Macalester College
DigitalCommons@Macalester College

Psychology Honors Projects | Psychology Department

4-18-2012

Facial Expression Processing is Holistic or Feature-Based Depending on Stimulus Format: Evidence from the Composite Face Illusion and Gaze-Contingent Stimulus Presentations

Emily R. Prazak
Macalester College, prazak.emily@gmail.com

Facial Expression Processing is Holistic or Feature-Based Depending on Stimulus Format: Evidence from the Composite Face Illusion and Gaze-Contingent Stimulus Presentations

Emily Prazak

Advised by Professor E. Darcy Burgund, Psychology Department

Macalester College, Saint Paul MN

Submitted April 18th, 2012

Acknowledgements

This project would not have happened without a number of people to whom I owe an enormous round of applause. First of all, to the best family in the world for their omnipresent support; to my housemates Jo, Hannah, and Will for putting up with my crazy; to my eye-tracking gurus Paul Rebman and Micah Mumper for programming my eye-tracking experiment and answering my incessant e-mails; to Melanie Rivera for her countless hours of data collection; to Professor Brooke Lea and Professor Ronald Brisbois for sitting on my honors committee and making my defense such a delightful experience; and most importantly to my goddess of cognitive neuroscience and advisor, Professor Darcy Burgund. By no stretch of the imagination would this project have been possible without you.

Abstract

Controversy exists over whether facial expression recognition is a holistic or feature-based process. The present research explored whether stimulus format (photographic vs. schematic) affects the type of processing used. In a composite/noncomposite expression recognition task, holistic processing was observed for photographic stimuli and feature-based processing was observed for schematic stimuli. Moreover, holistic processing in the photographic condition increased when more than one individual was presented. Results suggest that facial expression processing is holistic under natural viewing conditions and provide a potential resolution to the previous controversy. Such findings may be corroborated by an ongoing follow-up study using gaze-contingent stimulus presentations.

Facial Expression Processing is Holistic or Feature-Based Depending on Stimulus Format: Evidence from the Composite Face Illusion and Gaze-Contingent Stimulus Presentations

Extensive research has shown that facial identity recognition is a holistic process where the viewer observes the face as a whole instead of focusing on individual facial features (Farah, Wilson, Drain, & Tanaka, 1998; Maurer, Le Grand, & Mondlock, 2002; Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987). However, the debate is still ongoing as to whether facial expression recognition is also a holistic process (i.e., combining all the features as a whole to recognize emotions), or a feature-based process (i.e., focusing on individual features such as the eyes or mouth). There are many discrepancies in the literature, with research supporting both sides of the argument.

Calder, Keane, Young, and Dean (2000) obtained support for holistic processing of facial expression recognition through use of the composite face illusion. In this paradigm, composite stimuli are made by combining the top half of one face with the bottom half of another face (see Figure 1C), and noncomposite stimuli are made by taking the two halves of a composite face and separating them (see Figure 1D; Young, Hellawell, & Hay, 1987). Separating the two halves disrupts holistic processing by altering the first order feature relations of the face (i.e., the eyes are no longer directly above the nose and mouth). Calder et al. (2000) had participants identify expressions displayed in composite and noncomposite faces whose top and bottom halves displayed different emotions. Participants were significantly slower at identifying top-half and bottom-half expressions of composite faces compared to noncomposite faces. This pattern of results suggests facial expression recognition is a holistic process because the different expression in the bottom half hindered identification of the expression in the top half and vice versa for the composite condition only. In the noncomposite condition, subjects were able to focus on each half individually without the other half interfering because the two halves were separate and therefore impossible to process holistically. If facial expression recognition were a feature-based process, there would have been no difference between the composite and the noncomposite conditions.

Support for holistic processing has also been found through use of inversion (McKelvie, 1995). Inversion has been shown to hinder holistic processing (Tanaka & Farah, 1993; Valentine, 1988) because flipping a face upside down also alters the first-order relational features of the face. The results of McKelvie (1995) showed that inversion significantly reduced accuracy for all expressions except happiness. Such findings suggest that overall, holistic processing is used in the identification of facial expressions because the disruption of holistic processing resulted in reduced facial expression recognition.

Calder and Jansen (2005) also used inversion to disrupt holistic processing of composite faces. Participants viewed two composite faces side-by-side where the bottom halves of the faces were always different, but the top halves were either the same or different. The task involved making a simple same/different decision for the top halves. Results revealed that same-top composites were processed faster and more accurately when inverted, showing that when holistic processing was disrupted by inversion, participants were able to focus on individual features to make the decision. However, decisions for different-top faces were faster and more accurate with upright stimuli presentation. When the faces were upright, participants would holistically process faces that had both different bottoms and different tops, thus the difference would be very easy to see in the upright condition.

However, not all studies find support for holistic processing in facial expression recognition. A study by Ohman, Lundqvist, and Esteves (2001) looked at the result of inversion on the threat superiority effect, where threatening faces are perceived faster and more accurately than non-threatening faces. Using a visual search paradigm, the researchers had subjects view matrices of faces both upright and inverted. Subjects were instructed to decide whether faces were all the same or if one face was different. Regardless of orientation, subjects were faster at finding the threatening faces than the non-threatening faces, suggesting that the threat superiority effect is mediated by a feature-based strategy of facial expression recognition, rather than a holistic strategy. Similar results were found by Lipp, Price, and Tellegen (2009).

Ellison and Massaro (1997) found support for feature-based processing using computer simulated faces in which the researchers manipulated only the eyebrows and the mouth. The eyebrows varied in degrees from curved upward (prototypically happy) to straight (prototypically angry), and the mouth varied in degrees from having the corners curved all the way up (prototypically happy) to straight (prototypically angry). The varying degrees of these features were combined in the faces across a range so that some faces had the same emotion on top and bottom, while others had different emotions on the top and bottom. Participants were instructed to identify the face as happy or sad.

Results showed that participants tended to identify the expressions based on the feature that was closest to the extreme prototype. For example, if the eyebrows were more prototypically happy than the mouth was angry, the face would be identified as happy. Subjects also rated the faces on a scale of 1 (completely angry) to 9 (completely happy) with 5 being neutral. Regardless of whether or not the two features showed the same or different emotions, the average ratings for happiness increased as the eyebrows elevated and arched, and as the corners of the mouth curled up, while the average ratings for anger increased as the eyebrows straightened, and the corner of the mouth curled down. So even when two features were sending different signals, subjects were still able to focus on the most informative feature, suggesting that they were utilizing a feature-based strategy

Further support for feature-based processing comes from the finding that subjects are capable of identifying complex facial expressions from viewing the eyes alone (Baron-Cohen, Wheelwright, & Jolliffe, 1997). Additionally, the threat superiority effect is found among participants viewing the eyes alone (Fox & Damjanovic, 2006), and from manipulation of only the eyebrows (Tipples, Atkinson, & Young, 2002). Furthermore, Hall, Hutton, and Morgan (2010) made a connection between female superiority in facial expression recognition and a higher amount of time spent fixating on the eyes, which suggests increased focus on the eyes improves facial expression recognition. Taken as a whole, these studies show that viewing an entire face is not necessary for accurate facial expression recognition. However, these studies only suggest that identification through feature-based processing is possible—they do not prove that feature-based processing is the primary method for facial expression recognition.

In sum, evidence supports both holistic processing and feature-based processing in facial expression recognition and a resolution to the discrepancy has not been proposed. One possible resolution is that the discrepancy between studies supporting holistic processing and those supporting feature-based processing is due to differences in stimulus format. In particular, the majority of studies supporting feature-based processing use either schematic faces (e.g., Lipp et al., 2009; Ohman et al., 2001) or computer-simulated faces (e.g., Ellison & Massaro, 1997), while the studies that support holistic processing use photographs of real human faces (e.g., Calder & Jansen, 2005; Calder et al., 2000). Schematic and computer-simulated stimuli may be processed using a feature-based strategy because they do not display the global changes that occur in a human face during different facial expressions. For example, in a happy expression, the facial muscles pull the cheeks up towards to the eyes. This change is easily portrayed in photographic stimuli but is more difficult to show in a schematic face. If feature-based processing is utilized for schematic faces, using an identical task with both schematic faces and photographic faces should yield different results for the two types. More specifically, the results from the schematic faces should support feature-based processing and results from the photographic faces should support holistic processing. The present study tests this hypothesis.

In addition, the present study examines a related explanation. The variability within facial expressions is higher in photographic stimuli than schematic stimuli because multiple individuals may be used with photographic stimuli whereas only one identity is typically used with schematic faces. With only a single identity to process, general face processing would be easier and the viewer could potentially develop a more efficient feature-based strategy for expression identification. Because of this possibility, two photographic conditions are called for: one in which only one model is used (single identity condition), and one in which multiple models are used (multiple identity condition).

