proportion correct 0.85 and 0.82, respectively) was not significant ($p = .38$). Bonferroni-adjusted post hoc contrasts showed that when collapsed across orientation, accuracy for shape changes was significantly worse than any other type of change (all $p < .001$, see Figure 5). Although overall detection accuracy for categorical relation changes was greater than for coordinate relation changes (mean proportion correct 0.89 and 0.85, respectively), this difference was not significant ($p = .19$). As seen in Figure 5, the d' data showed the same pattern as the accuracy data. Both the main effects were significant, while the interaction was not significant. In addition, a Bonferroni-adjusted post hoc contrast showed that overall sensitivity was not significantly different for categorical and coordinate changes ($p = .22$). These d' results indicate that our experimental manipulations altered participants' sensitivity to change, rather then their response bias.

Insert Figure 5 about here

A 3x4 repeated measures ANOVA for RT, also showed significant main effects for change type, $F(2,56) = 8.08$, $p < .01$ and orientation, $F(3,84) = 5.86$, $p < .01$, with no significant interaction between change type and orientation, $F(6,168) = 0.88$, $p = .51$ (see Figure 6). Bonferroni-adjusted post hoc contrasts showed that there was no difference in RT between the types of configural changes ($p = .83$) and that both types of configural change were detected significantly quicker than shape changes (both $p < .01$). A trend analysis was conducted for the main effect of orientation on RT. The linear contrast was significant ($p < .05$) indicating that RT increased as the degree of rotation increased. The quadratic contrast was not significant ($p = .07$).

Insert Figure 6 about here

In Experiment 2, shape changes were detected significantly slower and less accurately than other change types, regardless of object orientation. Categorical and coordinate configural changes were created using quantitatively equal sized shifts of parts (presumably creating equal sized changes in pixel change, contour, colour, luminance, etc), yet when viewpoint is held constant there is clearly a difference in subjects' ability to detect categorical and coordinate changes. However, there is no interaction between change type and orientation; across all orientations the speed and accuracy of both categorical and coordinate change detection remained similar. While changes to categorical spatial relations appear to be more accurately detected when the viewpoint remained constant, there was little evidence to support the contention of a general difference between categorical and coordinate spatial relations across object rotations in this experiment. Thus, the current data indicates that configural information in terms of *both* categorical and coordinate relations provides useful information for object discrimination across changes in viewpoint, an idea compatible with such models as Hummel and Stankiewicz's (1998) *MetriCat*, which encodes both categorical and coordinate spatial relations.

GENERAL DISCUSSION

Across all manipulations of object stimuli, detection of configural changes remained superior to that of shape and switch changes. In contrast to Johnston and Hayes (2000), both experiments showed that regardless of orientation in depth, configural changes to objects were detected more quickly and more accurately than part shape or switch changes. In attempting to understand the differences in results between these two studies, three task differences might be important. First, the Johnston and Hayes study used object sets that were constructed of a different number of parts. The set matched on spatial relations consisted of objects with three or four parts whereas objects matched on morphology each had five parts. This resulted in different basic configurations of the objects in each set, making comparison of the two difficult (indeed, Johnston and Hayes never made a direct comparison of recognition performance for the two sets of objects). The increased complexity (number of parts) of the objects discriminable by spatial relations may at least partially account for the relatively poorer performance Johnston and Hayes found for this condition. Second, the range of viewpoints used in the Johnston and Hayes (2000) study was much larger (up to 180°) than the range of the current study (up to 45°).

Third, and perhaps most important, changes to objects in the Johnston and Hayes (2000) study involved all components (i.e., morphology changes altered the shapes of all parts, and changes to spatial relations altered the connectivity of all components). Thus, participants in their study performed trials with a greater degree of change in object morphology, configuration and rotation size. Further, since in the morphology change experiments successful matching and recognition could be done based on just one part (any part), it may not be surprising that performance was better than for spatial relation change experiments in which the relations between all parts had to be encoded for successful matching and recognition. In contrast, the current study used object stimuli that consist of the same number of parts (each object has four parts) and conditions in which changes to part morphology (shape) and spatial relations each involved only one part. Using stimuli of this kind allowed us to more directly compare spatial relations between parts with the part shape. Further, it required participants to encode objects as a whole at study, since it was not known which part would change or how. Perhaps a more holistic or global encoding of objects leads to changes of a more global nature (such as configuration changes) being easier to detect than local changes (such as the shape of one part). Such a hypothesis needs to be investigated by future studies.

