Perception, 2011, volume 40, pages 1047 – 1062

doi:10.1068/p7084

Configural information is processed differently in human action

Jeff Loucks
Institute for Learning & Brain Sciences (I-LABS), University of Washington, Seattle, WA 98195, USA;
e-mail: jloucks@uw.edu
Received 25 July 2011, in revised form 6 September 2011

Abstract. Recent evidence indicates that observers' sensitivity to configural information in dynamic human action is disrupted when action is inverted, whereas sensitivity to featural action information is not. The current research involved two experiments that expand upon this basic finding. Experiment 1 revealed that featural and configural action information are processed similarly in static representations of action as in dynamic action. Experiment 2 indicated that configural processing is uniquely sensitive to orientation only in human action as compared to a similar control stimulus. These findings further support the idea that the perception of action recruits specialized orientation-specific configural processing, and parallel similar findings in face perception and visual expertise.

1 Introduction

Effective social cognition is grounded in processing countless instances of human action every day. Higher-level aspects of social cognition—for example, understanding mental states—are informed by the perception of action itself. Our ability to efficiently process the actions of others belies the complex nature of human action as a stimulus: actions are often carried out rapidly, without pause, and are only available to perception for a brief window of time. In light of this complexity, the apparent ease with which actions are processed by observers suggests that relatively sophisticated cognitive mechanisms are enlisted in the service of action perception (Baldwin 2005).

One aspect of action processing that has only recently come under investigation is action discrimination—that is, identifying differences among actions. Ultimately, discriminating action is key to everyday goal inference (eg discriminating lifting versus pushing an object) and also survival (eg discriminating harmful from helpful actions). In addition, the study of action discrimination can be used to uncover the basic perceptual and attentional processes involved in action perception.

In one of the first studies to examine action discrimination, Loucks and Baldwin (2009) adapted methodologies from studies of face perception. The face perception literature typically distinguishes two main sources of perceptual information that people utilize for identifying and discriminating faces: featural and configural information (Maurer et al 2002). Featural information is captured in the relatively local detail of facial elements, such as the eyes, nose, and mouth. In contrast, configural information is captured in the relatively more global spatial relations of these features within the face. These sources are thought to be distinct from one another, as they register in different spatial frequencies (Goffaux et al 2005), and because processing of configural information is more heavily impaired when faces are inverted (Freire et al 2000; Mondloch et al 2002; Murray et al 2000). Some have argued that this greater effect of inversion for configural information reflects the fact that faces are processed holistically (Rossion 2008; van Belle et al 2010), and evidence for holistic face processing is found in the composite face illusion (Young et al 1987).

Loucks and Baldwin (2009) explored whether people might also make use of similar perceptual sources of information when processing dynamic human action. They defined featural action information as relatively local detail concerning fine body motion, such as a particular type of grasp used on an object. Configural action information was defined as relatively global spatial/relational information, such as the trajectory used to move an object. Across two experiments, participants' sensitivity to detecting changes in these two sources of information was measured. The first experiment demonstrated that both featural and configural information are useful for discriminating action, in that participants were able to detect changes in both sources when action was upright. However, only processing of configural information was selectively impaired with inversion; that is, participants were significantly less sensitive to configural changes for inverted action, while sensitivity to featural changes was unimpaired. The second experiment verified that these two sources of action information rely on relatively distinct bands of spatial frequency for processing. In addition, across both experiments, featural action information was selectively attended to over configural action information, even when configural changes were objectively larger than featural changes as measured by the overall amount of pixel change across time.

Interestingly, this last finding is also a property of featural and configural information in face processing (eg Freire et al 2000; Mondloch et al 2002). Thus, processing of action parallels processing of faces: for both stimuli, sensitivity to low-spatial-frequency featural information is selectively attended to and relatively insensitive to inversion. Loucks and Baldwin speculated that these parallels suggest action processing may rely on shared mechanisms that subserve both face processing and expert visual processing.

The present study aimed to further investigate the nature of featural and configural processing in action perception. In particular, the current research explored the degree to which the featural/configural distinction is a function of the nature of the stimulus. In other words, are featural and configural sources of information processed differently in human action in comparison to some other class of stimuli?

The face-perception literature offers a model for examining this question within action perception. The face has long been viewed as a unique visual stimulus. Yin's seminal research on the face-inversion effect demonstrated effects of inversion on recognition that were restricted to faces and did not occur for other objects (Yin 1969). Tanaka and Farah (1993) also found a part – whole recognition effect in face perception: a particular feature is recognized more accurately when presented within a face than when presented in isolation. This effect was not found in recognizing the particular elements of houses, however, suggesting faces are processed more holistically than houses. Neuropsychological findings also support the view that faces are processed with distinct neural mechanisms. Research using event-related potentials (ERPs) has uncovered a particular electrical signature, known as the N170, that is thought to specifically underlie processing of faces (Bentin et al 1996). In addition, processing of faces also recruits a selective region of the fusiform gyrus known as the fusiform face area (FFA—Grill-Spector et al 2004; Kanwisher et al 1997; Yovel and Kanwisher 2004).

The belief that face processing relies on unique mechanisms has been debated, however. Diamond and Carey (1986) argued that faces are only special in that they are a class of stimuli processed with expertise, and that expertise with processing results in greater sensitivity to configural information. Along these lines, Gauthier and Tarr (1997) trained participants to become experts at processing novel "Greeble" stimuli, which shared similar features and spatial arrangements to face stimuli. Greeble experts, but not novices, displayed sensitivity to changes in configural information that were orientation specific, mirroring the face inversion effects. In addition, the FFA is recruited on tasks that require visual expertise (Gauthier et al 2000; Xu 2005), and processing objects with expertise also elicits a face-like N170 response (Gauthier et al 2003; Rossion et al 2002a, 2002b; Tanaka and Curran 2001).

The primary goal of the current study was to further examine the nature of the featural/configural distinction by measuring how perception of these two sources of information differs when people are processing action relative to another stimulus.

As such, the goal was not to specifically identify shared processing mechanisms across action and faces. However, the degree to which action perception mirrors face perception and expert visual perception is at least consistent with the idea that human action may be processed with expertise.

Methodologies from the face literature were again utilized to address this goal. In this literature, only one investigation has specifically compared attention to featural and configural information in the face versus an object similar in visual complexity. Leder and Carbon (2006) tested participants' detection of featural, configural, and color changes in both schematic faces and houses. They replicated the inversion effect for configural information for faces, but no inversion effect was found for configural information in houses. Experiment 2 of the current research modeled the basic approach of Leder and Carbon, adapted for action.

