Children are sensitive to mutual information in intermediate-complexity face and non-face features

Benjamin Balas

Department of Psychology, Center for Visual and Cognitive Neuroscience, North Dakota State University, Fargo, ND, USA ✉

Amanda Auen

Department of Psychology, Center for Visual and Cognitive Neuroscience, North Dakota State University, Fargo, ND, USA ✉

Alyson Saville

Department of Psychology, Center for Visual and Cognitive Neuroscience, North Dakota State University, Fargo, ND, USA ✉

Jamie Schmidt

Department of Psychology, Center for Visual and Cognitive Neuroscience, North Dakota State University, Fargo, ND, USA ✉

Assaf Harel

Department of Psychology, Center for Visual and Cognitive Neuroscience, North Dakota State University, Fargo, ND, USA ✉

Understanding developmental changes in children's use of specific visual information for recognizing object categories is essential for understanding how experience shapes recognition. Research on the development of face recognition has focused on children's use of low-level information (e.g. orientation sub-bands), or high-level information. In face categorization tasks, adults also exhibit sensitivity to intermediate complexity features that are diagnostic of the presence of a face. Do children also use intermediate complexity features for categorizing faces and objects, and, if so, how does their sensitivity to such features change during childhood? Intermediate-complexity features bridge the gap between low- and high-level processing: they have computational benefits for object detection and segmentation, and are known to drive neural responses in the ventral visual system. Here, we have investigated the developmental trajectory of children's sensitivity to diagnostic category information in intermediate-complexity features. We presented children (5–10 years old) and adults with image fragments of faces (Experiment 1) and cars (Experiment 2) varying in their mutual information, which quantifies a fragment's diagnosticity of a specific category. Our goal was to determine whether children were sensitive to the amount of mutual information in these fragments, and if their information usage is different from adults. We found that despite better overall categorization performance in adults, all children were sensitive to fragment diagnosticity in both categories, suggesting that intermediate representations of appearance are established early in childhood. Moreover, children's usage of mutual information was not limited to face fragments, suggesting the extracting intermediate-complexity features is a process that is not specific only to faces. We discuss the implications of our findings for developmental theories of face and object recognition.

Introduction

Studying when specific types of visual information are used for recognizing different object categories across development is an important way of understanding how learning and experience affect high-level visual representations (Nelson, 2001). Faces, as a category, are particularly interesting to consider both in terms of the specialized neural systems supporting face perception and dedicated modes of visual processing

that are specific to faces (Chung & Thomson, 1995; McKone, Crookes, & Kanwisher, 2009). Examining how representations supporting face categorization change during development can, thus, provide insight into how neural mechanisms for processing different objects become differentiated from one another as observers gain experience with their visual environment (Nelson, 2001).

Our goal in the present study was to examine developmental changes in the information used to distinguish faces from non-faces, a categorization task we will refer to as a *face detection* judgment. In general, face detection defined in this manner has not received nearly as much attention as face individuation or categorization tasks involving emotion recognition (see Bindemann & Burton, 2009 and Robertson, Jenkins, & Burton, 2017 for a discussion of this with regard to the adult literature), but there are several important results that suggest changes in information use for face detection over developmental time. In infancy, newborn infants have rudimentary face detection skills that initially appeared to be based on a simple "top-heavy" preference for faces that respect the crude first-order arrangement of facial features (Simion et al., 2002), but more recent work suggests that even at the ages of 3 to 5 months, infants may have a more refined representation of what faces look like (Chien, 2011). This is further supported by recent studies using a rapid-presentation electroencephalogram (EEG) design that yielded clear distinctions between face and object responses over right occipitotemporal recording sites (De Heering & Rossion, 2015). In childhood, there is far less data describing how representations of facial appearance for detection and face/non-face categorization may change, but there is some evidence that specific kinds of information use may require continued development during middle childhood. Using the Thatcher Illusion, Donnelly and Hadwin (2003), demonstrated that 6-year-old children were not sensitive to the illusion, but that 8-year-old children were. This result suggests that between ages 6 and 8, children's representations of the typical configuration of faces may be changing. This is supported by results using Mooney faces (Carbon, Gruter, & Gruter, 2013) and pareidolic faces (Guillon et al., 2016; Ryan, Stafford, & King, 2016), both of which depend on mature spatial integration abilities that children seem to lack early in childhood. Further, although de Heering and Rossion's (2015) rapid-presentation EEG data suggests clear distinctions between face and non-face neural responses, event-related potential (ERP) responses measured in childhood are not always as distinct across face and object categories (Taylor et al., 2004, although see Kuefner et al., 2010). Thus, the extant literature suggests that the crude face detection abilities that are evident in infancy may become refined during childhood, but there is as yet little data describing the

time-course of that tuning, or the link between face detection abilities and the use of specific visual features.

Face recognition differs from other object recognition tasks in terms of the visual information that best supports performance in different face recognition tasks, with some visual features being more useful than others. Therefore, understanding how mechanisms for face recognition depend on different visual features is an important way to characterize how face recognition develops, as it describes its unique properties in terms of specific computations. Because faces are complex, natural stimuli, there are many candidate features to consider. In large part, the developmental face recognition literature has focused on feature vocabularies that we will refer to as *high-level* and *low-level* visual features. By *high-level* visual features, we refer broadly to presumed mechanisms for face detection and recognition that use global features, and are typically described as reflecting "holistic" or "configural" face processing (Maurer, Le Grand, & Mondloch, 2002). Developmentally, there is substantial evidence that children's use of holistic or configural features for recognition develops during childhood (de Heering et al., 2007; Mondloch et al., 2007; Pellicano & Rhodes, 2003; Schwarzer, 2000; Sangrigoli & de Schonen, 2004). By *low-level* features, we refer to visual features that typically involve local processing of the image, and reflect basic image properties measured at early stages of the visual system (e.g. spatial frequency or orientation). In adults, there are measurable information biases favoring intermediate spatial frequencies (Nasanen, 1999) and horizontal orientation energy (Dakin & Watt, 2009) for various face recognition tasks. Developmentally, there is again substantial evidence that these biases change during infancy, middle childhood, and beyond (Leonard, Karmiloff-Smith, & Johnson, 2010; Gao & Maurer, 2011; De Heering et al., 2016; Balas, Schmidt, & Saville, 2015; Balas et al., 2015; Goffaux, Poncin, & Schiltz, 2015; Obermeyer et al., 2012). Although both of these bodies of research are intriguing in their own right, it remains difficult to understand how low-level information is integrated into high-level representations of face appearance, and how development proceeds at multiple scales of feature complexity. Without more ideas about how to build a bridge between low-level and high-level visual information, a comprehensive description of how information use changes developmentally in the context of face recognition will likely remain elusive.