Experiment 2 demonstrated that the advantage found for changes to spatial relations across object rotation did not depend upon whether the configural change was categorical or coordinate in nature. Changes of similar magnitude to the configuration of an object appeared to be detected at a similar speed and accuracy. Obviously, the finding of a null result makes it difficult to draw firm conclusions; however, it is interesting to compare these results with those of Hummel and Stankiewicz (1996). In their study, they compared recognition at a single viewpoint of objects that differed in categorical or coordinate spatial relations. Similar to the current results and those of

Keane et al. (2003), when viewpoint was fixed they found less confusion between objects sharing categorical relations than coordinate relations. Of interest is whether such differences would remain if they had examined performance across changes in viewpoints. The results of Experiment 2 suggest that this is an important condition to be considered.

The data from the present study suggest a key role for coarse configural information (in terms of both categorical and coordinate relations) in the perception of novel objects across rotation in depth (at least up to rotations of 45^{o}), and a reduced role for the shape of specific components. On the whole, these results provide support for the general idea that structural information may be encoded along with multiple viewpoint specific representations of objects (Tarr & Bülthoff, 1998). More specifically, the current findings together with those of Keane et al. (2003) are compatible Edelman and Intrator's (2000, 2001, 2003) CoF theory of object representation in which object structure is the principal element of object representation and the parts of an object are coarsely coded shape fragments. Indeed, we found that configural changes, and not shape changes, were detected accurately and quickly regardless of the object's orientation in depth. Further, viewpoint dependent performance was found across all change types, with configural changes (both coordinate and categorical) being consistently detected most quickly and accurately.

Important to note, however, is that although coarse configural information is used quickly and accurately across different object viewpoints, it appears subject to the same decay across changes in viewpoint as other shape information. That is, configural information appears to be qualitatively similar to other kinds of shape information; it simply gives a stronger quantitative signal at all viewpoints. This is compatible with Foster and Gilson’s (2002) of a model of object recognition that combines both structural and view-based information in a simple additive model that includes independent terms for object structure and 2D view information. This model accounts for the current results in that it is sensitive to structural features, but generally viewpoint dependent.

References

Biederman, I. (1987). Recognition-by-Components: A theory of human vision understanding. *Psychological Review*, 94, 115-145.

Bülthoff, H.H., Edelman, S.Y., & Tarr, M.J. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex*, 3, 247-260.

Edelman, S., & Intrator, N. (2000). (Coarse coding of shape fragments) + (retinotopy) ≈ representation of structure. *Spatial Vision*, 13, 255-264.

Edelman, S., & Intrator, N. (2001). A productive, systematic framework for the representation of visual structure. In T. Leen (Ed.) *Proceedings of the 2000 Conference on Neural Information Processing*, 10-16, MIT Press.

Edelman, S., & Intrator, N. (2003). Towards structural systematicity in distributed, statically bound visual representations. *Cognitive Science*, 27, 73-110.

Favelle, S.K., Palmisano, S., Burke, D., & Hayward, W.G. (in press). The Role of Attention in Processing Configural and Shape Information in 3D Novel Objects, *Visual Cognition*.

Foster, D.H., & Gilson, S.J. (2002). Recognizing novel three-dimensional objects by summing signals from parts and views. *Proceedings of the Royal Society of London B: Biological Sciences*, 269, 1939-1947.

Hayward, W.G., & Tarr, M.J. (1997). Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1511-1521.

Hummel, J.E., & Stankiewicz, B.J. (1996). Categorical relations in shape perception. *Spatial Vision*, 10, 201-236.

Hummel, J. E., & Stankiewicz, B. J. (1998). Two roles for attention in shape perception: A structural description model of visual scrutiny. *Visual Cognition,* 5, 49-79.

Johnston, M.B., & Hayes, A. (2000). An experimental comparison of viewpoint-specific and viewpoint-independent models of object representation. *The Quarterly Journal of Experimental Psychology*, 53A, 792-824.

Kimchi, R. (2003). Relative dominance of holistic and component properties in perceptual organization of visual objects. In M. Peterson, & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes.* Oxford University Press.

Keane, S.K., Hayward, W.G., Burke, D. (2003). Detection of three types of changes to novel objects. *Visual Cognition*, 10, 101-127.