In Experiment 1, subjects completed the composite paradigm identification task utilized by Calder et al. (2000) with either photographic faces (from multiple identities or a single identity) or schematic faces. Because Calder et al. (2000) used this task to show holistic processing for photographic stimuli, an inability to replicate Calder et al.'s findings with schematic faces would support the possibility that the discrepancies in the literature are due to stimulus format. An additional inability to replicate with photographic stimuli from only one model would suggest the discrepancies are due to the identity variation issue. Subjects viewed composite and noncomposite faces and were asked to identify the expression in the top half of the face. If stimulus format is the cause of the discrepancy in the literature, subjects should be more accurate at identifying the noncomposite faces compared to the composite faces in the photographic conditions, but not in the schematic condition. In the schematic condition there should be no difference between the composite and noncomposite faces because feature-based processing would be functional for both composite and noncomposite faces. In addition, if the discrepancy is caused by the inability to vary the identity of schematic stimuli, responses to noncomposite faces should be faster than responses to composite faces in the multiple identity photographic condition only. In the single-identity photographic condition and the schematic condition, responses to composite and noncomposite stimuli should not differ.

Experiment 1

Method

Participants. Ninety-nine participants (72 females) were recruited from the Macalester College Psychology Participant Pool. Two participants were excluded due to at-chance performance, resulting in a total of 97 participants. Thirty-four of these were in the photographic multiple identity condition, 32 were in the photographic single identity condition, and 31 were in the schematic condition.

Design. A 2 x 3 factorial design was used in which stimulus type (composite vs. noncomposite) was examined as a within-subject variable and stimulus format (single-identity photographic vs. multiple-identity photographic vs. schematic) was manipulated as a between-subject variable.

Materials. Photographic stimuli were created using faces from the Radboud Faces Database (Lagner et al., 2010). Faces of five male models (models 7, 23, 28, 30, and 71) displaying five basic emotions (anger, fear, happiness, sadness, and surprise) were gray-scaled and cropped to include only the face (see Figures 1A and 1B). Adobe® Photoshop Elements 10 was used to crop the top halves and bottom halves and combine the halves of different expressions to make the composite stimuli (see Figure 1C). All 20 possible expression combinations were used. For noncomposite stimuli, the bottom half of the face was shifted to the right or to the left so that the nose in the bottom half lined up with the border of the face on the top half (see Figure 1D). The direction of the shift (right or left) was counterbalanced across stimuli.

Schematic stimuli were modeled after the schematic faces used by Ohman et al. (2001) (see Figures 2A and 2B). As described above for the photographic stimuli, composite faces were made by combining the top half of one expression and the bottom half of a different expression (see Figure 2C). Noncomposite stimuli were made by shifting the bottom half to the right or left (see Figure 2D).

Stimuli were presented in the middle of the screen against a white background and subtended a visual angle of approximately 7.1º x 6.9º. Participants placed their chins in a chin-rest positioned 24 inches from the computer screen. Response recording and stimulus presentation were controlled using PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) on a Macintosh computer.

Procedure. Participants were randomly assigned to one of three conditions. One group of participants viewed schematic stimuli, a second group viewed photographic stimuli created from images of four different models, and the remaining group viewed photographic stimuli created from only one model (the particular model was counterbalanced across participants). As the three conditions were identical except for stimulus format and the number of identities, only the multiple-identity photographic condition is described in detail.

The experiment began with a round of practice trials in which participants viewed 20 photographic faces displaying prototypical expressions. Their task was to identify the expression by pressing one of five buttons assigned to the emotions (i.e., happiness, sadness, anger, fear, and surprise). Participants were shown a fixation cross for 500 ms, followed by a blank screen for 500 ms, followed by the stimuli. No time restrictions were set; the participant's response initiated the next trial. A second round of practice trials involved only the top half of the prototypical expressions. Participants again identified the expression by pressing the corresponding button.

Following these practice trials, participants viewed 6 randomly selected composite and noncomposite faces to familiarize them with the two types of stimuli. Their instructions were to identify the expression in the top half of the face. After the composite/noncomposite practice trials, participants viewed 160 composite and noncomposite faces in random order. Again, they were instructed to identify the expression in the top half of the face and encouraged to be as fast and accurate as possible. Response times and error rates were recorded.

The designs of the single-identity photographic and schematic conditions were the same as described above. However, one important difference should be pointed out. Because there were fewer possible stimuli in the single-identity photographic and schematic conditions than in the multiple-identity photographic condition, participants in the two former conditions viewed each individual stimulus four times during the experiment to ensure that an equal number of composite and noncomposite trials were used for every condition.

Results

Response times. Response times were analyzed using a two-way repeated-measures analysis of variance (ANOVA) in which stimulus type (composite vs. noncomposite) was a within-subjects variable, and stimulus format (multiple-identity photographic vs. single-identity photographic vs. schematic) was a between-subjects variable. Only response times for correct responses were analyzed, and outliers were pruned by excluding (noniteratively) response times that were 2.5 standard deviations above the mean. Results revealed a main effect of stimulus type in which noncomposite responses (M = 1304 ms) were faster than composite responses (M = 1380 ms), $F(1, 94) = 16.22$, $p < .001$. More important, results revealed a two-way interaction of stimulus type by stimulus format, $F(2, 94) = 5.71$, $p = .005$ (see Figure 3). Post-hoc paired t-tests comparing composite and noncomposite stimuli in the three conditions showed that noncomposite faces were processed faster (M = 1320 ms) than composite faces (M = 1473 ms) in the multiple-identity photographic condition, $t(33) = 4.43$, $p < .001$. The same pattern was seen in the single-identity photographic condition, where noncomposite faces were processed faster (M = 1274) than composite faces (M = 1337), $t(31) = 2.81$, $p = .009$. However, in the schematic condition noncomposite and composite faces were processed equally (Ms = 1323 ms vs. 1319 ms, respectively), $t < 1$.

To compare the magnitude of the stimulus type effect between conditions directly, three additional 2 x 2 ANOVAs were conducted. These ANOVAs revealed that the stimulus type effect was larger in the multiple-identity photographic condition compared to the schematic condition, $F(1, 63) = 8.95$, $p = .004$, for the two-way interaction of stimulus type by format (multiple-identity photographic vs. schematic), and compared to the single-identity photographic condition, $F(1, 64) = 4.69$, $p = .034$, for the two-way interaction of stimulus type by format (multiple-identity photographic vs. single-identity photographic). In contrast, although the effect of stimulus type was significant in the single-identity photographic condition but not in the schematic condition, the two-way interaction comparing the effects in each condition did not reach significance, $F(1, 61) = 1.95$, $p = .168$.

Error rates. Error rates were also analyzed using a two-way repeated-measures ANOVA. Similar to the response times, composite faces had higher error rates than noncomposite faces (*M*s = 36% vs. 30%, respectively), $F(1, 94) = 60.10$, $p < .001$, for the main effect of stimulus type. In contrast, the two-way interaction of stimulus type by format was not significant, $F < 1$, however the trend was in the same direction as that seen in the response times and suggest there was no speed-accuracy trade off (see Figure 4). In addition, the main effect of stimulus format was significant with the greatest errors made in the schematic condition (41%), followed by the multiple-identity photographic condition (33%), followed by the single-identity photographic condition (23%), $F(2, 94) = 13.96$, $p < .001$.

Discussion

Experiment 1 sought to examine whether the discrepancy in the literature may be due at least in part to a difference in stimuli. Overall, results from the composite/noncomposite identification task support this hypothesis. Subjects were faster at identifying the top half of the noncomposite faces compared to the composite faces in both photographic conditions, but were equally fast at processing noncomposite and composite faces in the schematic condition. This corroborates the results of Calder et al. (2000). Such findings support holistic processing for photographic stimuli because noncomposite faces cannot be processed holistically, which allows subjects to focus on the top half without having the expression in the bottom half interfere.

Additionally, the results support feature-based processing for schematic faces because subjects were able to focus on the top half alone in both composite and noncomposite stimuli in the schematic condition. However, the significant two-way interaction between the two photographic conditions and the non-significant two-way interaction between the single-identity photographic and schematic conditions suggest that identity variation may also play a role in the discrepancy. Taken together, these findings suggest that humans utilize holistic processing under natural viewing conditions (i.e., viewing real faces of multiple individuals), but are able to develop feature-based strategies under experimental conditions when viewing faces from only one individual or schematic faces. This idea is explored further in Experiment 2.

Experiment 2

While the composite face illusion allowed us to measure the types of processing used under various stimuli conditions, eye-tracking technology permits us to directly manipulate which mode of processing a participant must employ. The majority of studies using eye-tracking technology to investigate face processing have used it to quantify eye fixations and scan paths, and most have only examined facial identity recognition (e.g., Althoff & Cohen, 1999; Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006;

Hedwig & Alpers, 2011; William & Henderson, 2007), although some eye tracking studies have focused on facial expression recognition as well (Adolphs, Gosselin, Buchannan, Tranel, Schyns, & Damasio, 2005; Hedwig & Alpers, 2011). A few studies have even used eye-tracking to differentiate between the eye movements involved for holistic and feature-based processing of facial identity (Barton et al., 2006; De Xivry, Ramon, Lefevre, & Rossion, 2008; Hsiao & Cottrell, 2008; Schwarzer, Huber, & Dummler, 2005). However, only one study has used eye-tracking to directly manipulate face processing.