Presently, we adopt a third approach to characterizing the features used for face recognition, one that relies on applying tools for identifying meaningful *intermediate-level* features for face detection (determining that an image depicts a face rather than another object class) that are on the one hand category-specific (unlike low-level features), and at the same time are easy to

describe computationally (unlike high-level features). By an intermediate-level feature, we refer to a description of the object that relies neither on purely local nor purely global information. There are several reasons to consider the contribution that such features make to face recognition. First, intermediate-scale features are capable of capturing global-scale information about object appearance (Peterson, 1996), and may be a more flexible means of measuring larger image structures. Second, for more complex objects, intermediate features offer distinct computational advantages in multiple tasks. Intermediate-level image fragments are more useful than global views for face classification (Ullman & Sali, 2000) and segmentation (Ullman, 2007), both of which are challenging computational tasks. Finally, physiological studies of the ventral visual system have demonstrated that some cortical areas are comprised of cells that appear to be sensitive to intermediate-level visual features (Tanaka, 1996), suggesting a neural code for objects in terms of intermediate-level descriptors. Intermediate-level visual features are, thus, of clear computational value, appear to be measured in the ventral visual stream, and make contributions to behavioral visual recognition tasks, yet have not been used in developmental studies of face detection or recognition.

Our goal, therefore, in the current study is to begin to build a bridge between low- and high-level features by investigating developmental changes in sensitivity to visual features of *intermediate complexity.* Specifically, we use a computational measure first proposed by Ullman, Vidal-Naquet, and Sali (2002) to classify image fragments of face and other object categories as a function of the *mutual information* they provide about their category (see Figure 1 for example of face fragments). Mutual information can be used to index the diagnosticity of an image fragment for basic-level categorization (i.e. determining that an image contains a face) by measuring the probability of how likely we are to find that pattern of intensity values in other face images and how likely we are to find that pattern in non-face images. A fragment will have high mutual information (or diagnosticity) if it is both highly likely to be found in a face image and unlikely to be found in non-face images, thus balancing the need for specificity and generalization. Mutual information values vary a great deal across different face fragments (Figure 1), indicating that some portions of a face are better indicators of category than others. Notably, fragments with high mutual information tend to be of intermediate complexity: They are neither very large or highly detailed fragments, nor are they very small and generic. In terms of behavioral responses, adult observers are faster and more accurate to categorize face fragments the higher mutual information they carry (Harel et al., 2007, Harel et al., 2011) indicating that this computational index is a perceptually
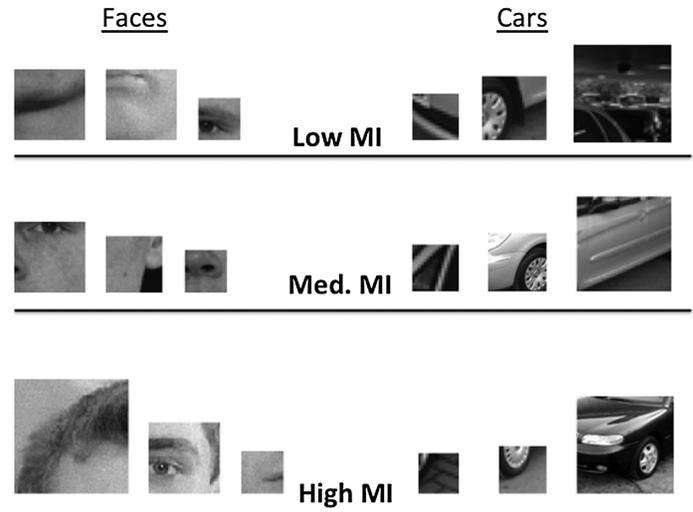

Figure 1. Examples of face and car fragments with varying levels of mutual information. We have included fragments of different sizes from all three levels of mutual information (MI) to demonstrate the variability within our stimulus set.

meaningful characterization of how diagnostic a fragment is. In terms of their neural correlates, face fragments with high mutual information also elicit more robust face-sensitive cortical responses then low mutual information face fragments (ERP: Harel et al., 2007; functional magnetic resonance imaging [fMRI]: Lerner et al., 2008), which further suggests that mutual information can serve as a useful means to characterize diagnostic face information at an intermediate level of representation.

Our investigation of intermediate representations supporting face categorization relies on an a priori definition of mutual information that we use to quantify how useful particular image fragments should be for signaling the presence of a face. This differs from other techniques for examining intermediate-level representations of faces that empirically determine what facial features observers rely upon to carry out specific recognition tasks. The "Bubbles" methodology, for example, is a form of reverse correlation in which a random mask of bubble-like apertures is applied to the face images used in a categorization task, allowing the experimenter to determine which parts of the face tended to be visible when the task was carried out correctly (Gosselin & Schyns, 2001). This methodology can be elaborated upon to investigate the use of specific spatial-frequency sub-bands (Willenbockel et al., 2010) to yield a "Bubbles" image that can be interpreted in terms of intermediate-level feature use. This technique is challenging to use with children, however, due to its dependence on a relatively large number of trials (though see Humphreys et al., 2006 and Ewing et al., 2018 for a discussion of applying the "Bubbles" technique in developmental populations). Moreover,

although the resulting Bubbles images yield insights into diagnostic parts of the face for a given task, it is a non-trivial problem to use those images to make strong inferences about the underlying vocabulary of recognition at a specific level of the visual system. We suggest that our use of image fragments as a tool for studying intermediate level representations of faces is a useful complement to the Bubbles technique and offers an important and different perspective on the features that may contribute to face recognition tasks at intermediate stages.