MacMillan, N.A., & Creelman, C.D. (1991). *Detection theory: a user’s guide.* New York, NY: Cambridge University Press.

Tarr, M.J., Bülthoff, H.H., Zabinski, M., & Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint? *Psychological Science, 8*, 282-289.

Tarr, M.J., & Bülthoff, H.H. (1998). Image-based object recognition in man, monkey, and machine. *Cognition: Special issue on "Image-based Recognition in Man, Monkey, & Machine"*, 67, 1-20.

Tarr, M.J., Williams, P., Hayward, W.G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience*, 1, 275-277.

Acknowledgements

Thanks to Alex Lee for help with the analyses. This research was partially supported by a grant from the Hong Kong Research Grants Council (CUHK4232/02H).

Tables

Appendix A. Mean values for (a) percentage pixel change from “original” object and (b) percentage colour change from “original” object in Experiment 1 as a function of change type (including *same*) and object orientation. The standard error of the mean is given in parentheses.

Figures

Figure 1. Example of (a) the three different types of change (configuration, shape and switch) used in Experiment 1 and (b) one of the object stimuli rotated across 45^o.

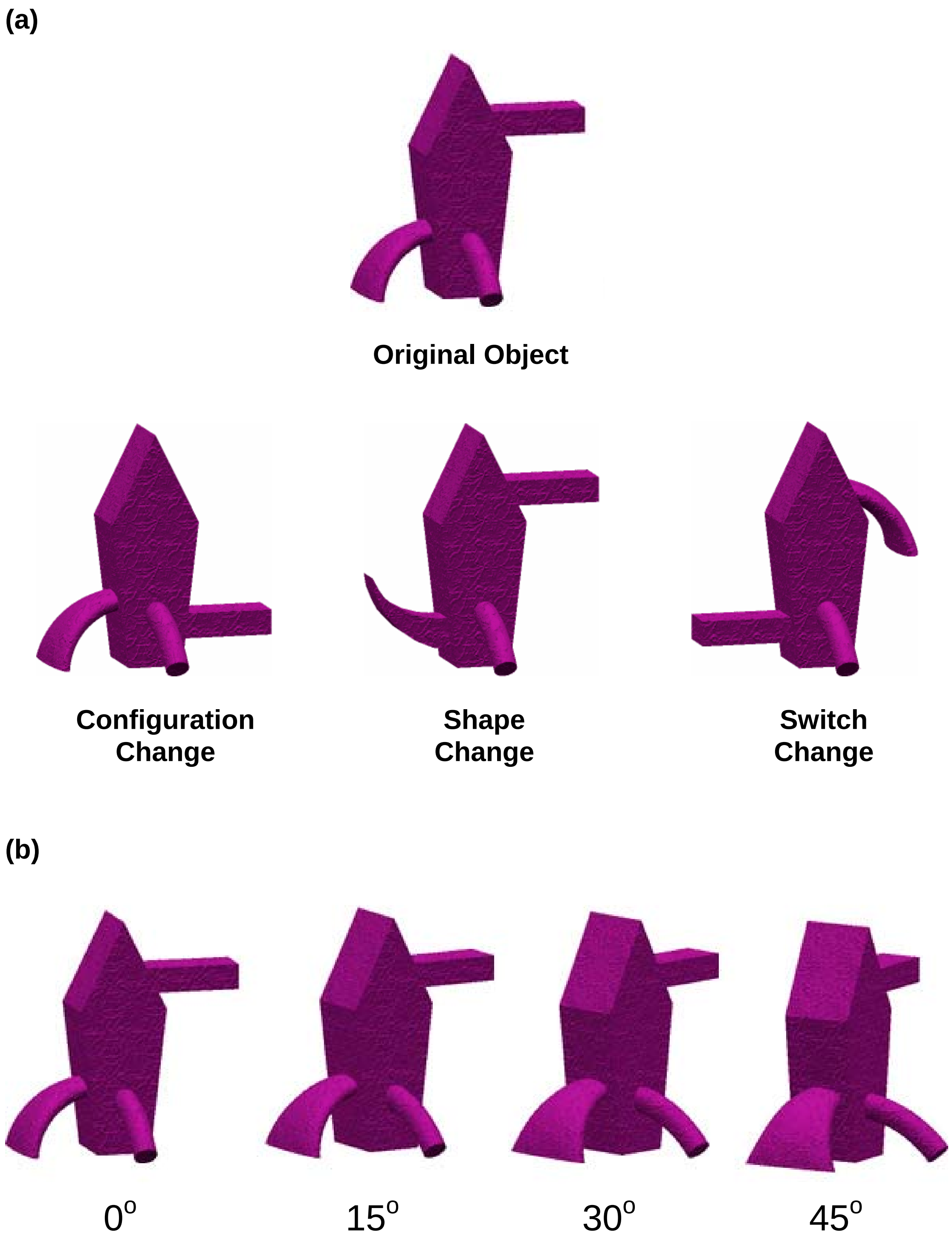
Figure 2. Mean accuracy (top panel) and sensitivity to change (bottom panel) on the change detection task in Experiment 1 as a function of change type and object orientation. Error bars represent standard errors of the means.