Developing the control stimuli for experiment 2 ultimately engendered an additional research question, however, which was addressed in experiment 1. One major difference between the work of Loucks and Baldwin (2009) and research in the face literature is that Loucks and Baldwin used dynamic action as stimuli, while the bulk of research on the face has used static faces as stimuli. Although faces are not static stimuli in the real world, the featural/configural distinction is nonetheless observable with static faces, suggesting that the very identification of a face as a stimulus recruits differential processing to these two perceptual sources. Is this also true for action? Does processing of static representations of action recruit similar perceptual and attentional mechanisms as the processing of dynamic action? If so, it would indicate that such mechanisms are engaged for processing action per se, and not just dynamic action.

If static representations of action do engender a featural/configural distinction, their use as stimuli for experiment 2 would provide some additional benefits. First, using static action stimuli would allow for the measurement of reaction times to detecting action changes. This was not possible in Loucks and Baldwin (2009), as the two types of changes could potentially be detected at different points in time in the videos. Measurement of reaction time could help support the distinction between these two perceptual sources: if configural change detection is uniquely slowed with inversion, it would suggest that configural relations may be extracted holistically with upright action but must be processed in a slower, piecemeal fashion when action is inverted. Second, the use of static action stimuli would simplify the creation of control stimuli, as matching the nature of the changes across the two types of stimuli is more straightforward with static rather than dynamic stimuli. Finally, investigating whether the discrimination of static representations of action is reliant on distinct sources of featural and configural information was an intriguing research question in its own right, as no previous research has investigated this possibility.

Evidence does indicate that processing static images of human bodies is susceptible to inversion. Reed et al (2003) discovered an inversion effect for configural relations in human bodies that did not exist for houses. Further studies have examined the particular elements of the human body that are susceptible to inversion (Reed et al 2006), and also the particular elements that might give rise to the inversion effect (Yovel et al 2010). While this work lends support to the idea that configural relations in static images of humans may also be orientation specific, the current research diverges from this literature in two important ways. First, the stimuli of Reed et al and Yovel et al involved computer-generated figures in meaningless postures, while the current research examined processing of static representations of meaningful action in context. There may be important differences in perception that emerge when bodies are presented in the midst of performing an intentional, object-directed action in a complex, naturalistic scene.

Second, no previous research has specifically isolated the processing of featural action information in comparison to configural action information. Processing of these two perceptual sources may differ, especially in the context of human action.

Thus, experiment 1 explored whether people make use of featural and configural action information in static representations of action. The results of this experiment set the stage for experiment 2, in which the processing of featural and configural information in action was compared against the processing of a similar control stimulus.

2 Experiment 1

The goal of experiment 1 was to explore whether distinct sources of featural and configural action information are utilized to process static representations of actions, as they are for dynamic actions. To this end, static versions of featural and configural action changes were created, akin to the dynamic featural and configural action changes of Loucks and Baldwin (2009). There were two hypotheses: (i) featural information would be selectively attended to over configural information in static action, and (ii) inversion of static action would specifically undercut discrimination of configural changes while leaving discrimination of featural changes intact.

2.1 *Method*

2.1.1 *Participants.* Fifty-nine University of Oregon undergraduates (twenty-one male) received partial course credit for their participation in the experiment. Thirty participants were assigned to the upright condition, and twenty-nine were assigned to the inverted condition. Data from one participant in the upright condition who had a mean accuracy more than 3 SDs below the mean for that condition was dropped.

2.1.2 *Stimuli.* Stimuli included 32 action scenarios. Each scenario depicted a scene of everyday human action. For each scenario, there were three images: the standard action, the featural change, and the configural change. Four example scenarios describing the standard action and the featural and configural changes can be found in figure 1. The images measured 720 (w) $\times$ 540 (h) pixels.

As in Loucks and Baldwin (2009), the featural changes were modifications in the local detail of action depicted, without altering the global spatial trajectories of body parts. For instance, for the standard action of lifting a coffee mug with a whole-hand grasp around the cup of the mug, the featural change was grasping the handle of the mug instead. Configural changes, on the other hand, were global modifications in the spatial configuration of body parts, without changing the local action details. Because the stimuli were static, this implied a modification in the spatial trajectory of the action depicted. For example, for the same standard action of lifting a mug, the configural change was altering the height of the arm and mug, which altered the implied trajectory of the mug and arm through space. In addition to arm trajectories, configural changes also included modifications of the head and torso trajectories. In sum, featural changes involved an alteration in local action detail but retained the broader spatial aspects of the way in which the action was executed, whereas configural changes retained the local detail of the action but altered the broader spatial elements of the action.

Since many face-processing studies find that featural changes are more easily discriminated over configural changes, Loucks and Baldwin (2009) specifically aimed to reduce this advantage by making configural action changes that were objectively larger physical changes relative to featural action changes. The current research adopted this same approach, in an attempt to more fully equate upright processing of featural and configural changes. The same pixel-change algorithm utilized by Loucks and Baldwin was adapted for use in the present experiment to explore whether configural changes were objectively larger physical changes over featural changes. The algorithm calculated the degree of pixel change using the following formula:

$$\sum_{i=1}^{h}\sum_{j=1}^{w}\sqrt{(R_{Cij}-R_{Sij})^2+(G_{Cij}-G_{Sij})^2+(B_{Cij}-B_{Sij})^2}\ ,$$

where R, G, and B represent the red, green, and blue color values of a pixel, C and S denote the change image and standard image, i and j represent the coordinate value of the pixel, and h and w represent the height and width of the image in pixels.