Van Belle, De Graef, Verfaillie, Rossion, and Lefevre (2010) utilized eye-tracking technology to control the type of processing used to identity faces through gaze-contingent stimulus presentation. In this framework, participants view faces in three ways: full-view where there are no viewing restrictions, window-view where only the fixated feature can be seen, and mask-view where the fixated feature is hidden. In the window-view participants are forced to use feature-based processing because they can only see one feature at a time, while in the mask-view they are encouraged to use holistic processing because they can gain more information from the surround of their fixation. Van Belle et al. (2010) used gaze-contingent stimulus presentation with upright and inverted faces in an identity recognition task and found that there were significant inversion effects for the full-view and mask-view conditions, but not for the window-condition. In addition, the mask-view had a significantly greater inversion effect than the full-view condition. Such findings support the hypothesis that inversion disrupts holistic processing because participants were equally capable of identifying upright and inverted faces when forced to use a feature-based strategy. As this is the only study to date that has used gaze-contingent stimulus presentation to manipulate face processing, it is unknown whether facial expression recognition would show the same pattern of results.

The present study remedies this by applying the gaze-contingent stimulus presentation used by Van Belle et al. (2010) to a facial expression recognition task. In addition, this experiment attempts to replicate the effects of stimulus format observed in Experiment 1. As in Experiment 1, participants viewed schematic faces or photographic faces (from multiple models or only one model) and were instructed to identify the expression. Stimuli were viewed through three different viewing conditions: a full-view condition, window-view condition, and a mask-condition. If holistic processing is used for photographic faces but not schematic faces, there should be significant inversion effects for the full-view and mask-view conditions, but not the window-view condition for the photographic stimuli. In the schematic stimuli there should be no significant inversion effects for any of the viewing conditions. If identity variation also plays a role in the discrepancy, the inversion effects should be greater in the multiple-identity photographic condition than the single-identity photographic condition.

Method

Participants. 60 participants were recruited from the Macalester College Psychology Participant Pool. All participants had normal or correct-to-normal visual acuity.

Design. A 2 x 3 x 3 factorial design was used in which stimulus orientation (upright vs. inverted) and viewing condition (full-view vs. window-view vs. mask-view) were examined as within-subject variables and stimulus format (single-identity photographic vs. multiple-identity photographic vs. schematic) was manipulated as a between-subject variable.

Materials. As in Experiment 1, photographic stimuli were created using faces from the Radboud Face Database (Lagner et al., 2010). Faces of five male models (models 7, 23, 28, 30, and 71) displaying five basic emotions (anger, fear, happiness, sadness, and surprise) were gray-scaled and cropped to remove all external features (see Figure 5A). Schematic stimuli were altered slightly from Experiment 1 by removing the background and making the face color gray so that the face could be found easily in the window-view condition (see Figure 5B). Stimuli were presented upright and inverted (see Figures 5C and 5D) through three different viewing conditions. One viewing condition was the full-view condition, where no viewing restrictions were in place (see Figures 6A and 6D), another was the window-view condition where only the fixated feature was viewable (see Figures 6B and 6E), and the third was the mask-view condition where the fixated feature was hidden from view (see Figures 6C and 6F).

Stimuli were presented against a white background in the middle of a 10 by 13.5 inch Accusync 900 computer screen. The stimuli subtended a visual angle of approximately 14.4º vertically x 10.5º horizontally and the window and mask subtended a visual angle of 2.64º vertically x 4.75º horizontally. Participants placed their chins in a movement-restricting head-rest positioned 21 inches from the computer screen. Response recording, stimulus presentation, and eye-movement tracking were controlled using SR Research Eyelink software with an SR Research Eyelink 1000 remote eye tracker set at a sampling rate of 1000 Hz and with a gaze position average error smaller than 0.5º. Participants responded by pressing the “s”, “d”, “f”, “j”, and “k” buttons on a standard keyboard, which were labeled with “anger,” “fear,” “happiness,” “sadness,” and “surprise” respectively.

Procedure. Participants were randomly assigned to one of three conditions. One group of participants viewed schematic stimuli, a second group viewed photographic stimuli created from images of five different models, and the remaining group viewed photographic stimuli created from only one model (the model was varied between participants).

The experiment began with two sets of practice trials. For the first set, participants viewed 25 trials of expression names (e.g., anger, happiness, fear, etc.) on the computer screen and were instructed to press the key with the corresponding label. The second set of practice trials included 25 trials of expressive faces and participants were instructed to identify the expression by pressing the corresponding key. Participants were given feedback for the practice trials. Calibration took place following the practice trials.

The experimental identification task was subdivided into 6 blocks of 30 faces. In every block, each of the 5 expressions was displayed 6 times (twice per viewing condition) in a random order. Stimulus orientation was varied in a blocked design where the first three blocks were either all upright or all inverted and the second three blocks were either all upright or all inverted. The order of stimulus orientation was counterbalanced between participants. Participants were given an optional break after the first three blocks.

A trial time-course is shown in Figure 7. Every trial began with a central drift correction to solve for any minor head movements. Following the drift correction, an average face was presented in the center of the screen. For the photographic condition, the average face was created from all five models displaying all five expressions using PsychoMorph software (Tiddemen, Burt, & Perret, 2001). For the schematic condition, the average face was simply the neutral expression schematic. To the left of the average face was a fixation cross. Participants were instructed to fixate the fixation cross. Upon doing so, the cross would disappear and participants could then fixate the average face. Once the participant's gaze reached the average face, it immediately changed to the expressive face to be processed. Participants were encouraged to identify the expression as fast and accurately as possible and were not given feedback. No time restrictions were set; the participant's response initiated the next trial. Response times, error rates, and eye-movements were recorded.

Results

Experiment 2 is currently underway and results are pending. However, potential results can be considered. If the hypothesis that holistic processing is used for photographic faces and feature-based processing is used for schematic faces is accurate, we should observe significant inversion effects in the full-view and mask-view conditions, but not in the window-view condition for photographic stimuli. Specifically, subjects should identify expressions faster and more accurately for upright faces compared to inverted faces in the full-view and mask-view, but should identify upright and inverted expressions equally in the window-view condition. In contrast, for the schematic condition there should be no significant inversion effects in any of the viewing conditions (see Figures 8 and 9).

Conversely, it is also possible that the results will not support the hypothesis that the discrepancy is due to stimulus format. If this is the case, results should not differ between the photographic and schematic conditions. Because results from Experiment 1 support holistic processing of facial expressions, this scenario would most likely be reflected in significant inversion effects for the full-view and mask-view conditions in all three format conditions (see Figures 10 and 11). However, it is also possible that results of Experiment 2 will not corroborate the general findings of Experiment 1. Specifically, the results from Experiment 2 may support feature-based processing of facial expression. If feature-based processing is used for facial expression recognition, no significant inversion effects should be seen for any viewing condition in any format. Furthermore, because feature-based processing would be most hindered in the mask-view condition, the highest reaction times and error rates would be seen in this condition for all three formats (see Figures 12 and 13).

General Discussion

Previous research concerning facial expression processing has yielded mixed results, with some studies supporting holistic processing and others supporting feature-based processing. A notable trend is present in the literature where studies supporting holistic processing generally measure expression identification using photographs of real faces, while studies supporting feature-based processing tend to utilize schematic faces.

The current study sought to examine whether the discrepancy in the literature is due at least in part to this difference in stimuli.

Results from the composite/noncomposite identification task in Experiment 1 support this hypothesis. Participants were faster at identifying the top half expression of noncomposite faces compared to composite faces in both photographic conditions, but were equally fast at identifying the expressions of noncomposite and composite faces in the schematic condition. Such findings support holistic processing for photographic stimuli because shifting the top and bottom halves apart to make noncomposite faces disrupts holistic processing, allowing participants to focus on the top half without having the expression in the bottom half interfere with identification. In contrast, these results suggest feature-based processing for schematic faces because participants were able to focus on the top half independently, regardless of whether the two halves were separated or combined into a whole face.

Furthermore, the smaller composite effect in the photographic single-identity condition is informative. The significant interaction between the two photographic conditions combined with the nonsignificant interaction between the single-identity and schematic conditions suggests that identity variation may also affect expression processing. When stimuli are created from images of multiple models there is more variability within expressions. For example, models A and B may both be displaying angry expressions, but due to their unique bone and muscle structures, their angry expressions will not be identical. Although varying a schematic face's "identity" is feasible through alteration of the sizes and distances between features, this has not been done in previous research. Thus, studies that utilize schematic faces lack the expression variability achieved by studies using photographic faces. This relative absence of variation may allow participants to develop a more efficient feature-based strategy for identifying facial expressions in schematic faces.

In addition to providing an explanation for the discrepancy in the literature, the present findings provide insight to facial expression recognition more generally. In this study, the photographic multiple-identity condition was the closest approximation to natural viewing conditions. Therefore, this study suggests that humans utilize holistic processing of facial expression under natural viewing conditions (i.e., seeing real faces of multiple individuals). This must be taken into account regarding future research designs. Because photographic and schematic stimuli have been shown to yield different results, the best course of action would be to use photographic stimuli in an effort to most closely approximate actual facial expression processing. Furthermore, images of multiple models should be used since the results of this experiment suggest photographic stimuli from only one identity may enlist feature-based processing to some extent.