Thus, our goal in the current study was to examine children's sensitivity to mutual information level in face fragments and non-face object fragments (Experiment 1 and Experiment 2, respectively) in an object categorization task. To do so we asked children between the ages of 5 and 10 years of age to carry out a simple detection task (a go/no-go judgment for each target category) using fragments that varied according to the mutual information (MI) they provided regarding target category. This detection task was used in a previous study examining the contribution of different orientation sub-bands to face recognition (Balas et al., 2015), as it is easier and more intuitive for young children than the standard 2AFC categorization task (see methods below). We predicted that sensitivity to the mutual information level might be evident already during childhood, reflecting ongoing tuning of mid-level representations for facial appearance during this age range. Further, we hypothesized that the developmental trajectory associated with the mutual information level in face fragments might differ from that of non-face fragments, mirroring the emergence of other face-selective behavioral signatures during development (Balas et al., 2015; Balas et al., 2016; Gao & Maurer, 2011; Goffaux, Poncin, & Schiltz, 2015; Mondloch et al., 2004).

Experiment 1

In our first task, we examined children's and adult's sensitivity to mutual information level in a face detection task. In particular, we investigated whether or not children were sensitive to the higher diagnosticity of face fragments with a higher mutual information level, and if the degree of that sensitivity changed with age during middle childhood (5–10 years).

Methods

Participants

We recruited a total of 61 participants from the Fargo-Moorhead community to take part in this study.

Because we did not have clear a priori predictions regarding the specific time point in which sensitivity to mutual information will emerge, we divided our child participants into two groups based on age: 5 to 7-year-old children ($N = 19$; 12 girls; mean age = 6 years, 2 mos.) and 8 to 10-year-old children ($N = 22$; 10 girls; mean age = 8). Adult participants ($N = 20$; 12 women) were recruited from the NDSU Undergraduate Psychology study pool. This sample size was chosen based on a power analysis we conducted using estimated effect sizes derived from the results reported in Harel et al. (2007), which used an expanded set of MI levels in a categorization task. Based on the differences they reported between the levels of MI we used in this task (their lowest, mid-range, and highest MI levels), we estimated that a minimum sample size of 18 participants per age group would be adequate to detect the critical interactions between factors in our design assuming medium-size effects (consistent with the results reported in their manuscript). Children and their caregivers received compensation for their participation, whereas adult participants received course credit. Informed consent was obtained from adult participants and the legal guardians of child participants. Children 7 years and older also provided written assent to participate. All participants (or their caregivers) reported either normal or corrected-to-normal vision.

Stimuli

We presented our participants with gray-scale images of face and non-face (car) fragments. Both classes of fragments were previously used in behavioral and electrophysiological (Harel et al., 2007; Harel et al., 2011) studies designed to reveal how adult face processing is impacted by the extent to which image fragments are diagnostic of face/non-face categories. The most important feature of this stimulus set is that for each fragment, the *mutual information* provided by that fragment for category membership has been calculated. Briefly, this is measure is a way to quantify how informative a given fragment is regarding the presence of a category (e.g. "face"). The mutual information associated with each fragment is calculated by determining how likely one is to find the fragment in images containing a face and how likely one is to find the fragment in images that do not contain faces. A high value implies that the fragment is very diagnostic: it appears frequently in images containing a face, and appears rarely in images containing non-faces. A detailed description of how mutual information is calculated can be found in (Ullman, Vidal-Naquet, & Sali, 2002), but for our purposes, the critical property of the stimuli is that we know how this value varies across our stimulus set. Critically, all of the fragments in our stimulus set are intermediate complexity fragments, but

they vary in their mutual information level, and, thus, their diagnosticity.

We selected 96 non-face fragments depicting portions of cars and 96 face fragments. All car fragments used in Experiment 1 had a "medium" level of diagnosticity for their category. Face fragments were selected to include 32 "low" mutual information fragments, 32 "medium" mutual information fragments, and 32 "high" mutual information fragments. Within each of these three levels, fragment size varied, and the regions of the face depicted within each fragment also varied. Low MI face fragments had a mean width of 26.25 units (SD = 4.64), medium MI face fragments had a mean width of 28.0 units (SD = 5.7), and high MI face fragments had a mean width of 38 units (SD = 15.3). "Units" here refers to the square root of image area in pixels in the raw images, but all images were scaled up by a uniform factor for presentation during the task. This yielded average sizes of approximately 3.5 degrees of visual angle for low and medium MI face fragments and approximately 4.75 degrees of visual angle for high MI face fragments. We emphasize, however, that within each MI level, fragments differed in size and aspect ratio, and also that participants of different ages likely also varied in terms of the visual angle subtended by each stimulus due to variation in the comfortable placement of the monitor and chair during the testing session. Examples of fragment stimuli can be found in Figure 1.

Procedure

We asked participants to complete a "go/no-go" judgment using the full set of face and non-face fragments, similar to prior work examining the contribution of different orientation sub-bands to face recognition (Balas et al., 2015). We chose this task primarily because it is easier for our youngest participants to understand and execute than a 2AFC categorization task. Young children do not need to remember which button goes with which response, for example, nor do they have to provide a response to stimuli that they have difficulty identifying as a member of either category.

Stimuli were presented in a pseudo-randomized order for 3 seconds each, (see Balas et al., 2015) and participants were instructed to press a large red button as quickly as possible if and only if the fragment depicted part of a face. Otherwise, participants were asked to withhold making a response and wait until the image disappeared. We included a 1-second inter-stimulus interval between trials and presented each image once.