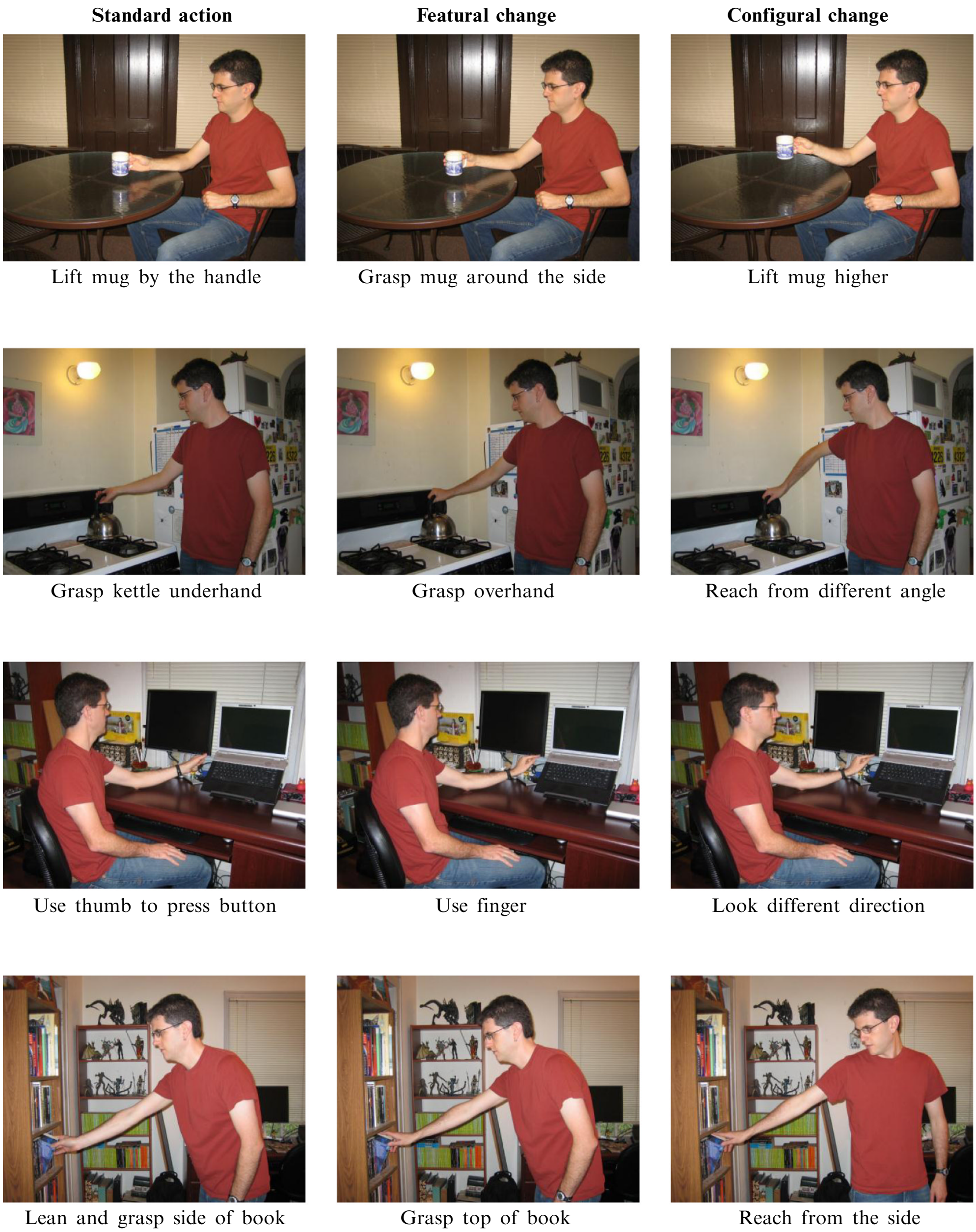

Figure 1. [In color online, see http://dx.doi.org/10.1068/p7084] Four example action scenarios used in the experiments.

For these images, configural changes ($M = 7\,712\,151$, $\mathrm{SD} = 2\,811\,122$) altered significantly more pixel information that featural changes ($M = 5\,856\,346$, $\mathrm{SD} = 1\,790\,135$), $t_{31} = 3.89$, $p < 0.001$, Cohen's $d = 0.69$. Thus, configural changes had an objective advantage in detection over featural changes.

2.1.3 *Design and procedure.* A mixed design was employed, in which change type (featural versus configural) was varied within-subjects and orientation (upright versus inverted) was varied between-subjects. On different trials, the standard image was paired with either the featural change image or the configural change image. The order of the images within different trials was counterbalanced, yielding a total of four different trials per scenario, for a total of 128 different trials. On same trials, the standard, featural, or configural image was paired with itself. The standard image was paired with itself twice, in order for there to be four same trials for each scenario. Thus, there were 128 same trials, balancing the number of same and different trials within the experiment.

On a trial, the first image was displayed for 2 s, followed by a 1 s blank screen, and then the second image was presented. The second image remained on the screen until participants responded. The inter-trial interval was self-paced: participants initiated the next trial using a button on the keyboard. In the upright condition, participants viewed all of the images in their normal upright orientation, while in the inverted condition all of the images were rotated 180° from normal. The order of presentation of the trials was completely random, with the exception that no scenario be repeated twice in a row.

A Macintosh G5 computer was used to present stimuli and record participant responses on a 19.5 inch × 12 inch cinema display. From where participants were seated, videos subtended approximately 16 deg × 12 deg of visual angle. Psychtoolbox (Brainard 1997) was used to present the trials and record responses.

After giving consent, participants were seated in front of the display and the experimenter provided instructions. They were informed that they would be watching pairs of action images and would be asked to decide whether the images in a pair were the same or different. Participants were informed that the changes would be subtle, and that on same trials the identical image file would be shown twice. Participants were asked to make their judgments as quickly and accurately as possible using assigned buttons on the computer keyboard. In the inverted condition, participants were also informed that the image pairs would be presented upside down.

2.2 *Results and discussion*

Statistical analyses were performed on d' scores, converted from participants' accuracy scores, in order to better assess sensitivity to the two change types in each condition. For reaction times, only reaction times from correct responses were analyzed; all reaction times that were beyond 2.5 standard deviations from a participant's mean correct reaction time were excluded. Preliminary analyses revealed no effect of gender on participants' responses, and thus this variable was excluded from further statistical analyses.

Mean d' scores for each of the change types in the upright and inverted conditions can be found in figure 2. A 2 (change type) × 2 (condition) mixed ANOVA revealed a main effect of change type ($F_{1,56} = 31.86$, $p < 0.001$, $\eta_p^2 = 0.36$), indicating that participants were more sensitive to featural changes ($M = 2.17$, $\mathrm{SD} = 0.62$) than to configural changes ($M = 1.77$, $\mathrm{SD} = 0.59$) overall. There was also a significant main effect of condition ($F_{1,56} = 6.27$, $p = 0.015$, $\eta_p^2 = 0.10$), indicating that participants were more sensitive to the changes in the upright condition ($M = 2.13$, $\mathrm{SD} = 0.50$) than in the inverted condition ($M = 1.81$, $\mathrm{SD} = 0.47$). Finally, the predicted interaction between change type and condition was significant ($F_{1,56} = 25.57$, $p < 0.001$, $\eta_p^2 = 0.31$).

Planned comparisons indicated that, in the upright condition, participants' sensitivity to featural and configural changes did not differ ($t_{28} < 1$) while, in the inverted condition, this difference was significant ($t_{28} = 7.00$, $p < 0.001$, Cohen's $d = 1.30$).