Gaze-contingent stimulus presentation presents an opportunity to strengthen this argument by yielding similar results through direct manipulation of facial expression processing. Such research is currently underway as described in Experiment 2. However, there are other face processing paradigms that may be used to further our claim as well. For example, previous research observing the threat superiority effect with both upright and inverted faces (Lipp et al., 2009; Ohman et al., 2001) has been used to support feature-based processing for facial expressions; however, these studies utilized schematic stimuli. Running their visual search task with both schematic and photographic stimuli may yield results similar to those found in the current study. The threat superiority effect may survive inversion in the schematic condition, but not in the photographic condition. Such results would support the argument that photographic stimuli enlist holistic processing while schematic stimuli utilize feature-based processing.

In conclusion, the present study attempted to resolve the controversy over facial expression processing. Results from a composite/noncomposite facial expression identification task revealed that both holistic and feature-based processing can be enlisted depending on stimulus format. Specifically, photographic stimuli enlist holistic processing and schematic stimuli enlist feature-based processing. Such findings explain the discrepancies in the literature and suggest that holistic processing is used for facial expression identification under natural viewing conditions. Further research testing this hypothesis using gaze-contingent stimulus presentations is underway.

References

Adolphs, R., Gosselin, F., Buchannan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433,* 68-72. doi:10.1038/nature03086

Adolphs, R., Tranel, D., Hamann, S., Young, A. W., Calder, A. J., Phelps, E. A., et al. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia, 37*, 1111-1117. doi:10.1016/S0028-3932(99)00039-1

Altoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 997-1010. doi:10.1037/0278-7393.25.4.997

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults and adults with autism or Asperger syndrome. *Visual Cognition, 4,* 311-331. doi:10.1080/713756761

Barton, J. S., Radcliffe, N., Cherkasova, M. V., Edleman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception, 35,* 1089-1105. doi:10.1068/p5547

Calder, A. J., & Jansen, J. (2005). Configural coding of facial expressions: The impact of inversion and photographic negative. *Visual Cognition, 12,* 495-518. doi:10.1080/13506280444000418

Calder, A. J., Keane, J., Young, A. W., & Dean, M. (2000). Configurational information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 527-551. doi:10.1037/0096-1523.26.2.527

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments, and Computers, 25,* 257-271.

Ellison, J. W., & Massaro, D. W. (1997). Featural evaluation, integration, and judgment of facial affect. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 213-226. doi:10.1037/0096-1523.23.1.213

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. W. (1998). What is “special” about face perception? *Psychological Review, 105,* 482-498. doi:10.1037/0033-295X.105.3.482

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion, 6,* 534-539. doi:10.1037/1528-3542.6.3.534

Hall, J. K., Hutton, S. B., & Morgan, M. J. (2010). Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition? *Cognition and Emotion, 24,* 629-637. doi:10.1080/02699930902906882

Hedwig, E., & Alpers, G. W. (2011). Happy mouth and sad eyes: Scanning emotional facial expressions. *Emotion, 11,* 860-865. doi:1037/a0022758

Hsiao, J. H., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science, 19,* 998-1006. doi:10.1111/j.1467-9280.2008.02191.x

Lagner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Kippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24,*1377-1388. doi:10.1080/02699930903485076

Lipp, O. V., Price, S. M., & Tellegen, C. L. (2009). No effect of inversion on attentional and affective processing of facial expression. *Emotion, 9,* 248-259. doi:10.1037/a0014715

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6,* 255-260. doi:10.1016/S1364-6613(02)01903-4

McKelvie, S. J. (1995). Emotional expression in upside-down faces: Evidence for Configurational and componential processing. *British Journal of Social Psychology, 34,* 325-334.

Ohman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 3,* 381-396. doi:10.1037/0022-3514.80.3.381

Orban de Xivry, J., Ramon, M., Lefevre, P., & Rossion, B. (2008). Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *Journal of Neuropsychology, 2,* 245-268. doi:10.1348/174866407260199

Schwarzer, G., Huber, S., & Dummler, T. (2005). Gaze behavior in analytical and holistic face processing. *Memory & Cognition, 33,* 344-354.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experiment Psychology, 46A,* 225-245.

Tiddeman, B., Burt D. M., & Perrett, D. (2001). Computer graphics in facial perception research**.** *IEEE Computer Graphics and Applications, 21*, 42-50.

Tipples, J., Atkinson, A. P., & Young, A. W. (2002). The eyebrow frown: A salient social signal. *Emotion, 2,* 288-296. doi:10.1037/1528-3542.2.3.288

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79,* 471-491.

Van Belle, G., De Graef, P., Verfaillie, K., Rossion, B., & Lefevre, P. (2010). Face inversion impairs holistic perception: Evidence from gaze-contingent stimulation. *Journal of Vision, 10,* 1-13. doi:10.1167/10.5.10

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747-759. doi:10.1068/p160747

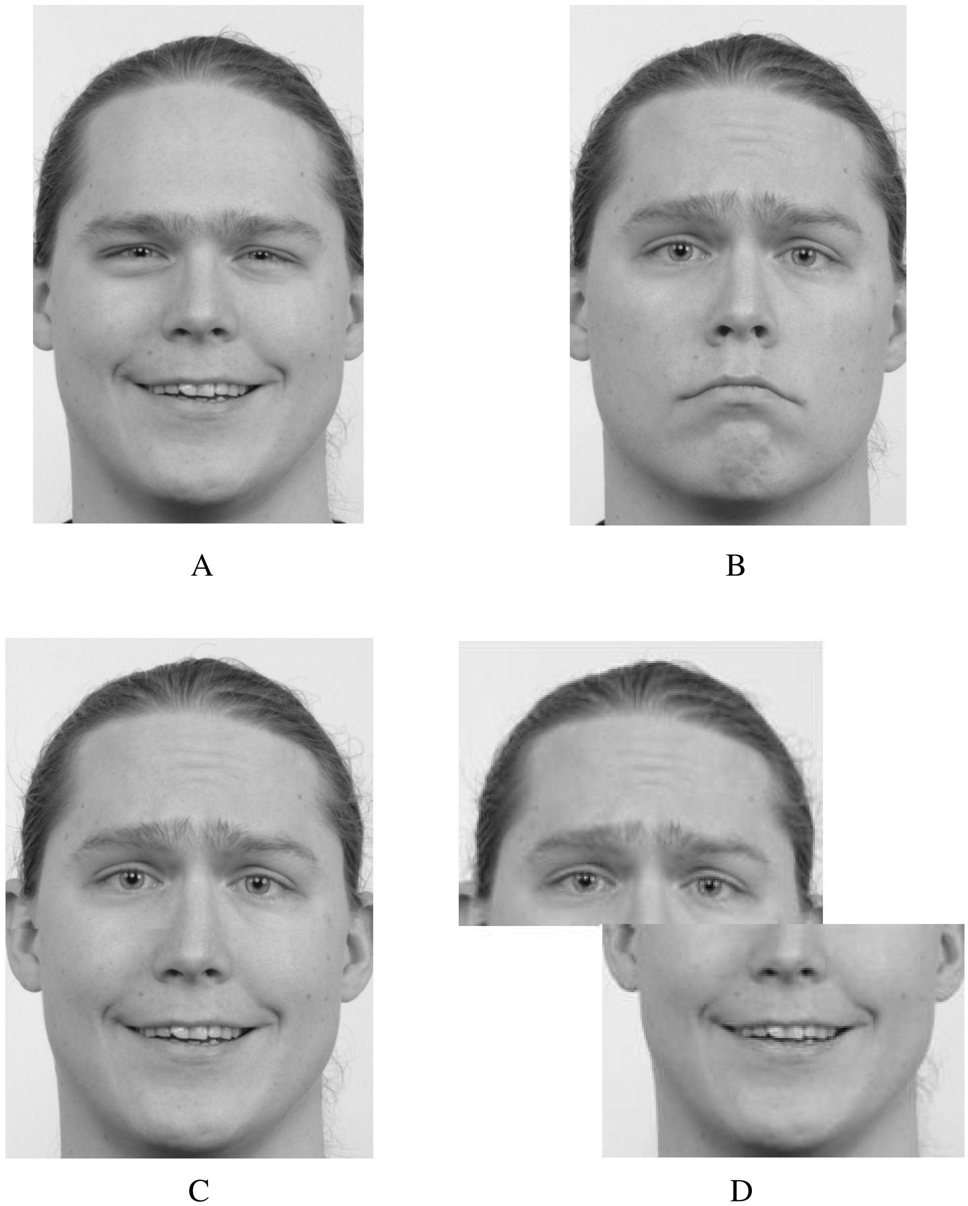


Figure 1. Photographic sad/happy composite (C) and noncomposite (D) made from prototypical happy (A) and sad (B) expressions.