Figure 3. Mean RT on the change detection task in Experiment 1 as a function of change type and object orientation. Error bars represent standard errors of the means.

Figure 4. Example of the three different types of change (categorical, coordinate, and shape) used in Experiment 2. The categorical change in this example involves the left part changing its categorical relationship with (i) the middle part from "below" to "in line with" and (ii) the right part from "below" to "above". The coordinate change in this example involves the left part changing its coordinate relationship with both the middle and right parts by 100 pixels distance.

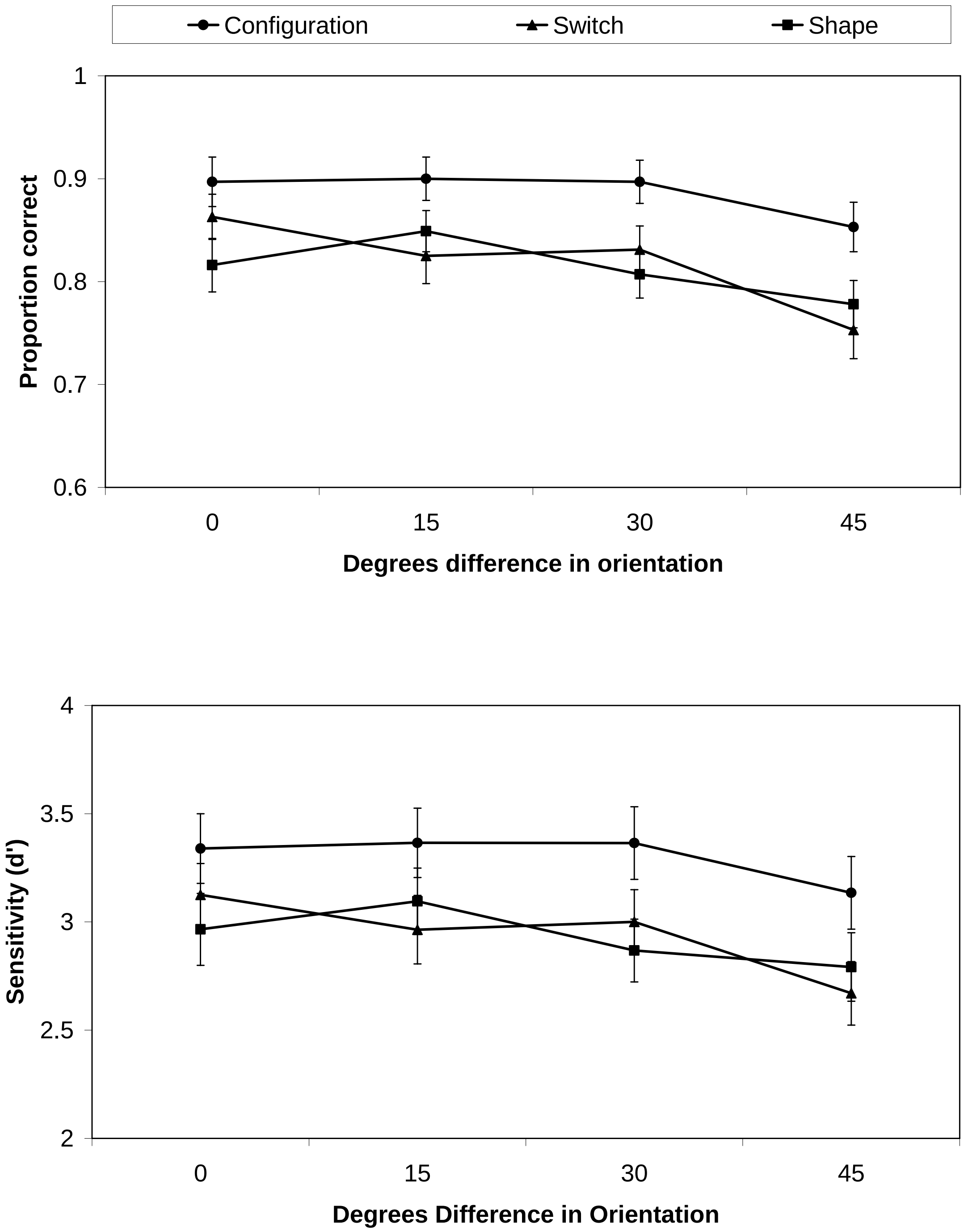
Figure 5. Mean accuracy (top panel) and sensitivity to change (bottom panel) on the change detection task in Experiment 2 as a function of change type and object orientation. Error bars represent standard errors of the means.

