Wellesley College
Wellesley College Digital Scholarship and Archive

Honors Thesis Collection

2013

The Importance of Configural Information and the Time-Course for the Perception of Animacy in Faces

Lila Fakharzadeh
lfakharz@wellesley.edu

Recommended Citation

Fakharzadeh, Lila, "The Importance of Configural Information and the Time-Course for the Perception of Animacy in Faces" (2013). *Honors Thesis Collection*. 137.
https://repository.wellesley.edu/thesiscollection/137

The Importance of Configural Information and the Time-Course for the Perception of Animacy in Faces

by

Lila Naheed Fakharzadeh

Submitted in Partial Fulfillment
of the
Prerequisite for Honors
in Neuroscience

May 2013

Abstract

It has been proposed that the processing of faces is a highly specialized function that is separate from the processing of other types of objects. Evidence for this appears with inversion, where it has been observed that face recognition is more affected than the recognition of other types of stimuli (Diamond and Carey, 1986). From these observations it can be inferred that holistic processing of faces makes face analysis unique. This thesis focuses on the question of whether configural information encoded in a face can be used quickly to determine the animacy of a face, or if a more detailed analysis of face features is necessary. By animacy we mean whether the image portrays a real human face or that of a “human-like” inanimate object such as a doll or a manikin. In the first part of this thesis face morphs between doll and human faces were inverted to see if animacy perception of faces would be affected with less configural information. In the second part of the thesis, the time-course of animacy perception in relation to face recognition is analyzed. The final part of the thesis examines the role of high and low spatial frequency ranges in encoding information used for animacy judgment. The results from the first experiment suggest that holistic processing of the face is necessary for animacy perception. From the temporal experiment it can be deduced that animacy perception occurs more quickly than face recognition, and may require less detailed face analysis. The results from the spatial frequency pilot experiment suggest that removing the high spatial frequencies or low spatial frequencies from the face images has no effect on animacy perception, however, more data is needed in order to form a strong conclusion.

Acknowledgments

I would like to thank my advisor Ellen Hildreth for being extremely supportive throughout the process. I would not have had a rich learning experience without your continuous thoughtful insight and guidance.

I would also like to thank Kami Koldewyn for making the experimentation for this thesis possible, and for your positive energy and enthusiasm.

Special thanks to everyone in the Kanwisher lab including Nancy Kanwisher, Daniel Dilks, Alex Kell, Jason Fischer, Idan Blank, Terri Scott, Caitlin Malloy, Beverly Cope, Jessica Pourian and Sam Norman-Haignere as well as Aude Oliva for being a pleasure to be around and to work with in the lab. I will always remember my experiences in the lab with all of you. Thanks to Bevil Conway for originally giving me the idea to expand the project for his class into a thesis and for always providing me with feedback. I would also like to include my Wellesley sisters and my Professors who have made my college experience truly memorable.

Finally, I would like to thank my grandmother for being my back-bone and source of happiness since I was young, and also my parents, brothers, aunts and uncles, and cousins for their unconditional love and sacrifices.

Contents

1. Introduction

The processing of faces is an important task for humans; information including gender, age, identity, animacy, and emotional state can be extracted from facial analysis (Looser and Wheatley, 2010). There is increasing evidence that the processing of faces is of special importance to people, and to primates in general, and holds a privileged place within the visual system (Farah, 1995; Kanwisher, McDermott, and Chun, 1997; Farah, Wilson, Drain, and Tanaka, 1998; Ro, Russell, and Lavie, 2001). The visual cues provided by the face contain information that aid social interaction between humans, and this processing may be more complex than the perceptual skills required for object recognition (Haxby et al., 2000). This thesis focuses on the question of whether configural information encoded in a face can be used quickly to determine the animacy of a face, or if a more detailed analysis of face features is necessary. By animacy we mean whether the image portrays a real human face or that of a "human-like" inanimate object such as a doll or a manikin. We address this question through experiments that examine the effects of inversion, exposure duration, and spatial frequency content on the perception of animacy in faces. By inverting the image of a face, we reduce the holistic processing of the face and can then observe the effects on animacy perception. We also compare the effect of exposure duration on animacy judgments to its effect on face recognition to explore how the time course of animacy perception compares to that of face recognition.

Previous experimental studies have found that faces are processed both featurally and holistically (Tanaka and Farah, 1993). Evidence in support of holistic processing comes in part from studies that examine the effect of inversion on face recognition performance. Face inversion disrupts the familiar spatial relationships between the facial features (Ellis et al., 1975). Face recognition is more heavily affected by inversion than recognition of other types of stimuli (Diamond and Carey, 1986), suggesting that face recognition may use more holistic processing than general object recognition (Le Grand et al., 2004). Configural information for analyzing faces has been found to be weaker for inverted faces (Tanaka and Gauthier, 1997). The orientation of a face has not been observed to have a significant effect on part-based or featural processing of a face; subjects are able to successfully identify alterations in the size of the eyes, mouth, or lips regardless of whether the face is upright or inverted (Leder and Bruce, 1998; Le Grand et al., 2001a, 2001b). Differences in performance for upright vs. inverted faces have been observed both in studies that employ recognition memory (Yin, 1969) and those involving comparison tasks between upright and inverted faces (Bartlett and Searcy, 1993; Kanwisher et al., 1998). The recognition of individual facial features is also improved when the feature is embedded in an upright face relative to when it was embedded in an inverted face (Thompson, 1980). The memory of a face is also influenced by the distances between face parts (Bartlett and Searcy, 1993). The performance differences in these tasks may be the result of different processing methods used by the brain to analyze the face in these two different orientations (McKone, 2004).

There is also evidence for some sort of holistic processing of the face based on observations made about patients with prosopagnosia. Experiments have shown that prosopagnosics do not show an inversion effect in the recognition of faces (Farah et al., 1995). The inability to take advantage of holistic processing in the analysis of upright faces may explain the lack of an inversion effect in prosopagnosics. It has been suggested that prosopagnosics identify faces using a feature by feature analysis.