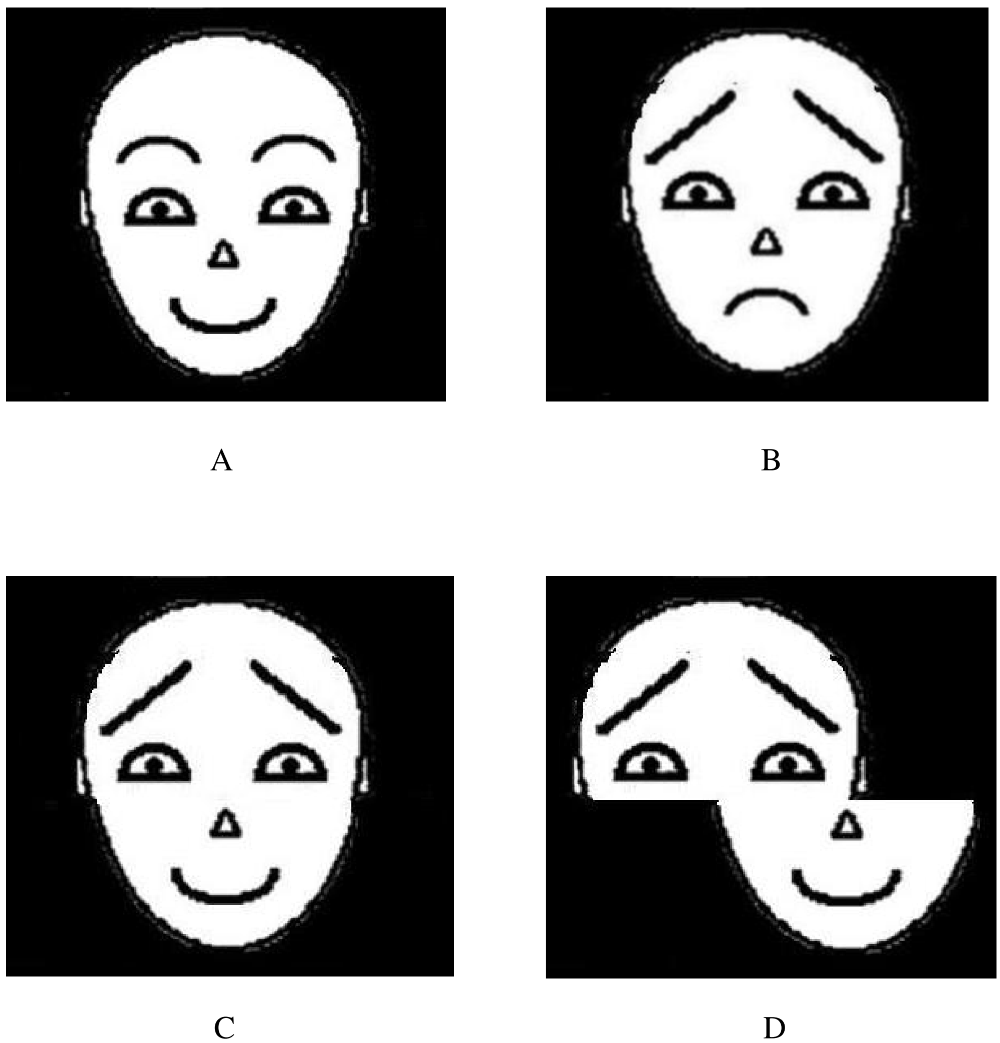


Figure 2. Schematic sad/happy composite (C) and noncomposite (D) stimuli made from prototypical happy (A) and sad (B) expressions.

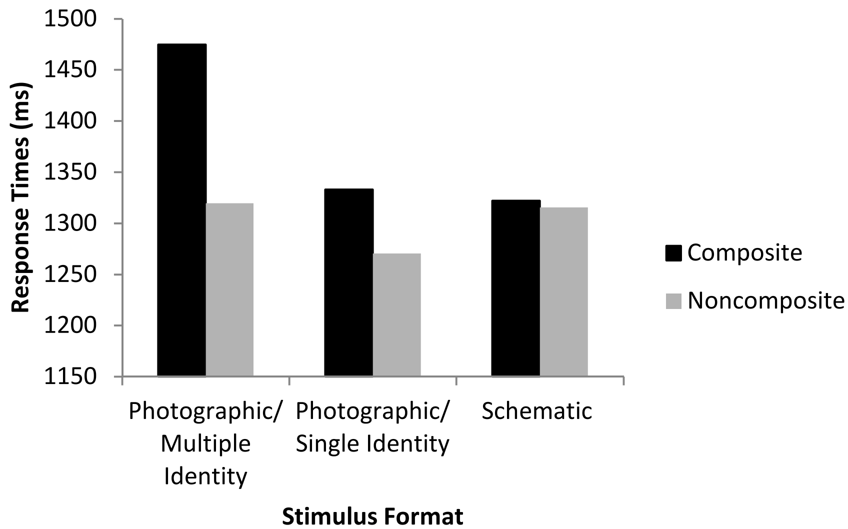


Figure 3. Mean reaction times displayed as a function of stimulus format (multiple-identity photographic vs. single-identity photographic vs. schematic) by stimulus type (composite vs. noncomposite).

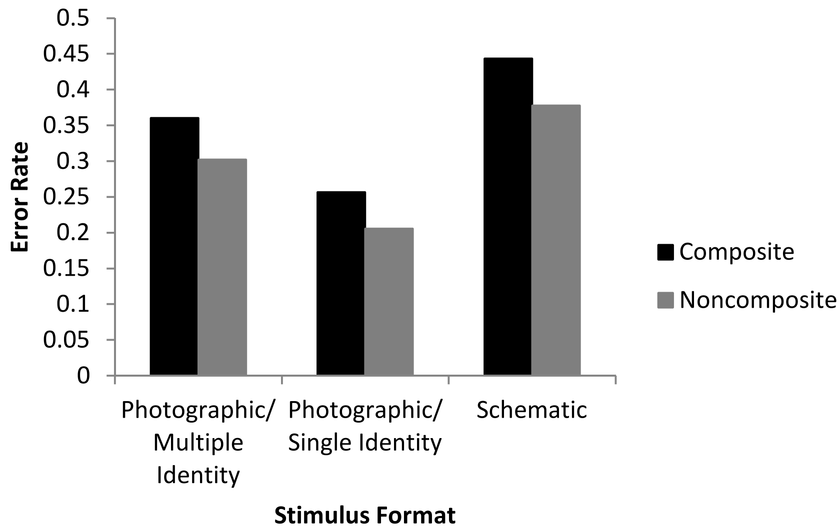


Figure 4. Mean error rates displayed as a function of stimulus format (multiple-identity photographic vs. single-identity photographic vs. schematic) by stimulus type (composite vs. noncomposite).

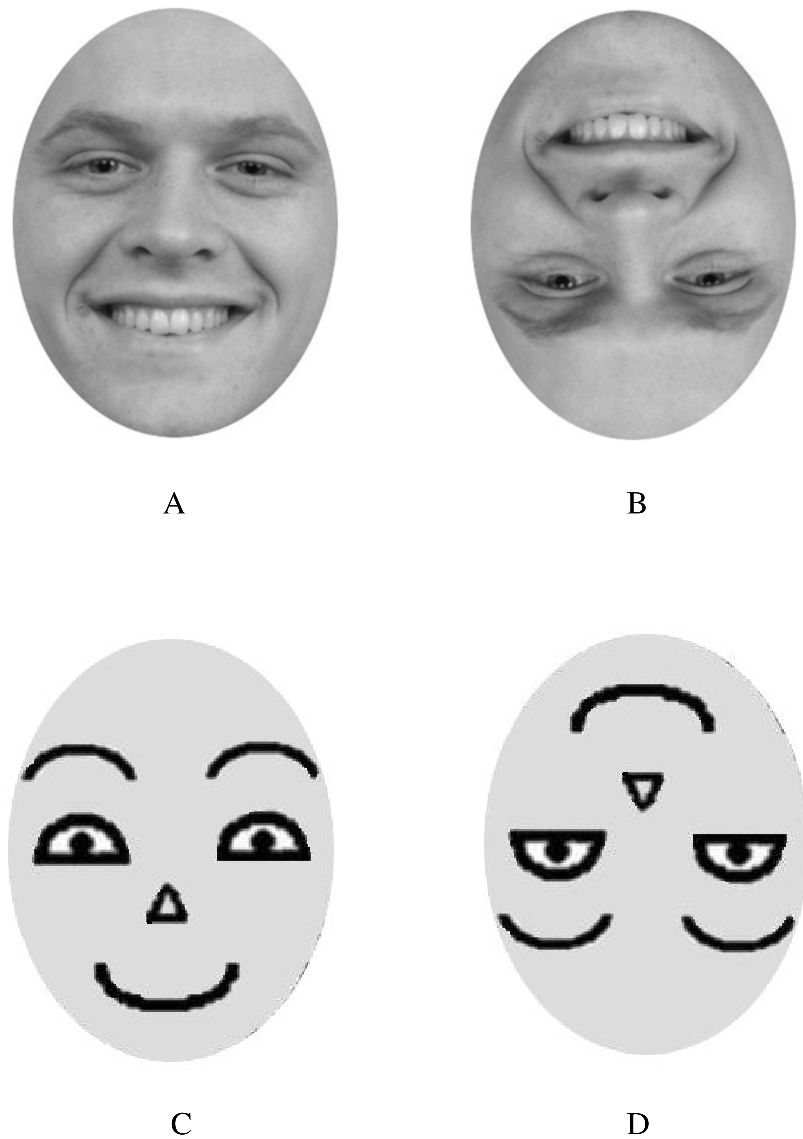


Figure 5. Examples of photographic stimuli displayed upright (A) and inverted (B), and schematic stimuli displayed upright (C) and inverted (D).