Figure 6. Mean RT on the change detection task in Experiment 2 as a function of change type and object orientation. Error bars represent standard errors of the means.



(a)
Original Object
Configuration Change
Shape Change
Switch Change
(b)
0°
15°
30°
45°

Figure 1



Configuration
Switch
Shape
1
0.9
0.8
0.7
0.6
Proportion correct
0
15
30
45
Degrees difference in orientation
4
3.5
3
2.5
2
Sensitivity (d')
0
15
30
45
Degrees Difference in Orientation

Figure 2

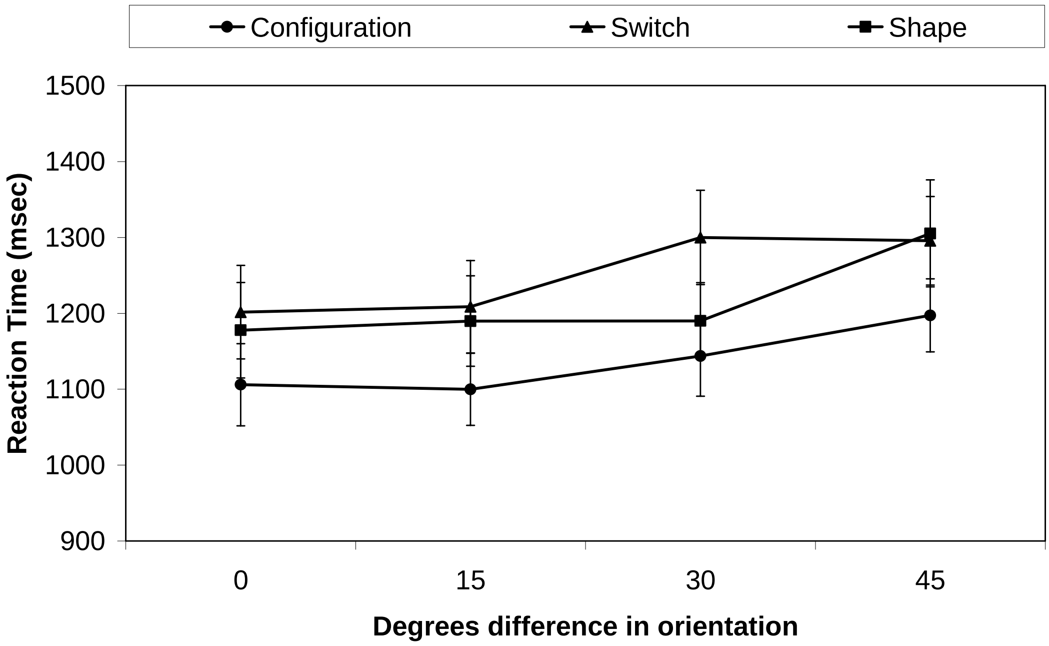
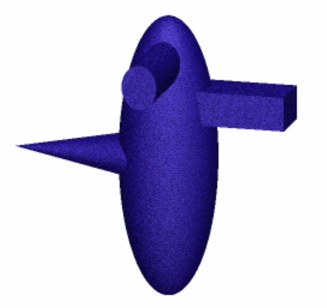
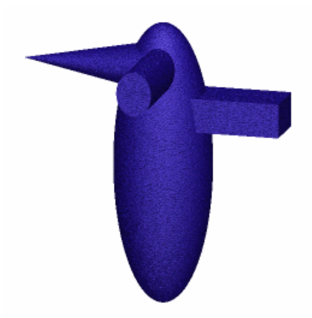
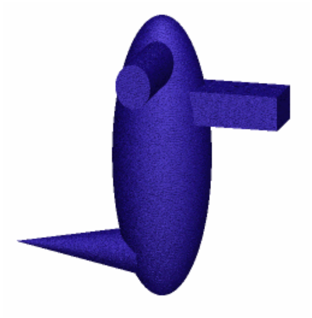