Further evidence for the importance of holistic processing for face recognition can be obtained from experimental results showing that caricatured faces, whose individual features are clearly distorted, are often recognized more easily than actual images of the face (Lee et al., 2000). This suggests that even when individual features of a face do not have the same proportions or shape in real-life, when placed in the context of a face it can be recognized. Holistic information may also play a role in modulating the featural analysis of a face; experiments have shown that the recognition of face parts is greater when these parts are presented in the context of an entire face than when presented in isolation (Tanaka and Farah, 1993; Kanwisher, 2000). Although these experimental results clearly show the importance of holistic processing, the mechanism or computation behind this holistic processing of the face is not clear.

When considering the role of holistic or configural processing versus detailed processing of facial features, it is important to distinguish between face detection and face recognition. Face detection and recognition may be two separate processes that require different forms of visual information and analysis. While face recognition requires the extraction of information from the face that makes it unique (Tsao and Livingstone, 2008), face detection requires only the recognition of a class of objects (i.e. faces as an object group). Detecting faces is required before identification, gender classification, and emotion can be extracted from the face (Torralba and Sinha, 2001). A study by Kanwisher et al. (1998) examined fMRI data from the fusiform face area (FFA), which is thought to be specialized for face detection (Allison et al., 2000; Kanwisher et al., 1997). It was observed that inverting a face impairs the recognition of a face but not the detection of a face (Kanwisher et al., 1998). Subjects in this experiment were presented with a series of inverted and upright faces and were asked to press a button when a face was repeated twice in a row. Subjects were more successful in identifying repeats when the faces were shown upright than inverted, indicating that they were able to process upright faces more effectively than inverted faces. The fMRI data, however, showed that the FFA was activated significantly in response to inverted faces, and that the activity in the FFA was decreased by only 15% when inverted pictures of faces were shown to subjects, compared to the activity measured for upright faces. The FFA may be activated in the presence of a face and may not be responsible for the processes needed to discriminate between faces, reinforcing the importance of distinguishing face detection and recognition.

The definition of configural processing of a face is still under debate by researchers (Maurer et al., 2002), but there is general agreement that some sort of integrative processing of the face occurs when they are upright. There are two interpretations of the nature of configural processing of a face. One interpretation is that the configural processing of a face cannot be decomposed into separate parts such as the eyes, nose, and mouth (Tanaka and Sengco, 1997), but rather the primitive lines and edges from early visual stages are integrated into one holistic perception of the face (Moscovitch et al., 1997). The second interpretation is that the features of the face (e.g. the eyes, mouth, and nose) are processed independently, and then the configural information is determined from the distances between these parts across the face (Bruce et al., 1991; Rhodes, 1988; Sergent, 1985; Goffaux et al., 2005). Another question raised by these definitions is what exactly are the stages of configural processing. It is still uncertain whether configural and part-based analyses occur in parallel or whether these systems operate sequentially (Moscovitch et al., 1997). In a study by Moscovitch and colleagues (1997), it was determined that part-based processing of a face can occur without configural processing. It was also observed that configural face analysis can occur independently of part-based analysis. Although these results suggest that the pathways may be parallel, rather than sequential, this does not necessarily mean that the two pathways cannot influence one another.

A study by Freiwald et al. (2009) with macaque monkeys provides some indication of the neural pathways involved in configural processing of faces. Recording from face sensitive cells in the macaque middle face patch, it was observed that out of the 272 face cells that were recorded from, 90% of the cells were receptive to at least one dimension of the face stimuli (Freiwald et al., 2009). These cells were not tuned to all of the dimensions in the face. Responses of 59% of the cells were modulated by the face aspect ratio, 46% by iris size, 39% by height feature assembly, 31% by inter-eye distance, and 27% by face direction. The researchers concluded face cells in the middle patch encode feature dimensions of faces. The activity of the cells was correlated with the degree to which the featural dimensions were different from the average value. This amplified response when featural dimensions transgress from the average value for that dimension implies that caricatured faces should be distinguished more than average faces (Freiwald et al., 2009). The gestalt of a face, therefore, increases the tuning of these cells to the facial dimensions.

The amount of time necessary for humans to perform visual tasks is closely linked to the time-course of the cognitive processes in the brain (Özbek and Bindemann, 2011). In a study by Keysers et al., where recordings of neurons selective for faces and other complex patterns were made, it was found that 65% of the neurons preserved their selectivity for a particular stimulus even at rapid serial visual presentation (RSVP) rates of 14 ms/image (Keysers et al., 2001). The perceptual performance of the subjects on the RSVP behavioral tasks was found to be correlated with the activity of single neurons during the RSVP tasks. The memory and detection of faces were also observed to be above chance at presentation rates of 14 ms/image. These results suggest that face detection can occur very quickly by the visual system, in as little time as 14 ms. Humans have the ability to recognize faces at very short exposure times of about 50-90 ms (Seeck et al., 1997). In a study by Yin, the brief presentation of faces was found to impair face recognition (Yin, 1970). In a face recognition study conducted by Veres-Injac and Schwaninger, subjects were given the tasks of matching the whole face, external features, or internal features of the face for different exposure durations. The exposure durations tested were 90, 120, 150 milliseconds, and self-paced (Veres-Injac and Schwaninger, 2009). External features consisted of the hair, ears, head, and face outline, while internal features were the eyes, eyebrows, nose, and mouth. The exposure durations of 120 ms, 150 ms, and self-paced did not yield a significant improvement in face recognition accuracy compared to that obtained from 90 ms exposure. Furthermore, the external features of a face were found to be recognized more easily than internal features (Veres-Injac and Schwaninger, 2009). In another study, it was also found that a short presentation time of 80 ms led to the processing of holistic information, and with long presentation time (2 s), a feature-matching strategy was used (Richler et al., 2009).