Participants completed the task seated approximately 60 cm away from a 1200 × 800 MacBook Pro laptop, although this distance differed across different participants depending on what was required for
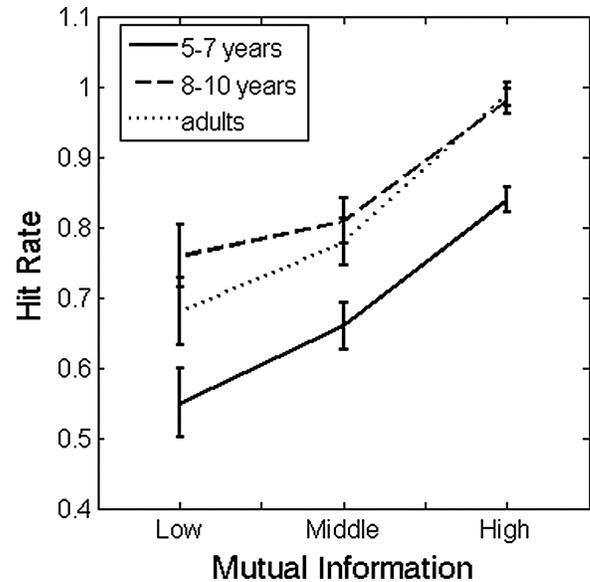

Figure 2. Average hit rate as a function of diagnosticity (mutual information) and age (5–7 years old, 8–10 years old, and adults) for face detection. Error bars indicate +/− 1 SEM.

comfortable seating during the task. All stimulus display and response collection routines were written using the Psychtoolbox extensions for Matlab (Brainard, 1997; Pelli, 1997). The task was administered in a dark, sound-attenuated room, and most participants completed the task in approximately 20 minutes.

Results

We examined three aspects of participants' performance: Their *hit rate* for correctly labeling face fragments as faces (Figure 2), their *response latency* for correctly labeled face stimuli (Figure 3), and their *false alarm rate* for incorrectly labeling non-face fragments as faces. We analyzed the hit rate and response latency data using a 3 × 3 mixed-design analysis of variance (ANOVA) with participant age (5–7 years old, 8–10 years old, and adults) as a between-subjects factor and face diagnosticity (low, medium, or high) as a within-subjects factor. To analyze false alarm rates, we carried out a 3 × 1 ANOVA with participant age as a between-subjects factor. We did not combine hit rates and false alarm rates into a composite accuracy score because the false alarm rate for each observer is shared across all three hit rates, making it not especially useful for converting these values to d' or other signal detection theory descriptors of performance. Specifically, differences in d' and response bias across conditions within one participant can only reflect hit rate differences given that there is only one source of false alarms.
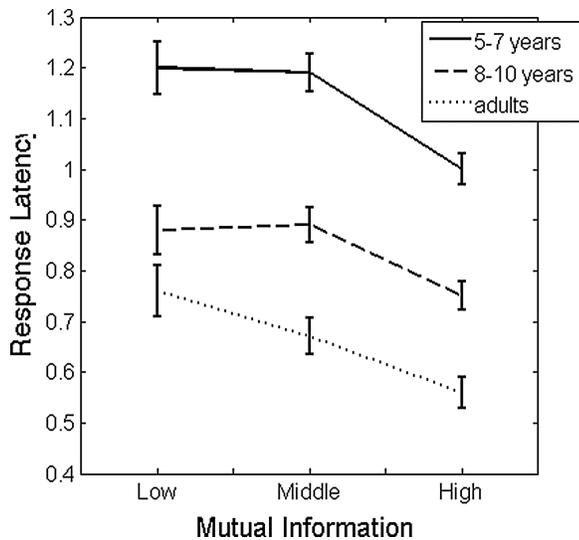
Figure 3. Average response latency for correctly labeled faces as a function of diagnosticity (mutual information) and age (5–7 years old, 8–10 years old, and adults). Error bars indicate +/− 1 SEM.

Hit rates

Our analysis of participant hit rates revealed a main effect of diagnosticity ($F (2, 116) = 99.66$, $p < 0.001$, partial $\eta^2 = 0.63$), and a main effect of age group ($F (2, 58) = 9.39$, $p < 0.001$, partial $\eta^2 = 0.98$). The interaction between diagnosticity and age group did not reach significance ($F (4, 116) = 1.18$, $p = 0.325$, partial $\eta^2 = 0.049$). The main effect of diagnosticity was the result of significant differences between all three levels of mutual information: hit rates to high MI fragments differed from medium MI fragments (95% confidence interval [CI] of the difference = 0.15–0.23) and low MI fragments (95% CI of the difference = 0.22–0.34), and performance with medium MI fragments also differed from low MI fragments (95% CI of the difference = 0.036–0.13). The main effect of age was the result of a significant difference between 5 to 7-year-old performance (M = 0.68, 95% CI = 0.62–0.74) and both 8 to 10-year-old performance (M = 0.85, 95% CI = 0.80–0.91) and adult performance (M = 0.82, 95% CI = 0.76–0.87). No other pairwise comparisons between age groups reached significance.

Response latency

Our analysis of response latency to correctly detected face images revealed significant main effects of diagnosticity ($F (2, 116) = 54.23$, $p < 0.001$, partial $\eta^2 = 0.48$) and age ($F (2, 58) = 43.55$, $p < 0.001$, partial $\eta^2 = 0.60$). The main effect of diagnosticity was the result of significant differences between

the response latency to images with "high" mutual information (M = 0.77, 95% CI = 0.74–0.81) and images with both "low" mutual information (M = 0.95, 95% CI = 0.89–1.00) and "medium" mutual information (M = 0.92, 95% CI = 0.88–0.96). The comparison between "low" and "medium" mutual information did not reach significance. The main effect of age was the result of significant differences in performance between all three age groups.

Besides these main effects, we also observed a significant quadratic contrast between face diagnosticity and age ($F (2, 58) = 3.62$, $p = 0.033$, partial $\eta^2 = 0.11$). To examine the nature of this interaction, we carried out follow-up tests to determine how adjacent levels of mutual information affected performance within each group. Specifically, within each age category, we determined whether "low" and "medium" response latency differed significantly, and whether "medium" and "high" response latency differed significantly using one-tailed, paired-samples *t*-tests (a directional test is appropriate here due to the expectation that higher mutual information should lead to faster response latencies). We found that although all three age groups exhibited a significant difference between "medium" and "high" latencies (5–7-year-olds: $p < 0.001$; 8–10-year-olds, $p < 0.001$; adults, $p < 0.001$), only adults exhibited a significant difference between "low" and "medium" response latencies (5–7-year-olds, $p = 0.71$; 8–10-year-olds, $p = 0.55$; and adults, $p = 0.043$). We note that this latter result for adult observers does not survive a Bonferroni correction for multiple comparisons, but nonetheless that it provides some insight into what drives this relatively small interaction effect.