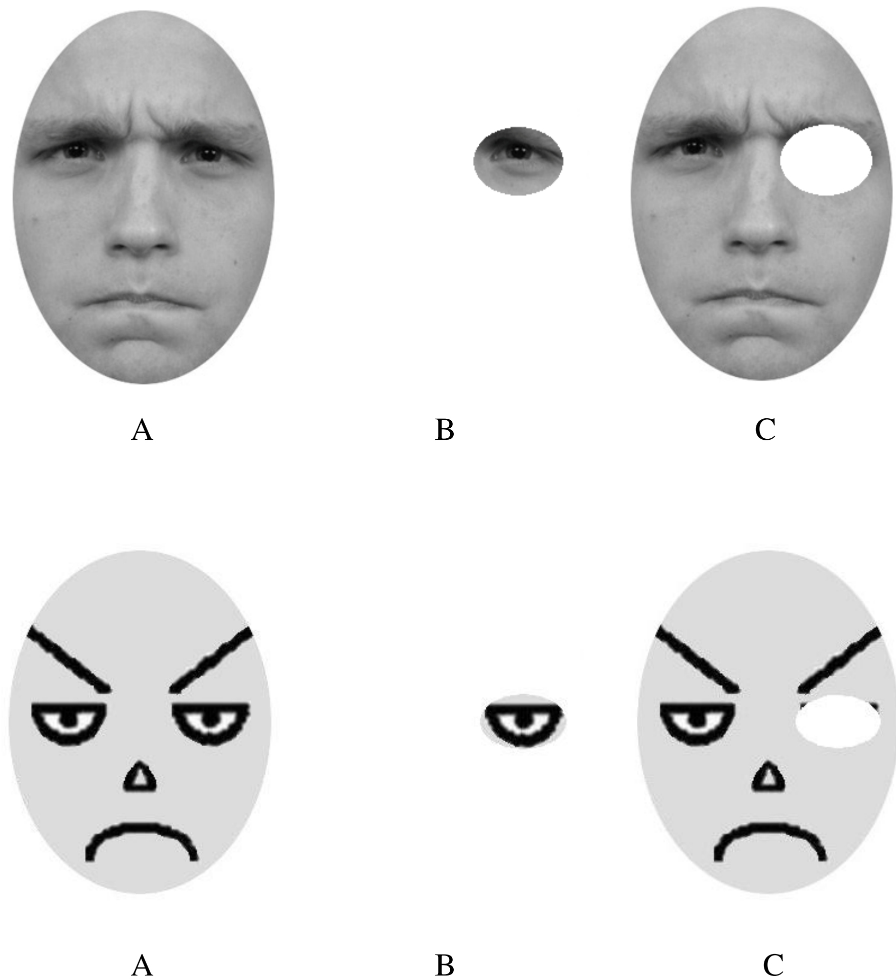


Figure 6. Examples of viewing conditions: full-view (A and D), window-view (B and C), and mask-view (C and F) with angry photographic and schematic stimuli.

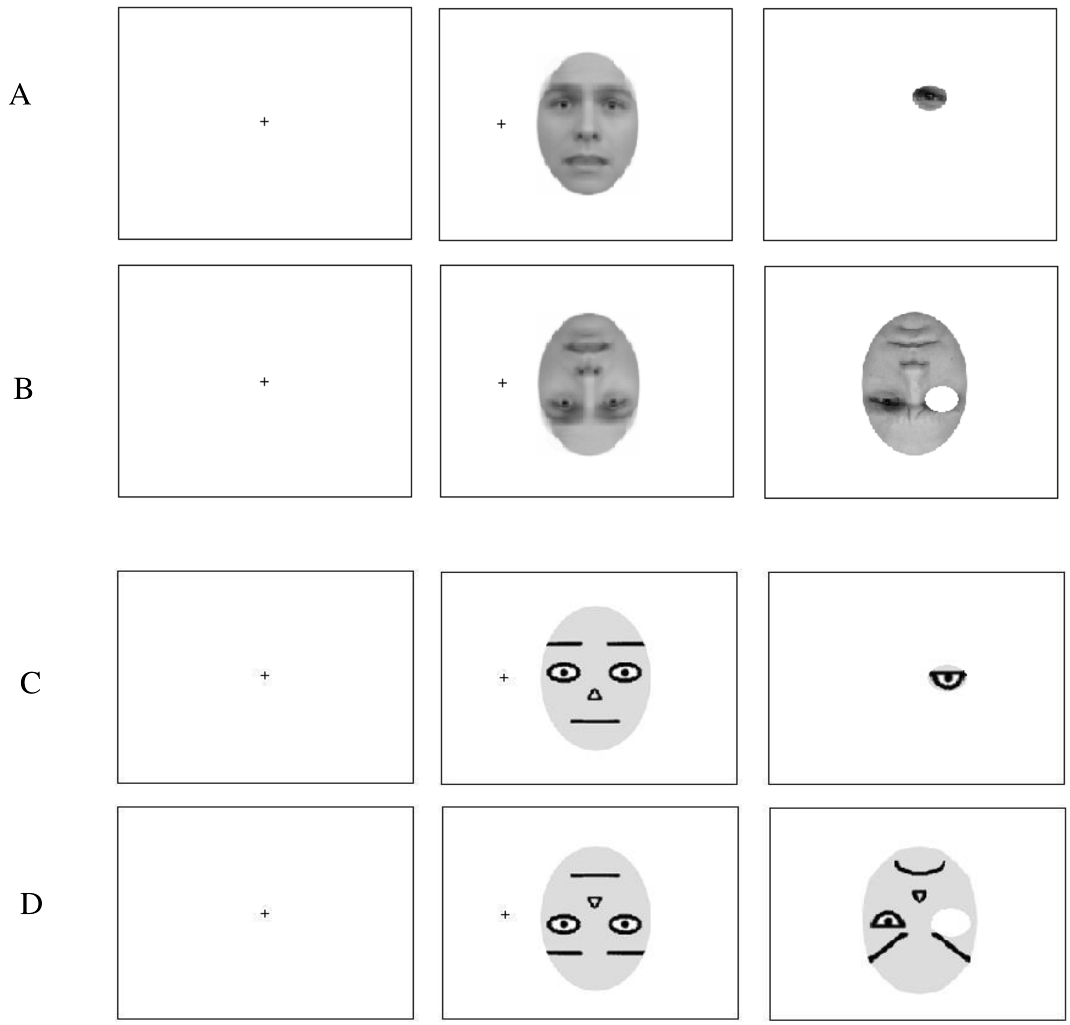


Figure 7. Trial time course examples for photographic upright window-view (A), photographic inverted mask-view (B), schematic upright window-view (C) and schematic inverted mask-view (D).