Configuration
Switch
Shape
1500
1400
1300
1200
1100
1000
900
Reaction Time (msec)
0
15
30
45
Degrees difference in orientation

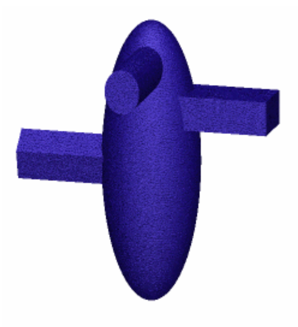
Figure 3



Original Object







Categorical **Coordinate** **Shape**

Figure 4

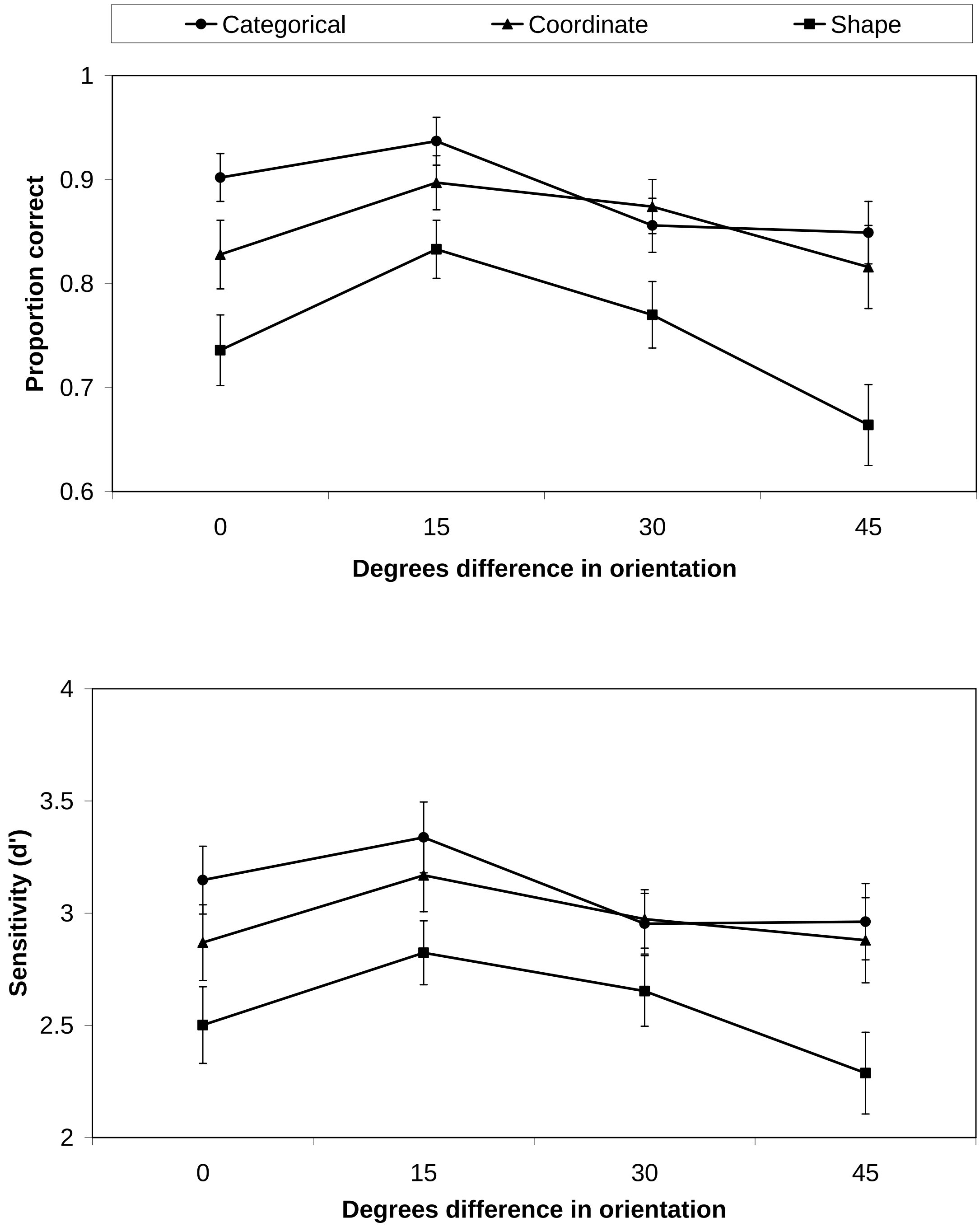
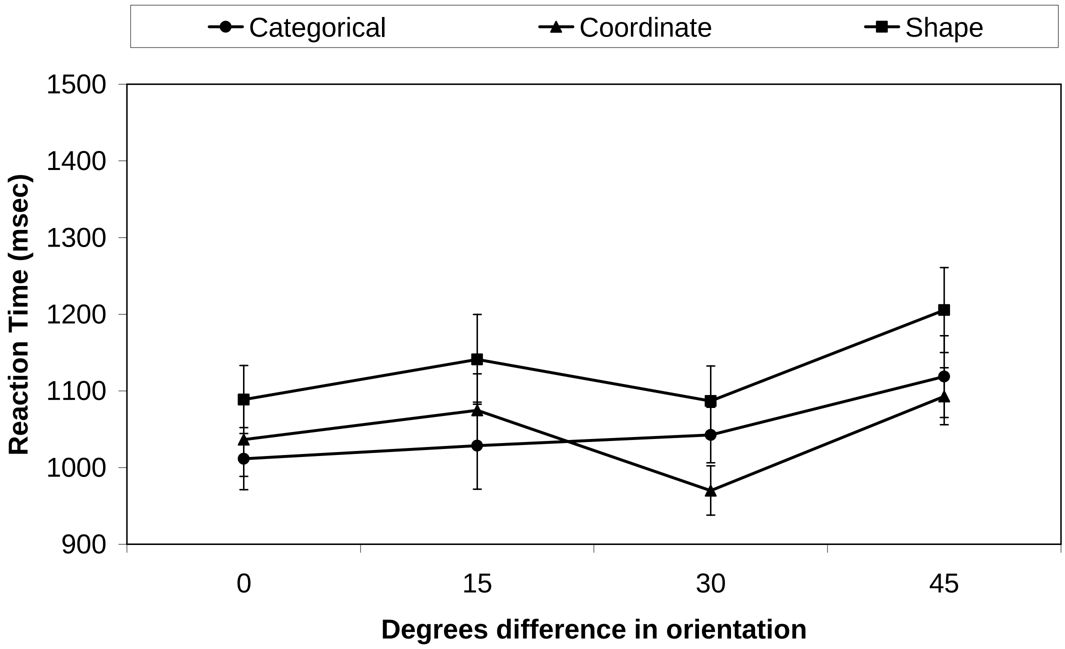