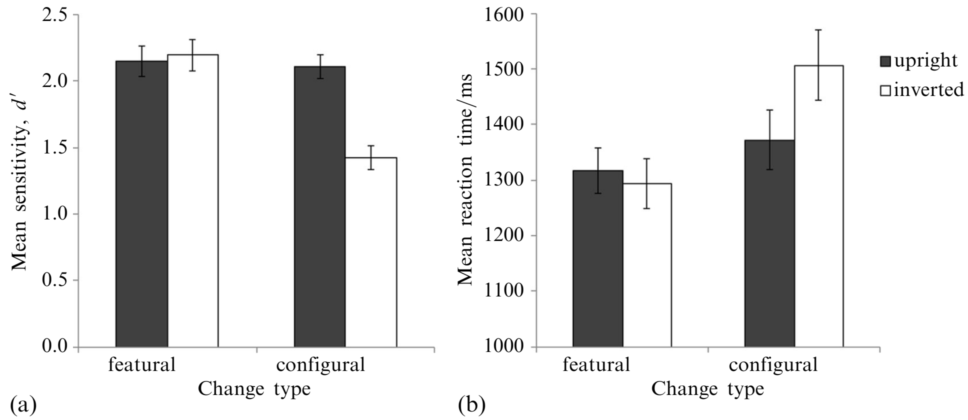

Figure 2. Mean sensitivity (a) and mean reaction time (b) to featural and configural changes in the upright and inverted conditions of experiment 1.

Comparing across conditions, sensitivity to featural changes was unaffected by inversion ($t_{56} < 1$), while sensitivity to configural changes was significantly disrupted with inversion ($t_{56} = 5.41$, $p < 0.001$, Cohen's $d = 1.03$).

Participants' reaction times to correctly detect the changes yielded a similar pattern of findings. Reaction times for each type of change in each condition can be found in figure 2. A 2 (change type) $\times$ 2 (condition) mixed ANOVA revealed a main effect of change type ($F_{1,56} = 31.75$, $p < 0.001$, $\eta_p^2 = 0.36$), indicating that participants were overall faster to identify featural changes ($M = 1305$, SD $= 228$) than configural changes ($M = 1438$, SD $= 318$). As in the d' data, there was also a significant interaction between change type and condition ($F_{1,56} = 10.94$, $p = 0.002$, $\eta_p^2 = 0.16$). There was no significant effect of condition ($F_{1,56} < 1$).

Planned comparisons revealed that, in the upright condition, the difference in reaction times to featural and configural changes was marginally significant ($t_{28} = 1.96$, $p = 0.06$, Cohen's $d = 0.39$). In the inverted condition, this difference between the change types was significant ($t_{28} = 5.55$, $p < 0.001$, Cohen's $d = 1.03$). Across conditions, the effect of inversion on reaction times was not significant for featural changes ($t_{56} < 1$), but approached significance for configural changes ($t_{56} = 1.63$, $p = 0.11$, Cohen's $d = 0.43$).

The results of experiment 1 indicate that static representations of action are processed similarly to dynamic action. Participants were sensitive to both featural and configural sources of information in static action, but when static action was inverted, sensitivity to configural information was uniquely impaired. Changes to configural information are also correctly identified significantly more slowly when action is inverted, in comparison to upright configural changes and inverted featural changes. Together, these findings further support a distinction between featural and configural information that extends to processing human action more generally, whether action is observed in a static or dynamic context.

For static representations of action, the selective attention effect for featural information may be relatively weak in comparison to the effect observed in Loucks and Baldwin (2009). For the stimuli used in the present experiment, participants were equally sensitive to the two change types in their normal orientation. However, there was a trend for featural changes to be detected somewhat more quickly than configural changes in the upright condition. Also note that sensitivity to these two change types was equal even though featural changes were significantly smaller objective psychical changes than configural changes. Perhaps at a minimum these findings suggest that selective attention to featural action information is strengthened when action is dynamic versus static.

3 Experiment 2

The main goal of the current research was to explore whether configural information is processed differently in action than in a comparable visual stimulus. Since the results of experiment 1 indicated that processing of static action is highly similar to processing dynamic action, it was now possible to create control stimuli for experiment 2 based on static representations of action.

Previous research in the face-perception literature has often used another class of stimuli as the control stimulus—most commonly, houses (Leder and Carbon 2006; Tanaka and Farah 1993; Yin 1969; Yovel and Kanwisher 2004). A drawback to this approach is that there is relatively little visual similarity between the two classes of stimuli. Specific manifestations of featural and configural information are necessarily different between faces, houses, and bodies. For example, changing the eyes in a face is not the same as changing the windows of a house, as these elements are embedded in very different visual contexts (eg framed by an oval versus a square, respectively).

To improve on previous research in this regard, in the current research control stimuli were visual analogues of the identical static representations of action used in experiment 1. That is, perceptual elements of the original action images were directly mimicked in the control stimuli. For example, one of the original action scenarios was an actor picking up a coffee mug. A control version of this stimulus involved a shape nearly identical in the shape of the arm, in an identical position on the screen. This arm shape was adjacent to a shape representing the hand, identical in position and similar in the complexity of the shape. This hand shape was overlapping a shape representing the cup, in the same manner as the hand does in the original action stimulus. In this way, the control stimuli were identical to the action stimuli in terms of where the featural and configural changes occurred, highly similar to the action stimuli in terms of the relational properties of visual elements, and also highly similar in terms of how the featural and configural changes manifested. Thus, the control stimuli were tightly matched to the action stimuli in terms of the key visual elements and spatial relationships among elements.

Thus, experiment 2 compared discrimination based on featural and configural sources of information for human action versus the control visual stimuli. The hypothesis was that processing of featural information would be similar in both conditions, and would not be impacted by inversion in the action or control condition, but that processing of configural information would differ between conditions, and would be selectively impacted by inversion only in the action condition.

3.1 *Method*

3.1.1 *Participants.* Thirty-seven University of Washington undergraduates (ten male) received partial course credit for their participation in the experiment. Eighteen participants were assigned to the action condition, and nineteen were assigned to the control condition.

3.1.2 *Stimuli.* Stimuli for the action condition included the same 32 action scenarios from experiment 1. Stimuli for the control condition were 32 control scenarios, adapted from the original 32 action scenarios. Creation of the control images was accomplished with graphical software.

For each scenario, starting with the standard action, five major visual elements of the scene were replaced with colored two-dimensional geometric shapes, in the identical positions of the original visual elements. First, to control for the featural changes, the hand in each image was always replaced by a five-pointed star shape. Second, to control for the configural changes, either the arm, head, or body was replaced by a similar-sized shape, depending on which body part represented the configural change in the scenario. Third, three other major shapes were also replaced—sometimes an object the actor was acting upon, another body part (head or body), or some other salient visual object in the scene. As mentioned previously, care was taken to ensure that the images did not resemble images of human action, and thus these last three visual elements occasionally took on slightly different shapes than the original elements. Finally, a uniform colored background was applied to the rest of the image.

From the standard control image, featural and configural control images were created. The featural change always involved an alteration in the star shape, just as the featural action change always involved an alteration of the hand shape. Configural action changes were alterations in the shape of the object that replaced the arm, head, or body, to the same degree as the original configural action alteration. Two sets of control scenarios, along with their original action scenarios, can be found in figure 3.