False alarms

Finally, our analysis of false alarm rates to non-faces revealed a significant effect of age group ($F (2, 58) = 4.27$, $p = 0.019$, partial $\eta^2 = 0.13$). Pairwise comparisons revealed that this main effect was the result of a lower false alarm rate in adults (M = 0.079, 95% CI = 0.03–0.12) relative to 5 to 7-year-olds (M = 0.18, 95% CI = 0.13–0.23). The false alarm rate observed in 8 to 10-year-olds (M = 0.16, 95% CI = 0.11–0.20) did not significantly differ from either age group.

Discussion

The goal of Experiment 1 was to examine how children's sensitivity to the information contained in intermediate-level visual features of faces changes during middle childhood. Starting with our results

from adult observers, we replicated and extended previous results obtained using mutual information as a criterion for defining face fragments that are more or less category-diagnostic (Harel et al., 2007; Harel et al., 2011). Specifically, face fragments with higher mutual information were more reliably categorized as faces in our go/no-go task by adults, and they were also categorized as such with a shorter response latency. This further demonstrates the utility of mutual information as a useful computational tool for identifying diagnostic face fragments.

Developmentally, our data demonstrates that both young children and older children are sensitive to mutual information contained in face fragments. In all age groups, hit rates were larger for high mutual information face fragments, and response latencies shorter. This suggests sensitivity to MI in mid-level face features may be relatively mature early in childhood, although we note that young children also tended to label low-MI fragments as faces at a much lower level than adults. This outcome differs from results examining low-level biases for face categorization in a similar go/no-go task, which indicated that biases for horizontal orientation information in face images develop relatively slowly during childhood (Balas, Schmidt, & Saville, 2015). Another important point to consider is that we did not find an interaction between age group and mutual information in our hit rate data, which could be interpreted as evidence supporting general cognitive improvement during childhood in the extraction of intermediate-level diagnostic information, that is, an improvement that perhaps does not reflect any face-specific mechanisms (Crookes & McKone, 2009). Besides the different conclusions that can be drawn from examining the hit rate data and the response latency data, we also cannot argue that these effects do or do not reflect face-specific aspects of visual development without examining how children's sensitivity to mutual information level changes for a non-face category. To address this question, we conducted Experiment 2, in which we examined the extent to which children of the same age range as in Experiment 1 are sensitive to the mutual information level contained in car fragments of varying levels of mutual information.

Experiment 2

In our second study, we examined the development of sensitivity to mutual information level in a non-face category (cars). Using the same methods as Experiment 1, we investigated how performance in car detection varied with age and mutual information level when face fragments were used as the distractor category.

Methods

Participants

We recruited a sample of 56 participants from the Fargo-Moorhead community to take part in Experiment 2. As in Experiment 1, the sample was comprised of three groups: 5 to 7-year-old children (N = 20, 13 girls, mean age = 6 years, 4 mos.), 8 to 10-year-old children (N = 20, 11 girls, mean age = 9 years, 1 month) and adults (N = 16, 10 women). Recruitment and consent procedures were identical to those reported in Experiment 1, and none of the participants included in this sample had taken part in the first experiment.

Stimuli

The stimulus set used in this experiment was drawn from the same larger set of images used to select stimuli for Experiment 1. The key difference between this stimulus set and the previously described images is that, in this case, we selected car fragments that varied according to the mutual information level provided by each image regarding the target category and chose face fragments from a fixed, intermediate, level of mutual information. Specifically, we selected 96 car fragments and 96 non-car (face) fragments for use in this task. The set of car fragments was comprised of 32 "low mutual information" car fragments, 32 "medium" fragments, and 32 "high" fragments. The set of face fragments was comprised of 96 face fragments all with a "medium" level of mutual information. Within each MI category, image size was variable: low MI car fragments had a mean width of 19 units (SD = 6.97), medium MI car fragments had a mean width of 32 units (SD =16.5), and high MI car fragments had a mean width of 32 units (SD = 21.6). "Units" here refers to the square root of image area in pixels in the raw images, but all images were scaled up by a uniform factor for presentation during the task. In terms of visual angle measurements, this led to approximate sizes of 2.5 degrees of visual angle for low MI car fragments and approximately 3.75 degrees of visual angle for medium and high MI car fragments. As we described in Experiment 1, these values must be considered in the context of the variable size and aspect ratio within each MI level and the varying position of the display and chair for child and adult participants. As such, these values of visual angle should be considered as estimates rather than as values that were strictly maintained for all participants during testing.

Procedure

We administered the same go/no-go task described in Experiment 1 to participants in this task. The only critical difference between tasks was that, in this case,

participants were instructed to press the response button only in response to car fragments and withhold responding if a face fragment was presented. Otherwise, all stimulus presentation, response collection, and display parameters were identical to those described in Experiment 1.

Results

As in Experiment 1, we examined participants' hit rate for responding correctly to car fragments, their response latency for correct responses to car fragments, and their false alarm rate in response to face fragments. We submitted the hit rate and response latency values to a 3×3 mixed-design ANOVA with age (5–7 years old, 8–10 years old, and adults) as a between-subjects factor and mutual information level (low, medium, and high) as a within-subjects factor. We analyzed false alarm rates using a one-way ANOVA with participant age as a between-subjects factor.

Hit rates

This analysis revealed only a main effect of mutual information level (F (2, 106) = 121.5, $p < 0.001$, partial $\eta^2 = 0.70$). Neither the main effect of age (F (2, 53) = 2.23, $p = 0.12$, partial $\eta^2 = 0.08$), nor the interaction between these factors (F <1) reached significance. The main effect of mutual information level was the result of lower hit rates in response to low information fragments relative to medium (mean difference = 0.28, 95% CI of the difference between means = 0.22–0.34] and high information fragments (mean difference = 0.30, 95% CI of the difference between means = 0.23–0.36). The average hit rates as a function of age and car diagnosticity are presented in Figure 4.