Categorical
Coordinate
Shape
1
0.9
0.8
0.7
0.6
Proportion correct
0
15
30
45
Degrees difference in orientation
4
3.5
3
2.5
2
Sensitivity (d')
0
15
30
45
Degrees difference in orientation

Figure 5



Categorical
Coordinate
Shape
1500
1400
1300
1200
1100
1000
900
Reaction Time (msec)
0
15
30
45
Degrees difference in orientation

Figure 6

(a)

Degrees Difference in Orientation	*Same*	*Configuration*	*Switch*	*Shape*
0	0 (0)	3.66 (0.22)	2.90 (0.54)	1.40 (0.19)
15	3.33 (0.60)	5.78 (0.60)	4.82 (0.87)	4.07 (0.64)
30	6.12 (0.78)	7.45 (0.83)	6.59 (1.04)	6.41 (0.88)
45	7.76 (0.93)	8.11 (0.94)	7.83 (1.11)	7.81 (1.01)

(b)

Degrees Difference in Orientation	*Same*	*Configuration*	*Switch*	*Shape*
0	0 (0)	3.01 (0.24)	2.83 (0.52)	1.64 (0.30)
15	3.35 (0.50)	4.98 (0.48)	4.43 (0.68)	3.89 (0.54)
30	5.28 (0.60)	6.14 (0.63)	5.64 (0.77)	5.52 (0.69)
45	6.33 (0.74)	6.58 (0.74)	6.43 (0.82)	6.43 (0.80)

Appendix A