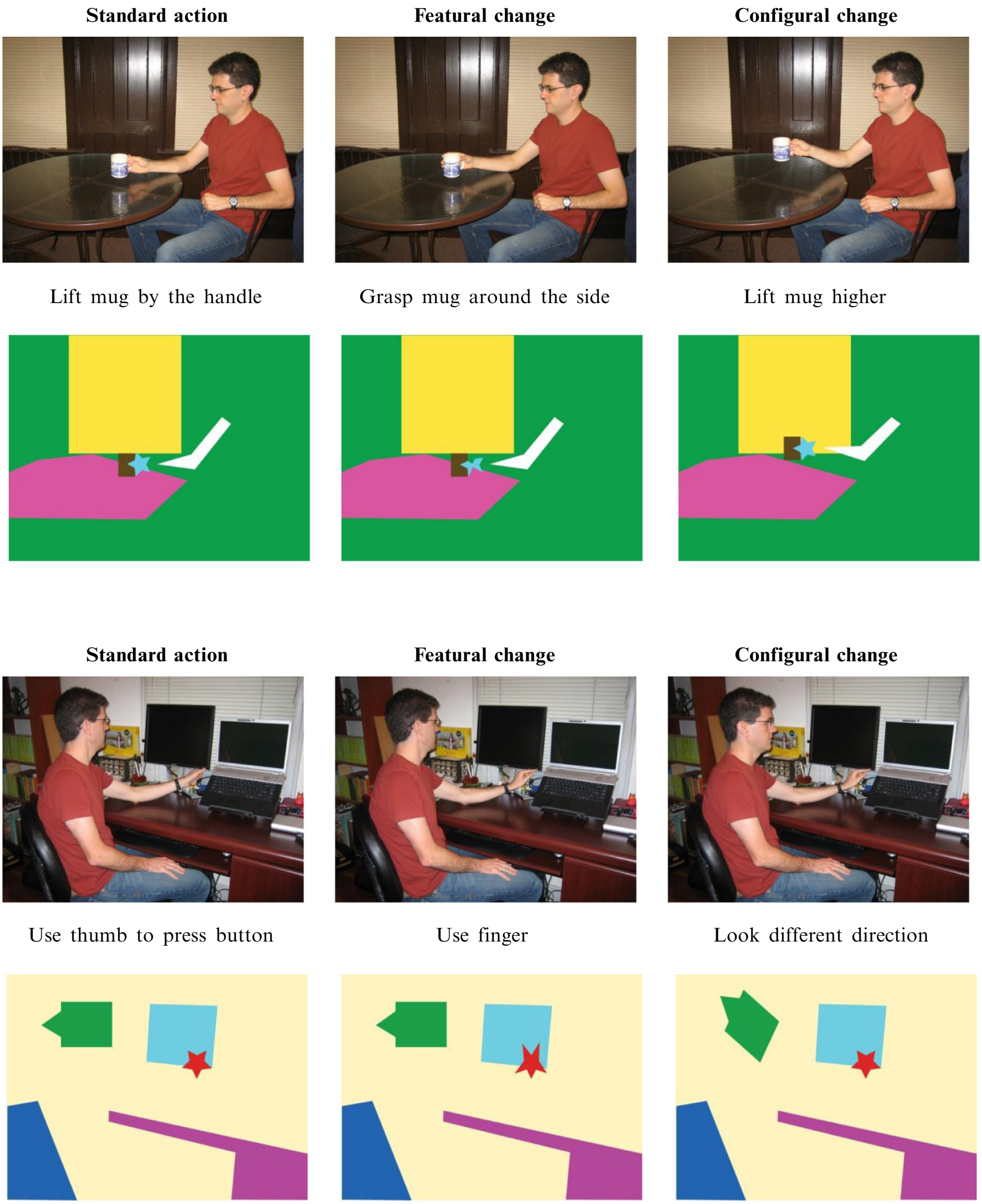

Figure 3. [In color online.] Two example control scenarios paired with their original action scenarios.

The types of shapes used and the particular elements selected in each scenario for replacement was also guided by two additional aims. First, it was important that none of the control stimuli looked like representations of human action. If the control stimuli resembled two-dimensional geometric action representations (ie a cubist work of art), then processing in the control condition would not be expected to differ from processing in the action condition. Second, it was important that detection rates across the action and control conditions were comparable.

We also measured the degree of physical change between the featural and configural control changes, using the same pixel-change algorithm described in experiment 1. As would be expected, given that the changes were derived from the original action scenarios, configural control changes ($M = 2\,627\,364$, $\mathrm{SD} = 1\,162\,492$) were significantly larger physical changers than featural changes ($M = 454\,853$, $\mathrm{SD} = 1\,186\,600$) $t_{31} = 7.59$, $p < 0.001$, Cohen's $d = 1.34$.

3.1.3 *Design and procedure.* We employed a mixed design, in which change type (featural versus configural) and orientation (upright versus inverted) were varied within-subjects, and condition (action versus control) varied between-subjects. Nearly all of the design elements were identical to experiment 1, and the structure of a trial within each condition was identical to experiment 1. However, with the addition of orientation as a within-subjects factor, two additional changes were made. First, to ensure that participants only saw a particular scenario in one particular orientation, different groups of participants saw different sets of upright and inverted scenarios. Second, within a trial, the orientation of all images were identical (all upright or all inverted).

A laptop computer was used to present stimuli and record participant responses on a 13.5 inch × 7.5 inch display. From where participants were seated, videos subtended approximately 16 deg × 12 deg of visual angle. Psychtoolbox (Brainard 1997) was again used to present the trials and record responses.

Following consent, participants were seated in front of the display and the experimenter provided instructions. They were informed that they would be watching pairs of images and would be asked to decide whether the images in a pair were the same or different. Participants were informed that the changes would be subtle, and that on same trials the identical image file would be shown twice. Participants were also told that half of the images would be presented inverted. They were asked to make their judgments as quickly and accurately as possible using assigned buttons on an attached mouse.

At the end of the experiment, all participants in the control condition were asked what the control images resembled, in order to assess for spontaneous identification of the control stimuli as representations of action. No participants spontaneously identified the pictures as images of action. After providing a response to this question, they were also asked directly if the pictures looked like people. Only one of the nineteen participants replied that "some of them do". Thus, the control images did not resemble the images of action from which they were derived.

3.2 *Results and discussion*

As in experiment 1, statistical analyses were performed on d' scores, only reaction times from correct responses were analyzed, and all reaction times beyond 2.5 standard deviations from a participant's mean correct reaction time were excluded. Preliminary analyses again revealed no effect of gender on participants' responses, and thus this variable was excluded from further statistical analyses.