Response latencies

This analysis revealed main effects of age (F (2, 53) = 5.19, $p = 0.009$, partial $\eta^2 = 0.94$) and mutual information level (F (2, 106) = 32.47, $p < 0.001$, partial $\eta^2 = 0.38$). The interaction between these factors did not reach significance (F <1). The main effect of age group was the result of a significant difference between the performance of children in our youngest age group and adults (mean difference = 0.25, 95% CI of the difference between means = 0.06–0.45). Neither the difference between adults and older children, nor the difference between younger children and older children reached significance. The main effect of the mutual information level was the result of slower response latencies in response to low information fragments relative to medium (mean difference = 0.21, 95% CI of the difference between means = 0.11–0.30) and high information fragments (mean difference = 0.22, 95%
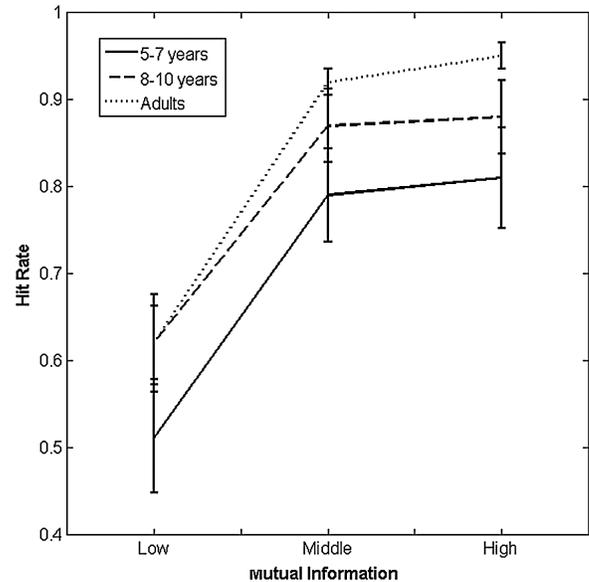

Figure 4. Average hit rate as a function of diagnosticity (mutual information level) and age for car detection. Error bars indicate +/− 1 SEM.
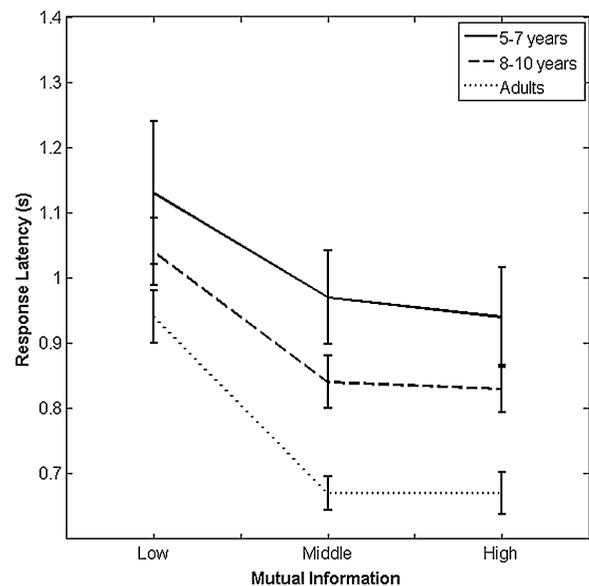

Figure 5. Average response latency for correctly labeled cars as a function of diagnosticity (mutual information) and age (5–7 years old, 8–10 years old, and adults). Error bars indicate +/− 1 SEM.

CI of the difference between means = 0.14–0.31). The average response latencies as a function of age and car diagnosticity are presented in Figure 5.

False alarms

Finally, our analysis of false alarm rates (response errors made to face fragments) revealed a marginally

significant effect of participant age (F (2, 53) = 2.71, $p = 0.076$, partial $\eta^2 = 0.093$). Although the average proportion of false alarms did appear to decrease (5–7-year-olds: M = 0.17, 8–10-year-olds: M = 0.10, and adults, M = 0.049) with age, no pairwise comparisons among age groups reached significance.

Discussion

Experiment 2 was conducted to assess the extent to which children in the middle childhood range are sensitive to the mutual information contained in a non-face category (cars) and not only to the face fragments. We found that participants across all age groups were sensitive to the diagnosticity of car fragments, as indicated by a main effect of MI level in both hit rate and response latency. Thus, children as early as the age of 5 years old can distinguish more diagnostic object features than less diagnostic object features in this task. Note that although this may suggest at first that the sensitivity to mutual information is not face-specific, the patterns of information usage in the two experiments are not identical, suggesting a more complex pattern of results. We discuss this below.

General discussion

The present study sought to determine how sensitivity to varying diagnosticity in intermediate complexity features for face and non-face categorization changes during middle childhood. We tested this question in two separate experiments, using the objective and quantifiable measure of mutual information. In both of our experiments, we have found evidence that children as young as 5 to 7 years old are sensitive to category diagnosticity in image fragments of complex objects (faces and cars). Like adults, children in both of our target age groups showed better accuracy and faster response latencies when categorizing fragments with high mutual information levels. This basic effect applied to both face and car categories, so this observation could be interpreted in terms of general visual development rather than any face-specific developmental trajectory that governs the recruitment of intermediate complexity features to support basic-level recognition. The present findings provide a first indication that children as young as 5 years of age (and perhaps even younger) are capable of extracting diagnostic features for categorization at an intermediate complexity level and not only at low- or high-level levels of visual representation.

Notably, in addition to the general effects of age and feature informativeness, we also observed critical differences in how participant age and mutual information level contributed to performance in our face and non-face detection tasks. In particular, the interaction we observed between age and mutual information level in the response latency data from Experiment 1 was not present in Experiment 2. This outcome could be interpreted to mean that for face recognition, there is ongoing development during middle childhood that leads to graded sensitivity to mutual information level in adulthood as opposed to the more nonlinear pattern of results we observed in our younger age groups. To put this more simply, middle childhood could be a period during which children are still learning how to use less diagnostic information for recognition, which means that the difference between "low" and "medium" amounts of mutual information is not meaningful to them. Adults, on the other hand, who have acquired the ability to make use of weakly diagnostic facial features, differ in that the increase in mutual information level from low to medium confers additional visual information that their visual system is capable of using for modest gains in recognition ability. This outcome does resemble recent results describing the development of low-level orientation biases for emotion recognition in faces (Balas et al., 2015) and bodies (Balas et al., 2016), in that young children in both of these studies demonstrated particularly poor abilities with suboptimal visual features. That is, although young children exhibited similar biases for orientation features in these emotion recognition tasks in terms of the direction of the effects, their absolute performance when the preferred features were not available was far worse than adults (in some cases, not different from chance). The current finding of lack of an Response time (RT) difference between low and medium levels of mutual information in children but not in adults may reflect a similar developmental trajectory whereby optimal features are identified early in development and contribute to face representations quickly, but less optimal features are uniformly "bad" at early stages of development and only gradually become useful. We suggest that this is also consistent with the poor performance of young children with regard to labeling low MI face fragments as faces in Experiment 1, which indicates that children in the 5 to 7 year age range may struggle to determine that weak indicators of the presence of a face may be useful for assigning the category label. Further, the fact that we observed this effect for faces and not for cars may mean that this is a unique feature of how face recognition develops, either because children's exposure to faces relative to cars differs across middle childhood (Sugden, Mohamed-Ali, & Moulson, 2014), or because there are differences in how cortical areas that support face and object recognition mature during this time period (Grill-Spector, Golarai, & Gabrieli, 2008).