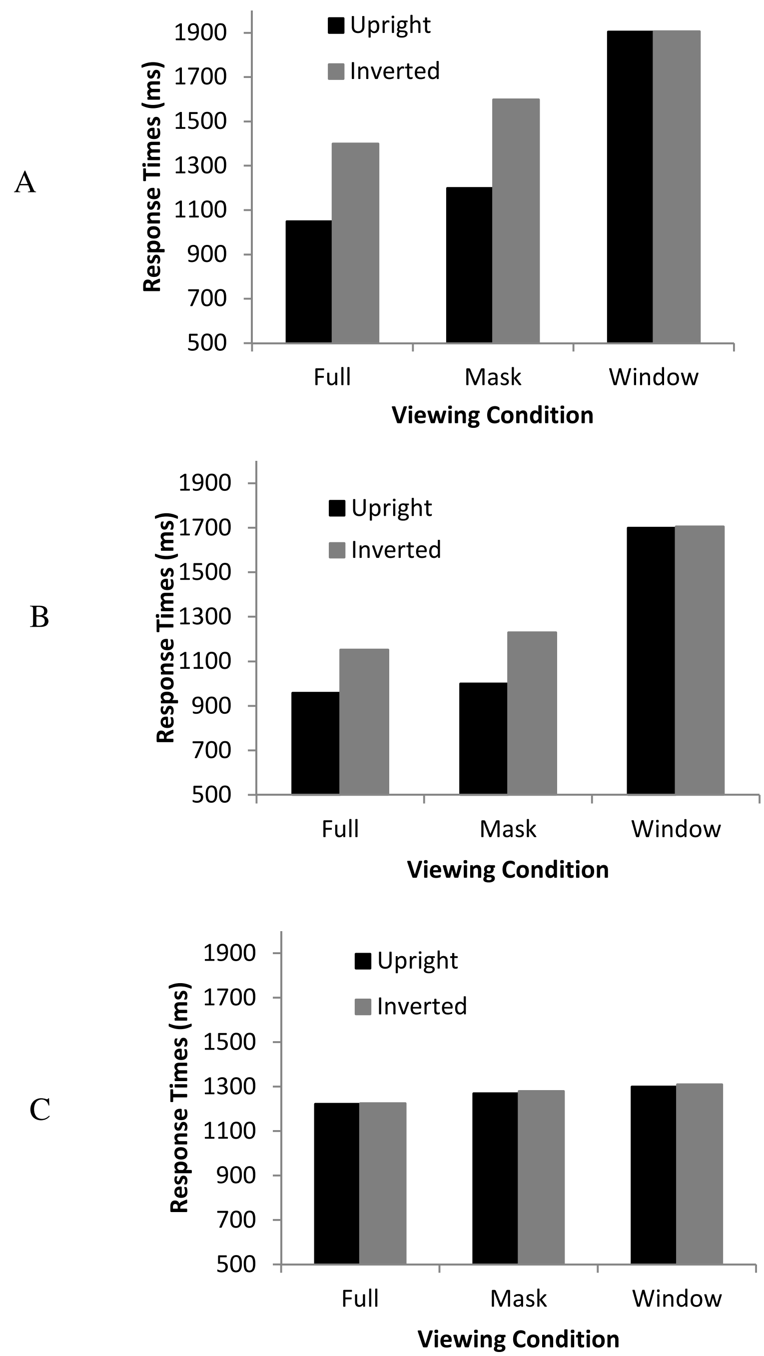


Figure 8. Hypothetical average response times for upright and inverted faces in the full, mask, and window viewing conditions for the photographic/multiple identity (A), photographic single/identity (B), and schematic (C) conditions.

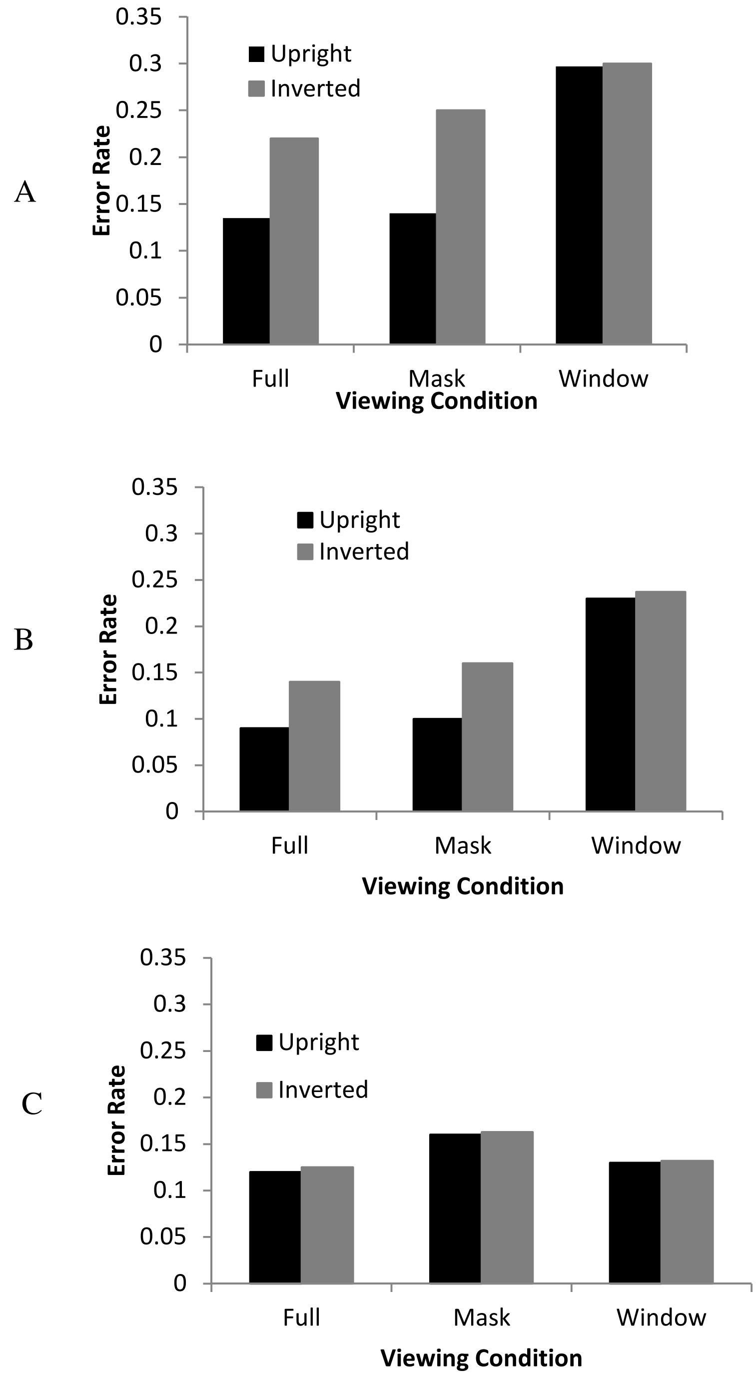


Figure 9. Hypothetical average error rates for upright and inverted faces in the full, mask, and window viewing conditions for the photographic/multiple identity (A), photographic single/identity (B), and schematic (C) conditions.

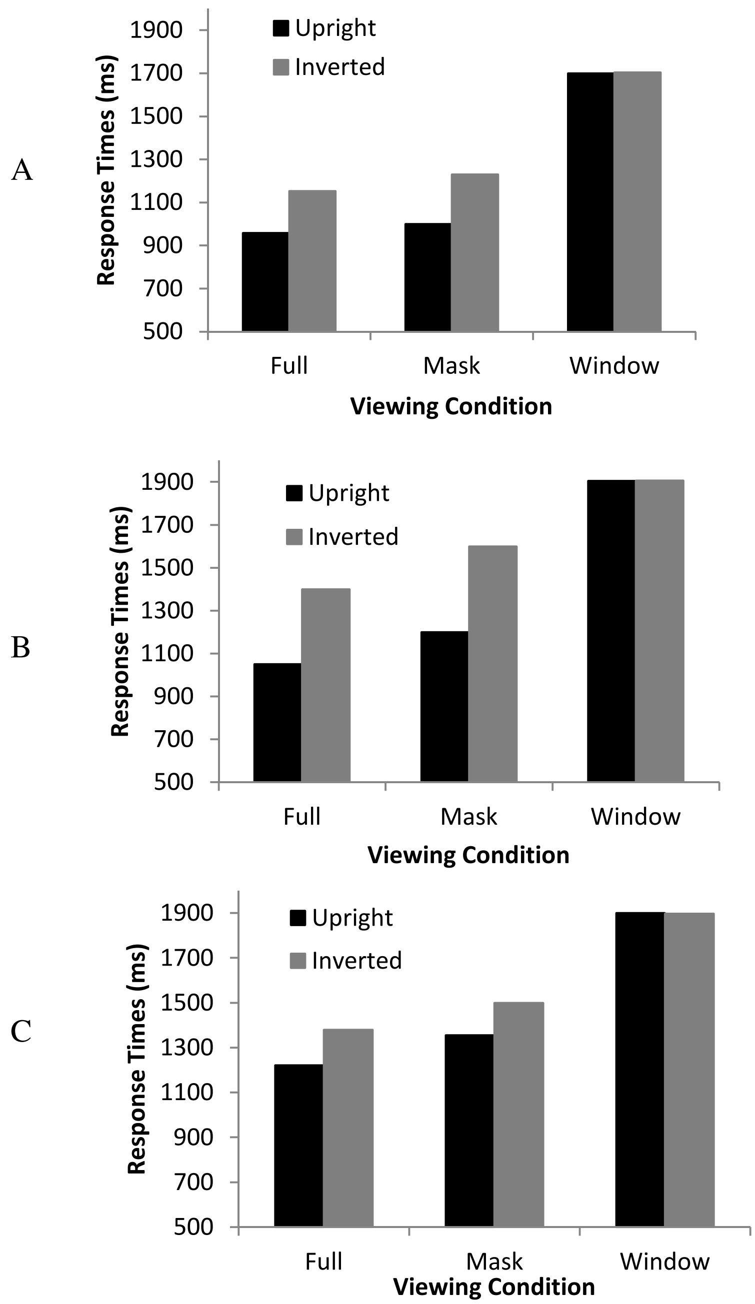


Figure 10. Hypothetical average response times for upright and inverted faces in the full, mask, and window viewing conditions for the photographic/multiple identity (A), photographic single/identity (B), and schematic (C) conditions.