Visual stimuli are analyzed at multiple scales by the human visual system using different ranges of spatial frequency (Ruiz-Soler and Beltran, 2006). There are underlying channels in the visual system that process low and high spatial frequencies independently (Livingstone and Hubel, 1988; Merigan and Maunsell, 1993). This suggests that configural and featural information are processed using different pathways (Halit et al., 2006). High spatial frequencies are carried by the parvocellular channels to the cortical ventral visual stream (Livingstone and Hubel, 1988; Merigan and Maunsell, 1993). This pathway analyzes visual stimuli at a finer resolution than the analysis of low-spatial frequency information but takes more time (Johnson, 2005). Lower spatial frequencies, on the other hand, are carried by the magnocellular channels to the superior colliculus and pulvinar- a process that occurs relatively quickly (Schiller et al., 1979). Characteristics of the face such as identity, gender, and expression are conveyed in different ranges of spatial frequency (Halit et al., 2006). It was noted above that holistic processing of the face occurs at relatively short exposure times (Richler et al., 2009). Given the different spatial and temporal properties of processing in the parvocellular and magnocellular pathway, this suggests that holistic processing may at least initially be based on the low spatial frequency content of face images. Subjects are attuned to certain ranges of spatial frequency based on the task at hand (Ruiz-Soler and Beltran, 2006). For example, if you are asked to identify which members of a group of people are overweight, you will be more inclined to use the low spatial frequencies in the image, while if you are looking for members who are wearing glasses, you will focus on the high spatial frequencies in the image (Hoeger, 1997). Determining the age of an individual, requires high-spatial frequency information such as wrinkles around the eyes or creases on the forehead and texture of the skin (Halit et al., 2006). For face recognition the intermediate spatial frequencies (8-16 cycles per face) and lower spatial frequencies have been found to be the most critical (Halit et al., 2006). When face images lacked high-spatial frequency information, subjects asked to recognized individual faces performed with the same accuracy as with the original unfiltered images (Harmon and Julesz, 1973). It is true, however that when an image contains a larger band of frequencies faces are generally easier to recognize because more spatial frequency channels are activated (Liu et al., 2000). The shading and texture of a face have also proven to be important features for face recognition, which means that coarse and fine spatial frequencies are both used (Liu et al., 2000).

Spatial frequency analysis of images occurs in the very early stages of visual processing (Goffaux et al., 2005). In a study by Goffaux et al., it was observed that in tasks that required detecting differences in configural information between faces, the use of low spatial frequencies led to better performance than the use of high spatial frequencies (Goffaux et al., 2005). Based on these results the authors concluded that low spatial frequencies are more important for configural processing, whereas featural processing is dependent on high spatial frequencies. Goffaux et al. also emphasized that configural information can be taken from both LSF and HSF faces. The high spatial frequencies are not only used for detailed featural analysis, but can also be used to obtain more precise metric relations between the features (Goffaux et al., 2005).

In a study by Torralba and Sinha, the threshold in image resolution needed for distinguishing between face and non-faces was determined experimentally (Torralba and Sinha, 2001). The visual system has the ability to identify objects even with very low spatial resolution (Torralba and Sinha, 2001). Torralba and Sinha found that even at very low resolutions, the subjects were able to reject non-face patterns. Face patterns could be discriminated from non-face patterns even at 1.3 cycles eye-to-eye resolution. It was also observed that the inversion of faces decreased face detection performance at full resolution (Torralba and Sinha, 2001).

A study by Looser and Wheatley (2010) explored the particular task of judging animacy in static photographs of faces using a morph continuum made with human and mannequin faces. The subjects were instructed to scroll through a morph continuum of doll and human faces and decide where the animacy boundary was located. It was determined that the threshold in the continuum is located near the human end of the spectrum. The researchers observed that the subjects' animacy judgments were closely linked to the features present in the face. Isolating the eyes, mouth, nose, and skin, they found that the most informative feature for animacy judgment in the face is the eyes. The ratings of the separate features were correlated with animacy ratings of whole faces, and the highest correlation was for the eyes (r= 0.87), then the nose (r=0.60), and lastly the skin (r=0.37). Surprisingly, the animacy ratings for these isolated features of the face did not exceed 68%, even with isolated features that were taken from 100% human faces. These results suggest that animacy judgment may rely heavily on a holistic processing of the face (Looser and Wheatley, 2010).

The first part of my thesis explores the visual cues used to make animacy judgments by comparing performance on upright vs. inverted faces. Morph continua of doll and human faces, ranging from 0% human (100% doll) to 100% human in increments of 10%, were used in this study. The subjects were asked to rate an individual face as either doll or human. The configural information for analyzing faces has been found in previous studies to be weaker for inverted faces (Tanaka and Gauthier, 1997). Holistic information about a face can be extracted very rapidly after a face stimulus is presented (Richler et al., 2009). In order to observe the effects of decreased holistic processing on the perception of animacy, we compared animacy judgments for upright and inverted faces. It was predicted that we would identify an 'inversion effect' on animacy perception, because based on the results of Looser and Wheatley it was observed that configural information is important for the formation of animacy perception. The results of this experiment show that the reduction of configural information increased the variability in animacy ratings and affected animacy perception of real human faces, but not doll faces.