We emphasize, however, that the critical interaction supporting this account is associated with a rather small effect size, and is only evident in response latencies and not in accuracy of performance. We also must be careful with regard to the comparison between faces and cars in this study. We opted here to match these two categories according to the range of MI values obtained for face and car fragments across a large number of images. That choice allows us to make clear statements about mutual information and category diagnosticity, but does not provide guarantees regarding low-level image properties like contrast, orientation energy, or even simpler properties like image size. To the extent that we have measured some differences across tasks when face and non-face fragments were used, we acknowledge that these differences may be attributable to image properties other than MI that have not been explicitly matched in these tasks, and would be very difficult to match while also being careful about the values of MI that we are interested in. On balance, we, therefore, suggest that both the hit rate data and the response latency data suggest that information biases favoring higher amounts of mutual information in face fragments show similar gradients from low-to-high at the earliest ages we tested, differing primarily in overall performance across levels with increasing age (younger children perform substantially worse). While the potential for face-specific development is interesting to consider in light of other recent results, the current data does not make a strong case for this conclusion at this moment.

An important avenue of research to consider to help clarify the extent to which the current findings may reflect face-specific effects is the measurement of neural sensitivity to mid-level visual features during childhood. Adult observers' ERP responses are sensitive to mutual information in face and car fragments (Harel et al., 2007; Harel et al., 2011), but little is known about how either low-level or mid-level information biases for face recognition develop. An investigation of how ERP components like the P100 and the N170/N290 respond to varying levels of mutual information in face and non-face fragments would be an important step toward linking the computational and behavioral evidence for fragments of intermediate complexity as a mid-level representation of face appearance to real neural outcomes. Recent results examining the sensitivity of these components to the low-level visual information in faces (horizontal versus vertical orientation energy) has revealed that face-sensitive ERP components change their tuning to orientation sub-bands during middle childhood in a category-selective way (Balas et al., 2017), supporting behavioral work indicating similar outcomes for response latencies (Balas et al., 2015). Applying these same methods to intermediate complexity features in faces and non-faces would, thus, provide an important complementary look at the development of visual recognition during this age range.

In summary, the present findings provide important and novel data demonstrating sensitivity to mid-level features for face and object recognition emerges as early as 5 years old of age. Complementing prior results examining children's use of visual information at low-level and high-level stages of face procession, the current study demonstrates near-adult utilization of information contained in intermediate complexity level features. Our findings are significant not only in establishing the feasibility of intermediate-level features as a "third way" for object representations, but also in providing explicit, objective, and quantitative means to describe what it means for a mid-level feature to be diagnostic. This will allow future investigations of how children develop "vocabularies" of intermediate complexity visual features to form different object categories, while avoiding previous dichotomies between low and high accounts of face and object categorization. Deeper insights into how the content of face representations changes during development will depend on identifying and testing quantitative models like the current one, and our results offer a compelling example of how such computational models can be adapted for use with developmental populations.

Keywords: face detection, visual development, object recognition

References

Balas, B., Auen, A., Saville, A., & Schmidt, J. (2016). Body emotion recognition disproportionately depends on vertical orientations during childhood. *International Journal of Behavioral Development, 42*, 278–283.

Balas, B., Schmidt, J., & Saville, A. (2015). A face detection bias for horizontal orientations develops during middle childhood. *Frontiers in Developmental Psychology, 6*, 772.

Balas, B., Huynh, C., Saville, A., & Schmidt, J. (2015). Orientation biases for facial emotion

recognition in early childhood and adulthood. *Journal of Experimental Child Psychology, 140*, 71–83.

Balas, B., van Lamsweerde, A. E., Saville, A., & Schmidt, J. (2017). School-age children's neural sensitivity to horizontal orientation energy in faces. *Developmental Psychobiology, 59*, 899–909.

Bindemann, M., & Burton, A.M. (2009). The role of color in human face detection. *Cognitive Science, 33*, 1144–1156.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Carbon, C-C., Gruter, M., & Gruter, T. (2013) Age-dependent face detection and ace categorization performance. *PLoS One, 8*, e79184.

Chien, S. H.-L. (2011). No more top-heavy bias: infants and adults prefer upright faces but not top-heavy geometric or face-like patterns. *Journal of Vision, 11*, 13.

Chung, M-S., & Thomson, D. M. (1995). Development of face recognition. *British Journal of Psychology, 86*, 55–87.

Crookes, K., & McKone, E. (2009). Early maturity of face recognition: no childhood development of holistic processing, novel face encoding, or face-space. *Cognition, 111*, 219–247.

Dakin, S. C., & Watt, R. J. (2009). Biological "bar codes" in human faces. *Journal of Vision, 9*(4):2, 1–10.

De Heering, A., Houthuys, S., & Rossion, B. (2007). Holistic face processing is mature at 4 years of age: evidence from the composite-face effect. *Journal of Experiment Child Psychology, 96*, 57–70.

De Heering, A., Goffaux, V., Dollion, N., Godard, Or., Durand, K., & Baudouin, J.-Y. (2016). Three-month-old infants' sensitivity to horizontal information within faces. *Developmental Psychobiology, 58*, 536–542.

De Heering, A., & Rossion, B. (2015). Rapid categorization of natural face images in the infant right hemisphere. *eLife. 4*, e06564.