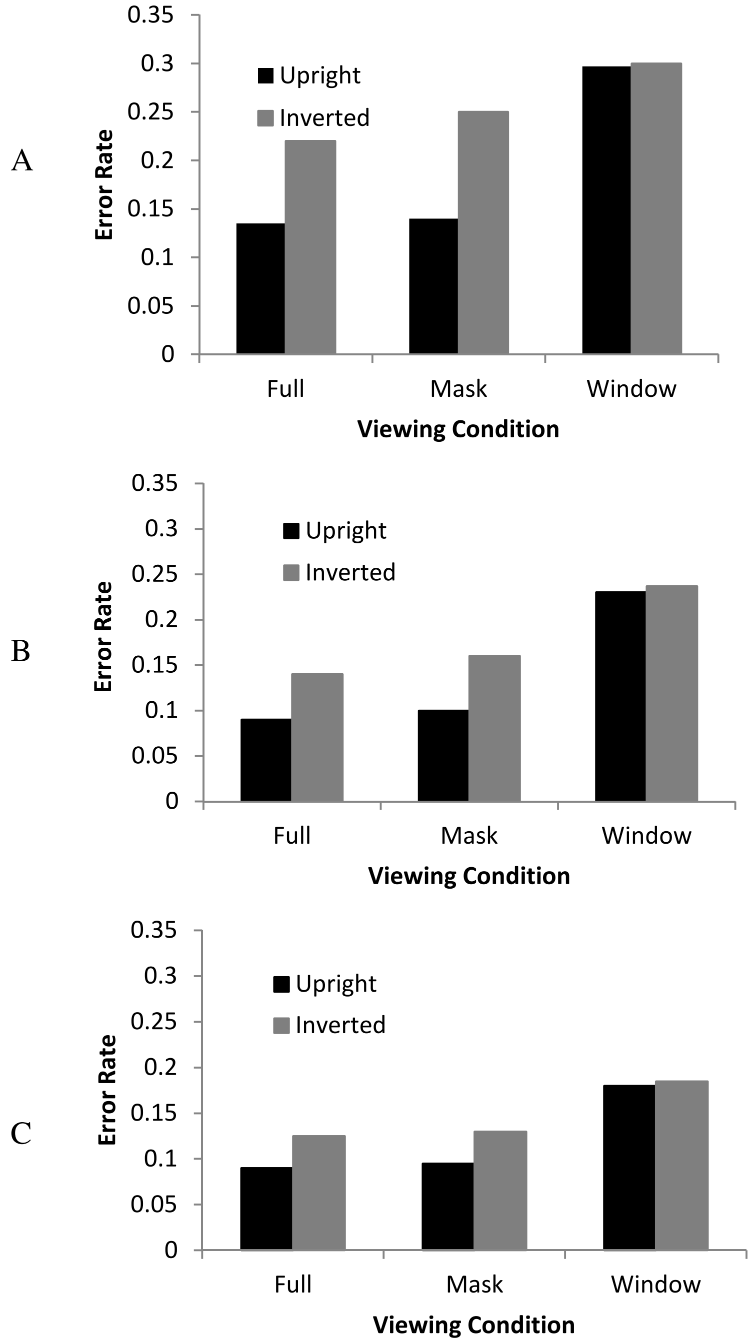


Figure 11. Hypothetical average error rates for upright and inverted faces in the full, mask, and window viewing conditions for the photographic/multiple identity (A), photographic single/identity (B), and schematic (C) conditions.

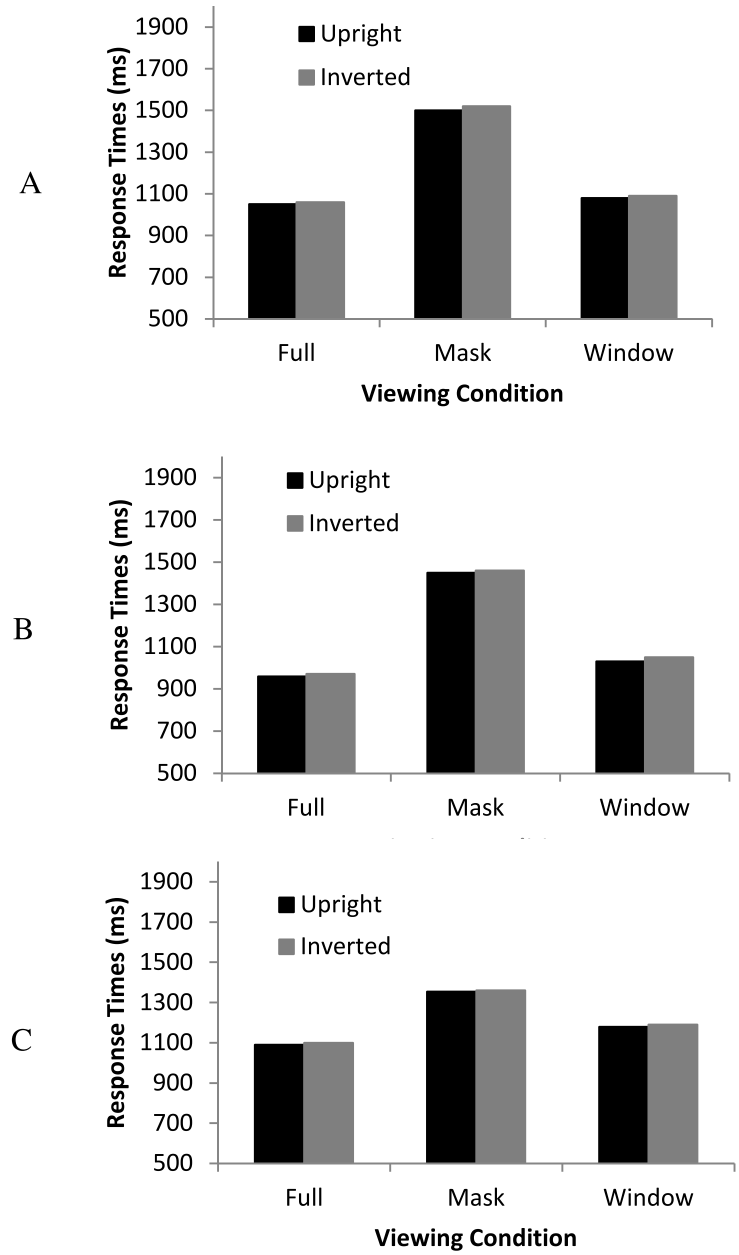
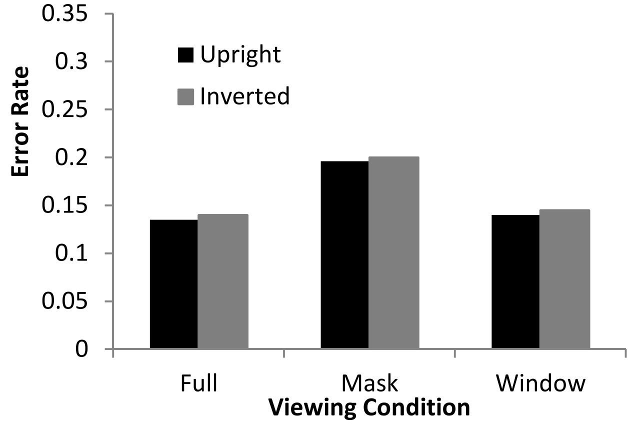
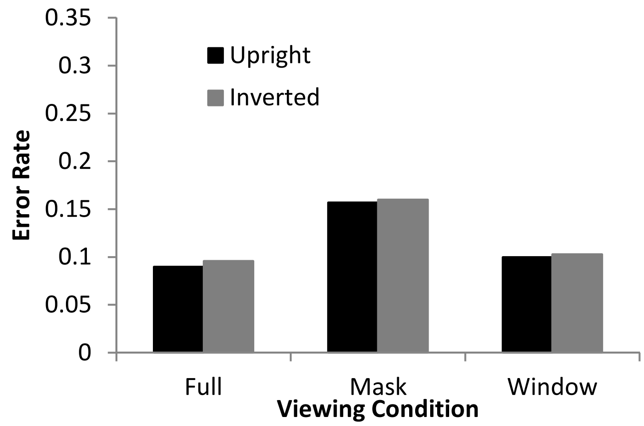


Figure 12. Hypothetical average response times for upright and inverted faces in the full, mask, and window viewing conditions for the photographic/multiple identity (A), photographic single/identity (B), and schematic (C) conditions.

A



B



C

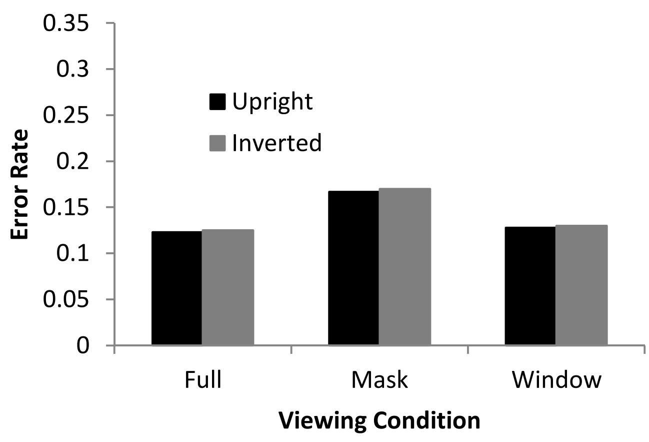


Figure 13. Hypothetical average error rates for upright and inverted faces in the full, mask, and window viewing conditions for the photographic/multiple identity (A), photographic single/identity (B), and schematic (C) conditions.