The second part of the study seeks to gain insight on the time-course of animacy processing. It was noted earlier that the amount of time necessary for humans to conduct visual tasks is closely linked to the time-course of the processes in the brain needed for the task (Özbek and Bindemann, 2011). Previous studies examined the time-course of face recognition, leading to the observation that humans can recognize faces at very short exposure times between 50-90 ms (Seeck et al., 1997; Veres-Injac and Schwaninger, 2009; Richler et al., 2009). In the experiments described in this thesis we limited the presentation times of the morph images to determine the amount of time necessary to form a judgment on the animacy level of a face. A face recognition task was also conducted to compare the time-course of animacy perception to that for face recognition. As noted above it was observed that approximately 90 ms is needed for face recognition (Veres-Injac and Schwaninger, 2009). Face recognition may need more stages of processing in the brain compared to animacy perception, and thus require more exposure time as a result. Simpler cues such as eye size and skin texture may be used in animacy judgment, possibly requiring less exposure time than the cues used for face recognition. The exposure durations of the visual stimuli were limited to 33 ms, 84 ms, and 2.02 seconds. The first part of the temporal experiment used a face recognition task with the face stimuli presented at the three exposure times. In the second half of the temporal experiment, the morph continua were used so that the perceived animacy of the faces with exposure times of 33 ms and 84 ms could be compared to the ratings from 2.02 seconds of viewing time, where there was a sufficient amount of exposure time for full analysis of the face. It was predicted that at 33 ms of exposure time, there is not enough time for processing the face in both the animacy and recognition tasks. With a 84 ms exposure duration, it may be possible for subjects to judge animacy, but not recognize a face. The results of this experiment show that animacy judgment occurs earlier than face recognition, suggesting that less detailed face analysis is needed for animacy perception.

The spatial frequency ranges for determining the identity, gender, and expression of a face have been examined in previous studies (Halit et al., 2006). In this experiment, we examined whether a particular range of spatial frequencies is critical in animacy perception. In the study by Looser and Wheatley, it was found that configural information is more important for animacy perception. From the results of the Looser and Wheatley experiments, it was predicted that the low spatial frequencies will be more important because holistic cues are most likely best encoded in the low spatial frequencies of the face. Featural or detailed cues are best encoded in the high spatial frequencies of the face. If eye size is an important cue for distinguishing real vs. doll faces, we might predict, on the other hand, that the high spatial frequencies, which preserve detailed spatial information, may be more important for animacy judgment. In the final part of the study, different spatial frequency ranges representing the low and high frequencies of the face were presented to the subjects in a pilot experiment. The ratings from the low and high spatial frequencies were compared to the ratings of the faces containing all of the spatial frequencies. It was predicted that the low spatial frequencies would be more important for animacy perception than the high spatial frequencies. From the pilot experiment, there was no significant difference between the point where the subjects' perceived the morph as being half human and half doll, for the the low spatial frequency images and the unfiltered images. A difference was also not observed between the high spatial frequency images and the unfiltered images. The subjects' certainties in animacy judgment were also not found to be significantly different when comparing the ratings for the low spatial frequency images with the unfiltered images, and when comparing high spatial frequency images with the unfiltered images.

The Inversion Experiment, Temporal Experiment, and Spatial Frequency Experiment are discussed in sections 2-4 of this thesis. For each of the three experiments in sections 2-4, a brief discussion of the reasoning behind the experimentation, the subjects that were tested, the stimuli that were used in the experiments, the procedure, and the results of the experiments are provided. A discussion on the interpretation of the results from the three experiments is given in section 5 of the thesis.

2. Experiment I: Inversion

Looser and Wheatley (2010), observed that the individual processing of the features of the face is not sufficient for animacy perception, but that rather a more holistic approach is necessary. The purpose of this experiment was to reduce participants' holistic processing of faces in order to determine how important holistic processing is for animacy judgments of faces. In order to reduce holistic processing, the faces in the continuum were inverted. Previous experiments have shown that face inversion makes the task of recognition more difficult (Goldstein, 1965). Studies have suggested that upright faces are processed at a global level, while an inverted face is analyzed element by element (Farah et al., 1995; McKone, 2004).

2.1. Method

2.1.1. Subjects

Twenty-two subjects (male and female between the ages of 19-60 years) were selected from the subject pool at the Brain and Cognitive Sciences Department at MIT. The subjects had normal or corrected-to-normal vision when completing the tasks. After completing the experiment, subjects were paid for their participation.

2.1.2. Stimuli

Four morph continua between real human and doll faces were created using the Abrosoft FantaMorph morphing software (http://www.fantamorph.com/index.html). The morph continua of doll and human faces ranged from 0% human (100% doll) to 100% human in increments of 10%, resulting in a total of eleven images for each morph continuum. Four female adult Caucasian female faces and adult-like female doll faces were morphed together. The faces were neutral in expression. Microsoft Photoshop was used to remove the ears and hair, but the remaining shapes of the faces were not altered. The size of the faces was normalized so that the nose, eyes, and mouth were approximately the same size between faces. A gray background was used for all the images in each morph continuum, and all images were presented in grayscale. The morph images were 256 by 256 pixels. Figure 1 shows the four morph continua.

A.

B.

C.

D.

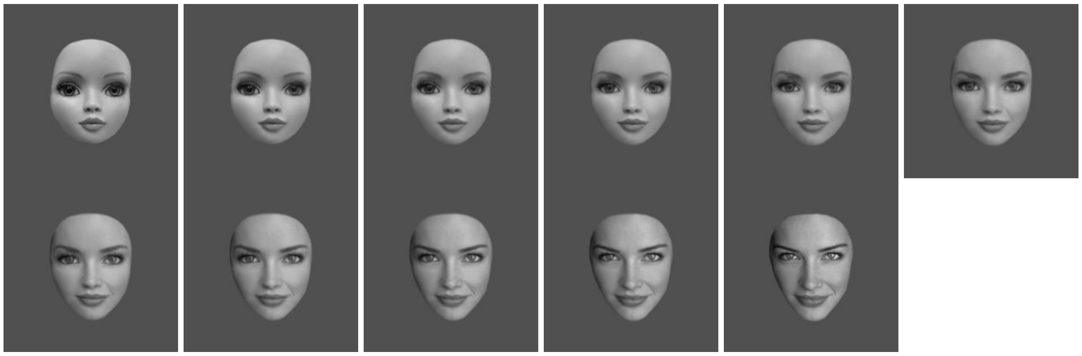


Figure 1- (A) Morph continuum 1 with eleven faces from 0% human to 100% human in increments of 10%. (B) Morph continuum 2. (C) Morph continuum 3. (D) Morph continuum 4.