Donnelly, N., & Hadwin, J. A. (2003). Children's perception of the Thatcher illusion: evidence for development in configural face processing. *Visual Cognition, 10*(8), 1001–1017.

Ewing, L., Pellicano, E., King, H., Lennuyeux-Conene, L., Farran, E.K., Karmiloff-Smith, A., . . . Smith, M.L. (2018). Atypical information-use in children with autism spectrum disorder during judgments of child and adult face identity. *Developmental Neuropsychology, 43*, 370–384.

Gao, X., & Maurer, D. (2011). A comparison of spatial frequency tuning for the recognition of

facial identity and facial expressions in adults and children. *Vision Research, 51*, 508–519. doi: 10.1016/j.visres.2011.01.011

Goffaux, V., Poncin, A., & Schiltz, C. (2015). Selectivity of face perception to horizontal information over lifespan (from 6 to 74 year old). *PLoS One, 10*, e0138812.

Gosselin, F., & Schyns, P.G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41*, 2261–271.

Grill-Spector, K., Golarai, G., & Gabrieli, J. (2008). Developmental neuroimaging of the human ventral visual cortex. *Trends in Cognitive Sciences, 12*, 152–162.

Guillon, Q., Roge, B., Afzali, M.H., Baduel, S., Kruck, J., & Hadjikhani, N. (2016) Intact perception but abnormal orientation towards face-like objects in young children with ASD. *Scientific Reports, 6*, 22119.

Harel, A., Ullman, S., Epshtein, B., & Bentin, S. (2007). Mutual information of image fragments predicts categorization in humans: Electrophysiological and behavioral evidence. *Vision Research, 47*, 2010–2020.

Harel, A., Ullman, S., Harari, D., & Bentin, S. (2011). Basic-level categorization of intermediate complexity fragments reveals top-down effects of expertise in visual perception. *Journal of Vision, 11*(8):18, 1–13.

Humphreys, K., Gosselin, F., Schyns, P.G., & Johnson, M.H. (2006). Using "Bubbles" with babies: A new technique for investigating the informational basis of infant perception. *Infant Behavior & Development, 29*, 471–475.

Kuefner, D., de Heering, A., Jacques, C., Palmero-Soler, E., & Rossion, B. (2010). Early visually evoked electrophysiological responses over the human brain (P1, N170) show stable patterns of face-sensitivity from 4 years to adulthood. *Frontiers in Human Neuroscience, 6*(3), 67.

Leonard, H.C., Karmiloff-Smith, A., & Johnson, M.H. (2010). The development of spatial frequency biases in face recognition. *Journal of Experimental Child Psychology, 106*, 193–207.

Lerner, Y., Epshtein, B., Ullman, S., & Malach, R. (2008). Class information predicts activation by object fragments in human object areas. *Journal of Cognitive Neuroscience, 20*, 1189–1206.

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

Mckone, E., Crookes, K., & Kanwisher, N. (2009). The cognitive and neural development of face

recognition in humans. In M. S. Gazzaniga, E. Bizzi, L. M. Chalupa, S. T. Grafton, T. F. Heatherton, & C. Koch, . . . B. A. Wandell (Eds.), *The Cognitive Neurosciences* (pp. 467–482). Cambridge, MA, USA: MIT Press.

Mondloch, C.J., Dobson, K.S., Parsons, J., & Maurer, D. (2004). Why 8-year-olds cannot tell the difference between Steve Martin and Paul Newman: factors contributing to the slow development of sensitivity to the spacing of facial features. *Journal of Experimental Child Psychology, 89*, 159–181.

Mondloch, C.J., Pathman, T., Le Grand, R., Maurer, D., & de Schonen, S. (2007). The composite face effect in six-year-old children: Evidence of adultlike holistic face processing. *Visual Cognition, 15*, 564–577.

Nasanen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research, 39*, 3824–3833.

Nelson, C.A. (2001). The development and neural bases of face recognition. *Infant and Child Development, 10*, 3–18.

Obermeyer, S., Kolling, T., Schaich, A., & Knopf, M. (2012). Differences between old and young adults' ability to recognize human faces underlie processing of horizontal information. *Frontiers in Aging Neuroscience, 4*, 3.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*, 437–42.

Peterson, M.A. (1996). Overlapping partial configurations in object memory. In M.A. Peterson, & G. Rhodes (Eds.), *Perception of Faces, Objects, and Scenes*. Oxford University Press, NY.

Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science, 14*, 618–622.

Robertson, D.J., Jenkins, R., & Burton, A.M. (2017). Face detection dissociates from face identification. *Visual Cognition, 25*, 740–748.

Ryan, C., Stafford, M., & King, R.J. (2016). Seeing the man in the moon: do children with autism perceive pareidolic faces? A pilot study. *Journal of Autism and Developmental Disorders, 46*, 3838–3843.

Sangrigoli, S., & de Schonen, S. (2004). Effect of visual experience on face processing: a developmental study of inversion and nonnative effects. *Developmental Science, 7*, 74–87.

Schwarzer, G. (2000). Development of face processing: the effect of face inversion. *Child Development, 71*, 391–401.

Simion, F., Valenza, E., Macchi Cassia, V., Turati, C., & Umilta, C. (2002). Newborns' preference for up-down asymmetrical configurations. *Developmental Science, 5*, 427–434.

Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology, 56*, 249–261.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19*, 109–139.

Taylor, M.J., Batty, M., & Itier, R.J. (2004). The faces of development: a review of early face processing over childhood. *Journal of Cognitive Neuroscience, 16*, 1426–1442.

Ullman, S. (2007). Object recognition and segmentation by a fragment-based hierarchy. *Trends in Cognitive Science, 11*, 58–64.

Ullman, S., & Sali, E. (2000). Object classification using fragment-based representation. In S.-W. Lee, H.H. Bulthoff, & T. Poggio (Eds.), *Biologically motivated computer vision, 1811*, 73–87. Springer-Verlag, Berlin.

Ullman, S., Vidal-Naquet, M., & Sali, E (2002). Visual features of intermediate complexity and their use in classification. *Nature Neuroscience 5*, 682–687.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J.W., Bub, D.N., . . . Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36*, 122–135.