Mean sensitivity scores as a function of change type and orientation in both conditions can be found in figure 4. The data for both conditions were first submitted to a 2 (change type) × 2 (orientation) × 2 (condition) mixed ANOVA, which revealed a significant three-way interaction between the variables ($F_{1,35} = 6.07$, $p = 0.019$, $\eta_\mathrm{p}^2 = 0.15$). Separate 2 × 2 repeated-measures ANOVAs were then performed on the sensitivity data

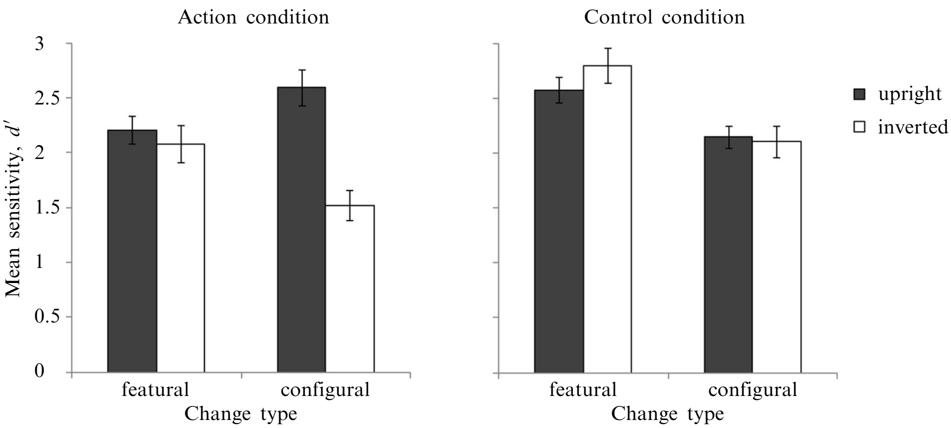

Figure 4. Mean sensitivity to upright and inverted featural and configural changes in both the action and control conditions of experiment 2.

for each condition, with change type and orientation as factors. In the action condition, there was a significant main effect of orientation ($F_{1,17} = 29.19$, $p < 0.001$, $\eta_p^2 = 0.63$), indicating that participants were more sensitive to changes in the upright orientation ($M = 2.40$, $SD = 0.55$) than in the inverted orientation ($M = 1.80$, $SD = 0.57$). There was no significant main effect of change type ($F_{1,17} < 1$). Finally, there was also the predicted significant interaction between change type and orientation ($F_{1,17} = 20.19$, $p < 0.001$, $\eta_p^2 = 0.54$).

Planned comparisons revealed that sensitivity to configural changes was significantly higher compared to featural changes in the upright orientation ($t_{17} = 2.83$, $p = 0.01$, Cohen's $d = 0.67$), but was significantly lower than featural changes in the inverted orientation ($t_{17} = 3.74$, $p = 0.002$, Cohen's $d = 0.88$). Comparing within change types, inversion did not affect sensitivity to featural changes ($t_{17} < 1$), but significantly affected sensitivity to configural changes ($t_{17} = 7.16$, $p < 0.001$, Cohen's $d = 1.69$).

The same 2×2 repeated measures applied to the sensitivity data in the control condition revealed only a significant main effect of change type ($F_{1,18} = 30.81$, $p < 0.001$, $\eta_p^2 = 0.63$), indicating that participants were more sensitive to control featural changes ($M = 2.68$, $SD = 0.52$) than control configural changes ($M = 2.13$, $SD = 0.46$). There was no significant main effect of orientation ($F_{1,18} < 1$) and no significant interaction between change type and orientation ($F_{1,18} = 2.17$, $p = 0.16$).

Planned comparisons indicated that sensitivity to featural changes was higher than sensitivity to configural changes in the upright condition ($t_{18} = 3.19$, $p = 0.005$, Cohen's $d = 0.73$), as well as in the inverted condition ($t_{18} = 5.06$, $p < 0.001$, Cohen's $d = 1.16$). However, inversion had no significant effect on sensitivity to featural changes ($t_{18} = 1.65$, $p = 0.12$) or configural changes ($t_{18} < 1$).

An overall similar pattern of findings was yielded from participants' reaction times to correctly detecting the changes in the action and control conditions. Mean reaction times as a function of change type and orientation in both conditions can be found in figure 5. Data from both conditions were again first submitted to a 2 (change type) $\times$ 2 (orientation) $\times$ 2 (condition) mixed ANOVA, which again revealed a significant three-way interaction between the variables ($F_{1,35} = 4.92$, $p = 0.033$, $\eta_p^2 = 0.12$). Thus, separate 2 (change type) $\times$ 2 (orientation) repeated-measures ANOVAs were performed for each condition. In the action condition, there was a significant main effect of orientation ($F_{1,17} = 24.33$, $p < 0.001$, $\eta_p^2 = 0.59$), indicating that participants were faster to identify changes in the upright condition ($M = 1132$, $SD = 259$) than in the inverted condition ($M = 1234$, $SD = 273$). There was no significant main effect

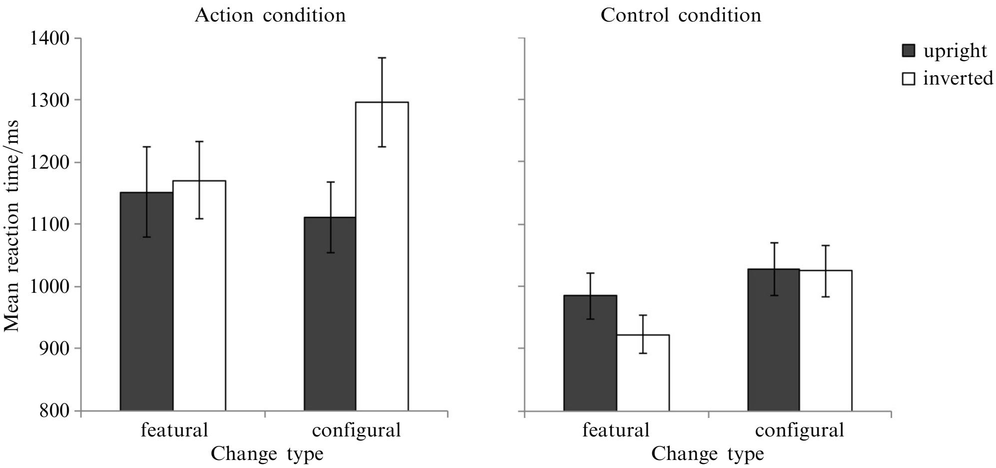

Figure 5. Mean reaction times to upright and inverted featural and configural changes in both the action and control conditions of experiment 2.

of change type ($F_{1,17} = 1.59$, $p = 0.22$). Finally, as in experiment 1, there was a significant interaction between change type and orientation ($F_{1,17} = 16.64$, $p < 0.01$, $\eta_p^2 = 0.50$).