2.1.3. Procedure

Upright and inverted conditions were tested in separate experimental runs, with the upright condition always tested first. Two morph continua were used for each orientation condition, and images within the two morph continua were presented in a random order. For an initial group of 18 of the 22 subjects who performed this experiment, the following procedure was used. The subjects were instructed to rate the morph images on a scale from 1 to 7, with 1 being completely doll and 7 being completely human. In the upright trials, the subject was presented with upright morphs from two morph continua. In the inversion trials, participants were presented with upside-down face images from the remaining two morph continua out of the four. The pair of morph continua used for the upright versus inversion trials was varied between subjects. The images were presented on the screen of an Apple Macbook Pro laptop for 1 second and then the subject rated the face by pressing a number key (1-7) on the keyboard. The MATLAB Psychtoolbox software (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) was used to display the images and record the responses of the subjects. Participants were given the following directions: “We morphed faces between real human and doll faces to create a continuum from doll to human. You will see a face on the screen, and I would like you to rate on a scale from 1 through 7, how animate-looking you think the face appears, with 1 being completely doll, and 7 being completely human”. Each of the eleven faces in the continuum was repeated ten times, yielding a total of 220 upright trials and 220 inversion trials.

For the remaining 4 subjects, the following procedure was used. Individual face images were drawn from the same two morph spectra (morph continua 2 and 4) for both the upright and inverted trials. The images from the morph continua were presented in a random order on a computer screen for 1 second and the subject was instructed to rate the face as either ‘doll’ or ‘real’ (0 = doll, 1 = real). Six faces were used in practice trials before beginning the upright and inverted experiments, so that the subject would become accustomed to the task. The faces used in the practice trials were not from the morph continua used in the actual trials. The participants were told the following: “We morphed faces between real human and doll faces to create a continuum from doll to human. When you see a face on the screen, please rate whether you think the face looks like a doll, by pressing 0, or whether you think it looks more like a real human, by pressing 1”. Similar to the procedure described earlier, upright and inverted conditions were tested in separate experimental runs, with the upright condition always tested first (Figure 2). The purpose for testing a small group of subjects with the modified procedure is two-fold. First, it allowed a comparison between results obtained using two different rating schemes, with one being a continuum (1 to 7) and the other being a two-alternative-forced-choice (doll=0 or human=1). The latter scheme was used in the temporal experiment described in section 3.1.3. Second, the strength of the cues used to judge animacy may vary across the four morph continua. If different morph continua are used for upright and inversion trials, a difference in performance between the two conditions could be due to the particular morph continua used, rather than the inversion per se. The second procedure above allowed a direct comparison between performance on upright and inverted face images from the same morph continua.

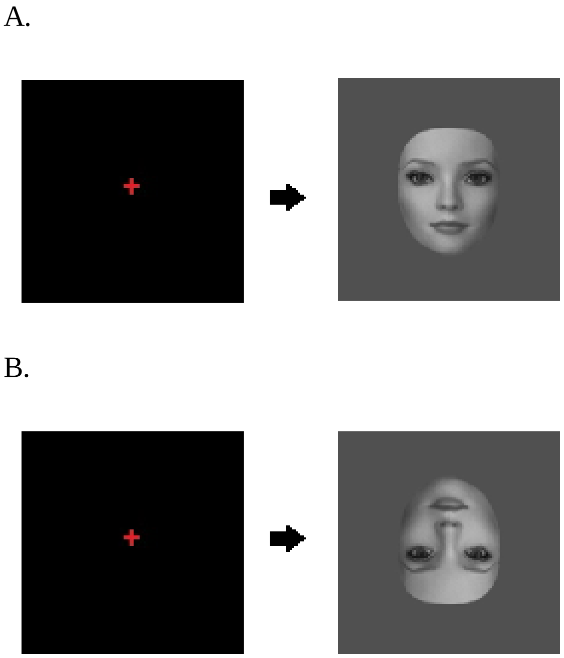


Figure 2- (A) An example of an upright animacy trial, where the subject is shown a morph image and asked either to rate the face on a scale of 1-7, with 1 being completely doll and 7 being completely human, or to indicate whether the face looks like a doll or human. A red fixation cross was placed on the screen before the image so that the subject could focus their attention to that region. (B) An example of

an inverted animacy trial, where the subject is presented with an inverted morph image on the screen and asked to make a similar judgment.

2.2. Results

The subjects' responses to each of the morph continua were analyzed for the upright and inverted cases. Given the possibility that the strength of the cues used for animacy judgment could vary across the four morph continua, we initially analyzed the data separately for the four continua. Differences in the strength of the cues could be captured in the variability of animacy judgments across subjects, for the different morph continua. Figure 3 shows the average animacy ratings for the upright images for each of the 18 subjects who specified their ratings on a scale from 1 to 7. Data for each of the four morph sequences is shown in separate graphs, as indicated by the titles above the graphs. The data for each subject is displayed with a distinct color, and data for each individual subject appears in two of the four figures. The average is calculated across the 10 repeats of each face image, for each morph level. Figure 4 shows the average ratings for each of the 18 subjects who used the 1-7 rating scale, for the inverted faces, again shown separately for each of the morph continua. For morph continua 1-3 the variability between subjects for inverted faces is noticeably greater than what is observed for the upright face morphs. For the fourth morph continuum, the variability between subjects does not appear to have increased substantially with inversion. It can be seen from this data that although there is variability across subjects, they are clearly able to perform the task, as there is a large difference between ratings at the two ends of the spectrum from human to doll.

The increase in variability between subjects' ratings for the inverted faces is an indication that participants had more trouble assigning animacy ratings for the face images when they were inverted. The fourth morph continuum, in comparison, may have a single featural characteristic or multiple featural cues that can be reliably used by the subjects to determine animacy. As a consequence, in the inversion condition, the disruption of the configural information of the face may not have as much of an effect on animacy perception of the face.

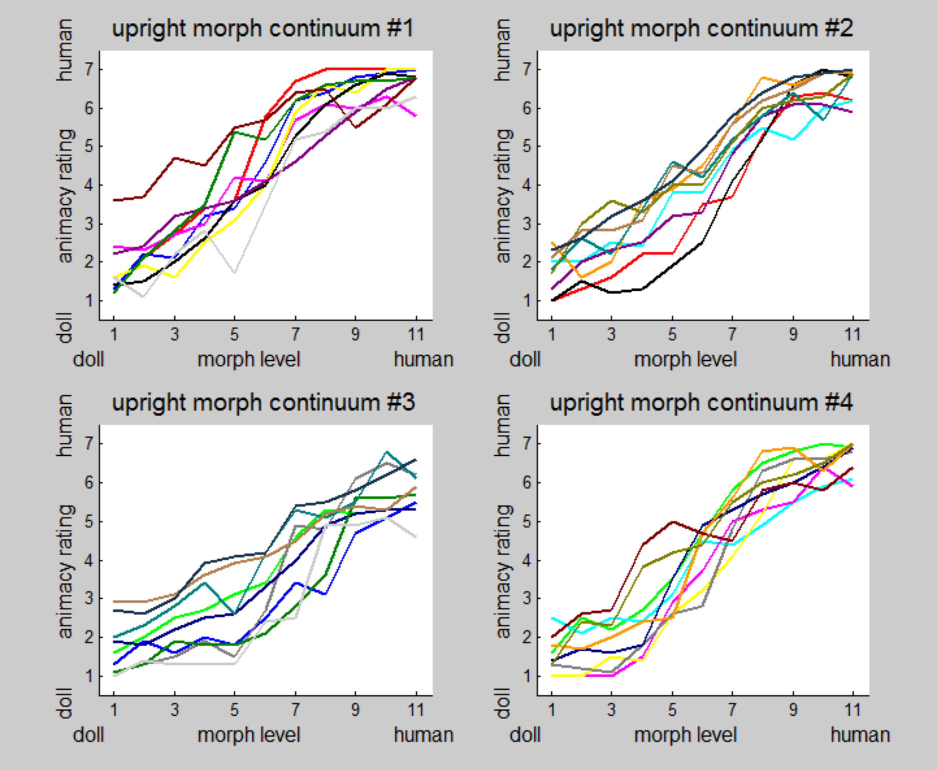


Figure 3- Graphs showing the average animacy ratings of 1-7 for morph images 1-11, for the four different morph continua in the upright orientation. Data for each subject is displayed with a distinct color, and data for each individual subject appears in two of the four figures. Each of the faces 1-11 were repeated 10 times and the ratings were averaged.

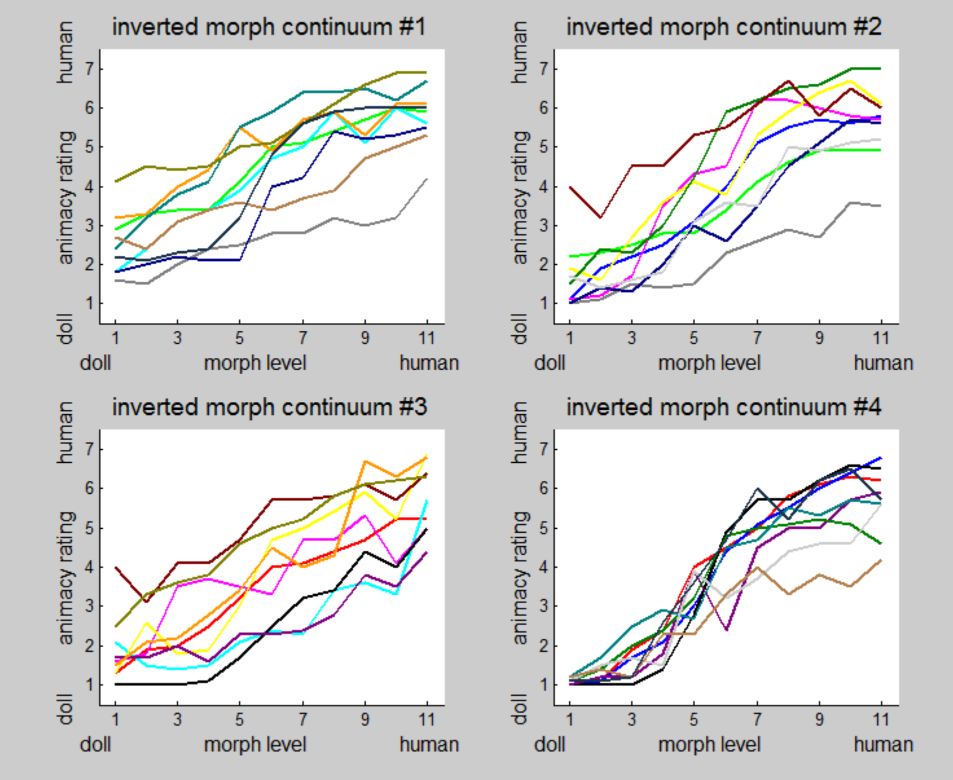


Figure 4- Graphs showing the average animacy ratings of 1-7 for morph images 1-11, for the four different morph continua in the inverted orientation. Data for each subject is displayed with a distinct color, and data for each individual subject appears in two of the four figures. Each of the faces 1-11 were repeated 10 times and the ratings were averaged.