Planned comparisons revealed that reaction times to featural and configural action changes did not differ significantly in the upright condition ($t_{17} < 1$), but were significantly higher in the inverted condition for configural changes ($t_{17} = 4.00$, $p < 0.01$, Cohen's $d = 0.94$). Inversion did not have a significant impact on reaction times for featural change detection ($t_{17} < 1$), but did significantly increase reaction times to configural change detection ($t_{17} = 6.06$, $p < 0.001$, Cohen's $d = 1.43$).

The same 2×2 repeated-measures ANOVA applied to the reaction times in the control condition revealed a significant main effect of change type ($F_{1,18} = 9.38$, $p < 0.1$, $\eta_p^2 = 0.34$), indicating that detection of featural changes ($M = 953$, SD = 132) was significantly faster than detection of configural changes ($M = 1026$, SD = 174). The main effect of orientation approached significance ($F_{1,18} = 3.21$, $p = 0.09$, $\eta_p^2 = 0.15$), as reaction times were lower in the inverted condition ($M = 974$, SD = 145) than in the upright condition ($M = 1006$, SD = 156). Finally, there was also a marginally significant interaction between change type and orientation ($F_{1,18} = 4.56$, $p = 0.047$, $\eta_p^2 = 0.20$).

Planned comparisons indicated that, in contrast to the action condition, inversion significantly reduced reaction times to control featural changes ($t_{18} = 2.44$, $p < 0.05$, Cohen's $d = 0.56$), but did not have any significant effect on reaction times to control configural changes ($t_{18} < 1$). There was also no significant difference between control featural and configural reaction times in the upright condition ($t_{18} = 1.54$, $p = 0.14$), while reaction times were significantly lower for featural compared to configural changes in the inverted condition ($t_{18} = 3.86$, $p < 0.01$, Cohen's $d = 0.88$).

Finally, inspecting the mean sensitivity to configural changes in the action condition versus the control condition, there appeared to be an advantage for detecting configural action changes in comparison to control changes. Indeed, this difference was significant ($t_{35} = 2.38$, $p = 0.023$, Cohen's $d = 0.78$). Thus, configural changes were easier to detect in the action condition, suggesting an advantage for configural processing of spatial relations in action in comparison to the control stimuli.

The results of experiment 2 provide compelling evidence that configural information is processed differently in human action in comparison to similar visual stimuli. Processing of featural information is insensitive to inversion for both types of stimuli, while processing of configural information is sensitive to inversion only in human action. This finding is especially striking given the fact that nearly identical types of configural transformations were presented in the control condition. However, participants detected these changes at equal levels of sensitivity in the two orientations in the control condition. This inversion difference was manifest in participants' sensitivity to configural action changes, as well as in their reaction times to detecting the changes. In addition, configural changes were easier to detect in the upright action condition, suggesting an advantage for such changes in the context of human action. Altogether, these results strengthen the claim that configural processing in human action differs qualitatively from comparable visual stimuli.

Once again there appeared to be no significant advantage for featural over configural change detection in the upright action condition—neither in the sensitivity nor reaction-time data. In fact, the sensitivity data revealed the reverse advantage for configural changes in this experiment. This finding may have been due to the fact that configural changes were objectively larger physical changes over featural changes. However, the objective differences cannot fully explain this finding, since it was not present with the identical stimuli in experiment 1. Furthermore, while there was a marginally significant reaction-time advantage for upright featural changes in experiment 1, this advantage did not replicate in experiment 2. Overall, this further supports the possibility that the selective attention effect for featural action information is stronger for dynamic over static action stimuli.

4 General discussion

As people observe the actions of others, they are sensitive to two distinct sources of perceptual information: featural information regarding fine motion elements, and configural information regarding the spatial relationships of action elements (Loucks and Baldwin 2009). The current experiments expand upon this featural/configural distinction in two ways. First, experiment 1 demonstrated that these sources of information are processed in similar ways for static representations of action as they are for dynamic action. Configural information is uniquely sensitive to orientation in static as in dynamic action; however, featural information does not appear to be selectively attended to over configural information in static action. Thus, once human action is identified in a scene processing of the configural relations in the action differs from processing of featural information, regardless of whether the action is identified in a static or dynamic context. Second, experiment 2 demonstrated that configural information is processed differently when observing human action as compared to a control visual stimulus. These results may indicate that configural relations in upright human action are extracted in a more holistic manner than they are in other types of visual stimuli.

The featural/configural distinction in action processing was borrowed from research on face perception, as the stimuli in both domains are complex to the degree that perception demands an efficient processing system. In the face and in action, perception seems to rely on similar, distinct sources of perceptual information to this end. In particular, the fact that configural information is selectively sensitive to orientation in both domains suggests that faces and human action utilize similar processing mechanisms. The current experiments strengthen the link between the way these sources of information are processed in these two domains. Importantly, comparable visual stimuli do not elicit the same featural/configural distinction across both domains (Leder and Carbon 2006). Processing of configural information in other stimuli certainly occurs, but it is not orientation-specific as it is in faces and action.

Another perspective on the current findings is that the featural/configural distinction is not unique to faces. One possibility is that qualitatively different configural processing occurs in action because action is processed with expertise. This idea was first proposed by Diamond and Carey (1986), and experimental evidence suggests that attention to orientation-specific configural information is enhanced when objects are processed with expertise (Gauthier and Tarr 1997). There is reason to believe that action may be processed with expertise; efficient processing is a crucial step towards inferring other people's goals and mental states, activities which are carried out rapidly countless times each day (Baldwin 2005). This idea may help explain why processing of upright configural information was superior in experiment 2 relative to the control stimulus. In the action condition, it might seem as though the configural changes would be harder to detect, as they are embedded in more complex visual scenes. However, the reverse finding emerged. This suggests that identifying human action within a scene actually enhances attention to configural information. As people become more adept with processing the actions of others, they may become more sensitive to the configural properties of action. This is an intriguing developmental question for future research (along the lines of Mondloch et al 2002).

Note that what is extracted as featural versus configural information in a scene is likely somewhat dependent on scale. One can imagine that a featural change in hand contact that is relatively local perceptual information at a viewing distance encompassing the whole body that could be extracted as configural information if one was examining the hand close-up. Featural and configural sources of information can be characterized as different ends of a spatial frequency continuum (eg Morrison and Schyns 2001). Importantly, the current results demonstrate a difference in processing these two sources of information that is specific to human action when observed at a fairly standard viewing distance.

An important next step in this line of work is to compare processing of dynamic action stimuli with a dynamic control stimulus. While the current findings suggest common mechanisms in the perception of dynamic and static action, they cannot fully capture the differences between action and other dynamic stimuli, since the selective attention for featural information is not present with static action stimuli. Using animated control stimuli akin to the static control stimuli used in experiment 2 could shed light on whether this selective attention is a property of action processing per se, or whether it is a property of dynamic motion perception more generally.