In order to quantify the variability of the subjects' ratings for each of the morph continua in the upright and inverted cases, we calculated the standard deviations of the ratings, across subjects for each animacy level in each continuum. Figure 5 shows the average rating across subjects, with error bars indicating the standard deviations of the average ratings for each upright morph image in each continuum. Rating variability is generally larger in the center of the morph continua and smaller at each end. This result is expected, because the stimuli at the center are the most ambiguous. Figure 6 is a graph of the average ratings across subjects for each of the morph levels of the inverted faces for each continuum, with error bars showing their standard deviations. Standard deviations in the ratings are larger for the inverted condition than the upright one for morph spectrums 1-3. For the first morph continuum, the average of the standard deviations for the middle region (levels 4-8 in the spectrum) was 0.764 for the upright images, and 1.06 for the inverted images (see Tables 1 and 2). The standard deviations of the average ratings for inverted and upright faces in morph continuum 1 were compared using t test and it was found that the standard deviations are significantly different (p=0.0016). In the second morph continuum the morph levels 4-8 for the upright faces have an average standard deviation of 0.736, compared to 1.17 for the inverted faces. Using a t-test it was determined that the standard deviations in the inverted and upright trials were also significantly different (p=1.02*10^-5). For the third morph continuum, the upright faces have an average standard deviation of 0.923 for levels 4-8, and the inverted faces have an average of 1.14; a significant difference was found between the standard deviations (p=0.00048). For the fourth morph continuum, the inverted faces had a slightly lower standard deviation average than the upright faces (a difference of 0.664), which is contrary to what was found for the other morph sequences. It may be that there is an important featural cue that was used in this case for animacy judgment, which was more easily attended to when configural information was removed. The average of the standard deviations between levels 4-8 in the morph continua for all of the upright morph continuum ratings was 0.924, and for the inverted morph continua it was found to be 1.053 (Table 3). No significant difference between the standard deviations for the average ratings between upright and inverted was found for this morph continuum (p=0.91). Figure 7 provides a direct graphical comparison between the standard deviations of the animacy ratings of the upright and inverted faces, across the morph level, highlighting the greater variability for intermediate morph levels and for the inverted faces in general.

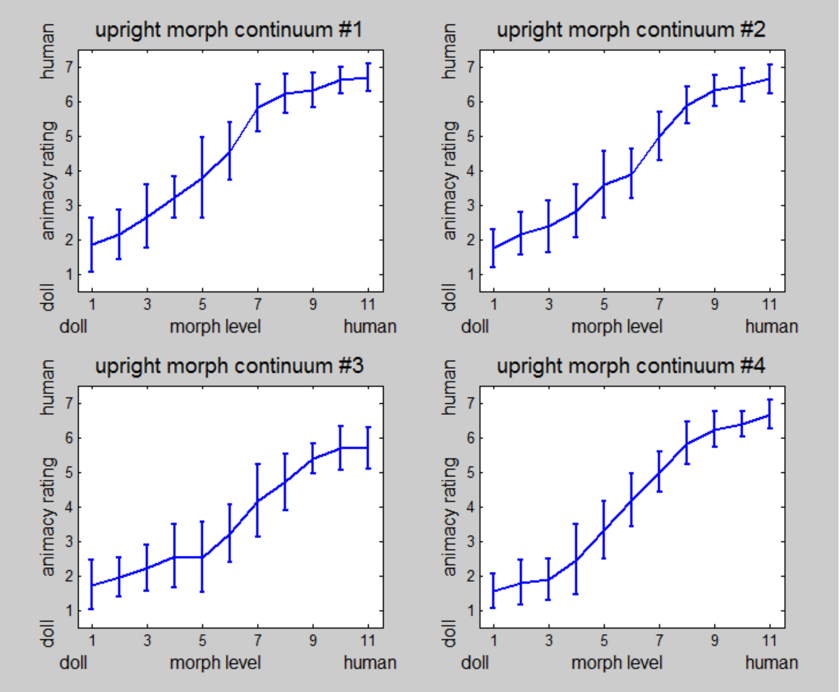


Figure 5- Graphs showing the ratings of 1-7 for each of the morph continua 1-4 in the upright orientation averaged across subjects. The standard deviations of the ratings for each of the morph levels 1-11 are shown by the error bars on the graph.

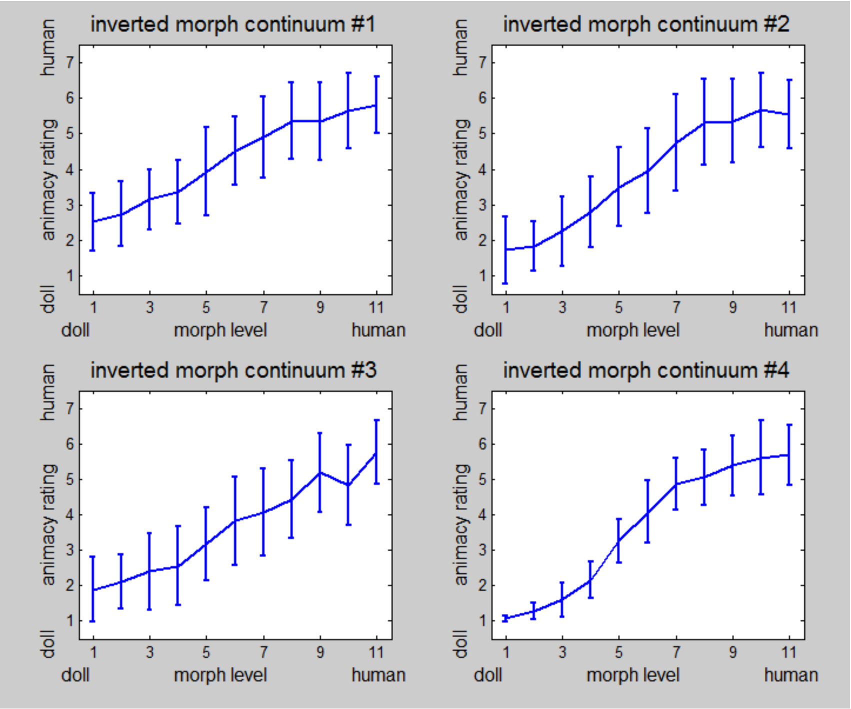


Figure 6- Graphs showing the ratings of 1-7 for each of the morph continua 1-4 in the inverted orientation averaged across subjects. The standard deviations of the ratings for each of the morph levels 1-11 are shown by the error bars on the graph.