Another important future direction is to assess the neural correlates of the featural/configural distinction in action processing. The results of experiment 1 suggest that static representations of action can be used in place of dynamic action, which is of great value to future research utilizing ERP techniques, as the presentation of each type of change can be equated in time (which is not possible with dynamic actions). Examining ERPs to human action could help strengthen the claim that action is processed with expertise, as several studies have documented that the N170 is elicited in contexts of visual expertise (Gauthier et al 2003; Rossion et al 2002a, 2002b; Tanaka and Curran 2001). It would be interesting to explore whether an N170 could also be elicited in the context of action discrimination. Very few studies have specifically examined ERPs to featural and configural processing, and these have been restricted to face stimuli (Mercure et al 2008; Scott and Nelson 2006). Exploring the neural correlates of featural and configural action processing would thus not only broaden our knowledge of the neural underpinnings of action processing, but would also aid in understanding the perceptual processes shared across domains of visual expertise.

The present research helps to clarify what forms of perceptual processing are unique—or at least specialized—for action processing. Evidence thus far indicates that configural relations in action are processed more holistically than in comparable stimuli, by virtue of the fact that configural information is more prominent in action and is uniquely disrupted by inversion. This research paves the way for future research into the shared and distinct perceptual and neural mechanisms that underlie action processing and perceptual expertise.

Acknowledgments. This research was supported by an NRSA postdoctoral fellowship (F32HD058445-01A2) awarded to me. An NSF grant (BCS-0214484) awarded to Dare Baldwin and an NICHD grant (1R03HD053616-01A1) awarded to Jessica Sommerville also partly supported this research. I am grateful to Meredith Meyer for helpful comments and suggestions on a previous version of this manuscript, and would also like to thank Pete Fisher and David Anderson for help with data collection.

References

Baldwin D, 2005 "Discerning intentions: Characterizing the cognitive system at play", in *The Development of Social Cognition and Communication* Eds B D Homer, C S Tamis-LeMonda (Mahwah, NJ: Lawrence Erlbaum Associates)

Belle G van, Graef P de, Verfaillie K, Rossion B, Lefèvre P, 2010 "Face inversion impairs holistic perception: Evidence from gaze-contingent stimulation" *Journal of Vision* **10**(5):10, 1 – 13

Bentin S, Allison T, Puce A, Perez E, 1996 "Electrophysiological studies of face perception in humans" *Journal of Cognitive Neuroscience* **8** 551 – 565

Brainard D H, 1997 "The psychophysics toolbox" *Spatial Vision* **10** 433 – 436

Diamond R, Carey S, 1986 "Why faces are and are not special: An effect of expertise" *Journal of Experimental Psychology: General* **115** 107 – 117

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

Gauthier I, Tarr M J, 1997 "Becoming a 'greeble' expert: Exploring mechanisms for face recognition" *Vision Research* **37** 1673 – 1682

Gauthier I, Curran T, Curby K M, Collins D, 2003 "Perceptual interference supports a non-modular account of face processing" *Nature Neuroscience* **6** 428 – 432

Gauthier I, Skudlarski P, Gore J C, Anderson A W, 2000 "Expertise for cars and birds recruits brain areas involved in face recognition" *Nature Neuroscience* **3** 191 – 197

Goffaux V, Hault B, Vuong Q C, Rossion B, 2005 "The respective role of low and high spatial frequencies in supporting configural and featural processing of faces" *Perception* **34** 77 – 86

Grill-Spector K, Knouf N, Kanwisher N, 2004 "The fusiform face area subserves face perception, not generic within-category identification" *Nature Neuroscience* **7** 555 – 562

Kanwisher N, McDermott J, Chun M M, 1997 "The fusiform face area: A module in human extrastriate cortex specialized for face perception" *Journal of Neuroscience* **17** 4302 – 4311

Leder H, Carbon C-C, 2006 "Face-specific configural processing of relational information" *British Journal of Psychology* **97** 19 – 29

Loucks J, Baldwin D, 2009 "Sources of information for discriminating dynamic human actions" *Cognition* **111** 84 – 97

Maurer D, Le Grand R, Mondloch C J, 2002 "The many faces of configural processing" *Trends in Cognitive Sciences* **6** 255 – 260

Mercure E, Dick F, Johnson M H, 2008 "Featural and configural face processing differentially modulate ERP components" *Brain Research* **1239** 162 – 170

Mondloch C J, Le Grand R, Maurer D, 2002 "Configural face processing develops more slowly than featural face processing" *Perception* **31** 553 – 566

Morrison D J, Schyns P G, 2001 "Usage of spatial scales for the categorization of faces, objects, and scenes" *Psychonomic Bulletin & Review* **8** 454 – 469

Murray J E, Yong E, Rhodes G, 2000 "Revisiting the perception of upside-down faces" *Psychological Science* **11** 492 – 496

Reed C L, Stone V E, Bozova S, Tanaka J W, 2003 "The body-inversion effect" *Psychological Science* **14** 302 – 308

Reed C L, Stone V E, Grubb J D, McGoldrick J E, 2006 "Turning configural processing upside down: Part and whole body postures" *Journal of Experimental Psychology: Human Perception and Performance* **32** 73 – 87

Rossion B, 2008 "Picture-plane inversion leads to qualitative changes of face perception" *Acta Psychologica* **128** 274 – 289

Rossion B, Curran T, Gauthier I, 2002a "A defense of the subordinate-level expertise account for the N170 component" *Cognition* **85** 189 – 196

Rossion B, Gauthier I, Goffaux V, Marr M J, Crommelinck M, 2002b "Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses" *Psychological Science* **13** 250 – 257

Scott L S, Nelson C A, 2006 "Featural and configural face processing in adults and infants: A behavioral and electrophysiological investigation" *Perception* **35** 1107 – 1128

Tanaka J W, Curran T, 2001 "A neural basis for expert object recognition" *Psychological Science* **12** 43 – 47

Tanaka J W, Farah M J, 1993 "Parts and whole in face recognition" *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology* **46** 225 – 245

Xu Y, 2005 "Revisiting the role of the fusiform face area in visual expertise" *Cerebral Cortex* **15** 1234 – 1242

Yin R K, 1969 "Looking at upside-down faces" *Journal of Experimental Psychology* **81** 141 – 145

Young A W, Hellawell D, Hay D C, 1987 "Configurational information in face perception" *Perception* **16** 747 – 759

Yovel G, Kanwisher N, 2004 "Face perception: Domain specific, not process specific" *Neuron* **44** 889 – 898

Yovel G, Pelc T, Lubetzky I, 2010 "It's all in your head: Why is the body inversion effect abolished for headless bodies?" *Journal of Experimental Psychology: Human Perception and Performance* **36** 759 – 767

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 40 2011

www.perceptionweb.com