Upright Error	**Morph 1**	**Morph 2**	**Morph 3**	**Morph 4**
1	0.4	0.4216	0.5947	0.4243
2	0.3905	0.477	0.6431	0.3655
3	0.4969	0.4664	0.4353	0.5268
4	0.5652	0.5231	0.8115	0.6325
5	0.6671	0.7032	1.0549	0.5916
6	0.8233	0.7167	0.8253	0.7563
7	1.1613	0.9589	1.0195	0.8363
8	0.603	0.7775	0.9028	1.0332
9	0.9055	0.753	0.6754	0.6078
10	0.7143	0.6167	0.5725	0.6364
11	0.778	0.5457	0.7018	0.4902
Average of levels 4-8	0.76398	0.73588	0.9228	0.76998

Table 1- The raw values of the standard deviations of levels 1-11 in the upright morph continua for the averaged animacy ratings across subjects. The standard deviations across levels 4-8 (30%-70% human) were averaged for each of the morph continua 1-4.

Inverted Error	Morph 1	Morph 2	Morph 3	Morph 4
1	0.7984	0.9644	0.8946	0.8378
2	1.0595	1.0454	1.1402	1.038
3	1.0876	1.1674	1.1054	0.8318
4	1.0783	1.2036	1.0933	0.7828
5	1.1325	1.3519	1.2186	0.7367
6	0.9493	1.1991	1.2518	0.8829
7	1.2319	1.1062	1.0392	0.6124
8	0.8918	0.9842	1.1114	0.5028
9	0.8457	0.9697	1.0735	0.4899
10	0.9126	0.6874	0.7541	0.2333
11	0.8012	0.9563	0.9103	0.0833
Average of levels 4-8	1.05676	1.169	1.14286	0.70352

Table 2- The raw values of the standard deviations of levels 1-11 in the inverted morph continua for the averaged animacy ratings across subjects. The standard deviations across levels 4-8 (30%-70% human) were averaged for each of the morph continua 1-4.

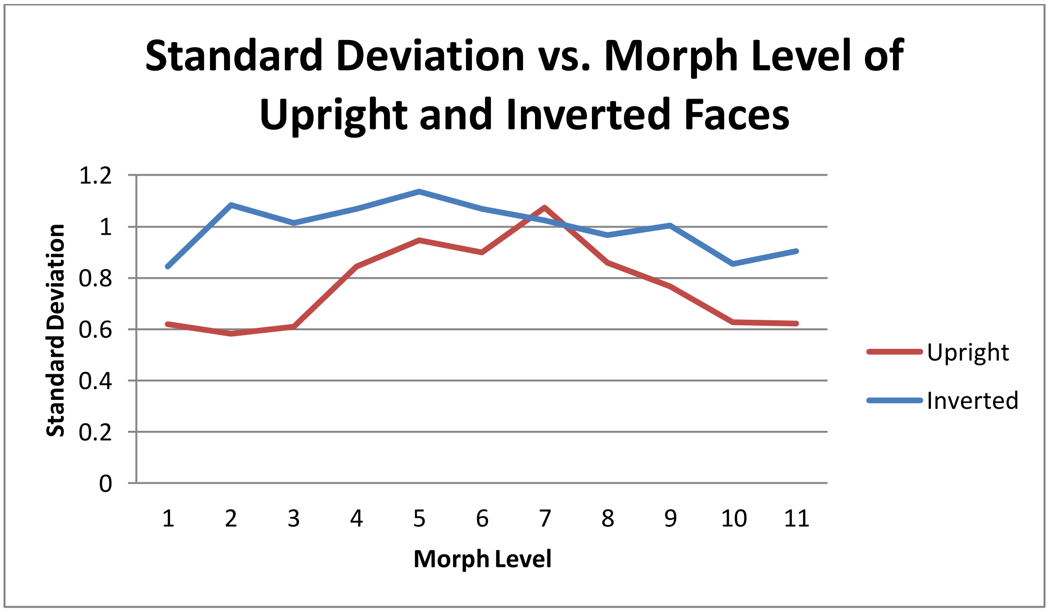


Figure 7- The average standard deviation values for each of the morph levels 1-11 in the upright (red) and inverted (blue) orientations, averaged across morph continua 1-4 and across all of the subjects.

Upright Error	**Morphs 1-4**
1	0.6208
2	0.5832
3	0.6106
4	0.8435
5	0.947
6	0.8987
7	1.0737
8	0.8582
9	0.7677
10	0.6274
11	0.6213
Average of Levels 4-8	0.92422

Table 3- The raw values of the standard deviations across upright face morph levels 1-11 for average ratings of all subjects for morph continua 1-4.

Inverted Error	**Morphs 1-4**
1	0.8437
2	1.0828
3	1.0127
4	1.0697
5	1.136
6	1.0675
7	1.024
8	0.9662
9	1.0033
10	0.853
11	0.9047
Average of Levels 4-8	1.05268

Table 4- The raw values of the standard deviations across inverted face morph levels 1-11 for average ratings of all subjects for morph continua 1-4.

The animacy ratings of all of the subjects for each of the morph spectrums were averaged, for the upright and inverted cases separately. Figure 8 shows the rating averages, with upright in blue and inverted in red. It can be seen that at the human-end of all of the curves, the completely human upright faces were rated as being more human-like than the completely human inverted faces in all of the morph continua. A significant difference between the ratings for the completely human upright and inverted faces was found (p= 0.0004), meaning that the upright faces were perceived as being significantly more human-like than the inverted faces. This suggests that some sort of holistic processing may be needed to make a more accurate judgment on whether a face is human or doll. It can be seen, however, that there is no consistent difference at the doll end. When the averages of all of the morph continua are combined, as shown in Figure 9, the same pattern can be seen in the animacy curves, with the only differences being at the human end. Figure 10 shows the two curves for upright and inverted faces with error bars indicating the standard deviations for each of the morph levels. It can be observed in this graph that the spread of the ratings for inverted faces is similar across morph levels, while for upright faces the spread is smaller at the two ends of the spectrum.

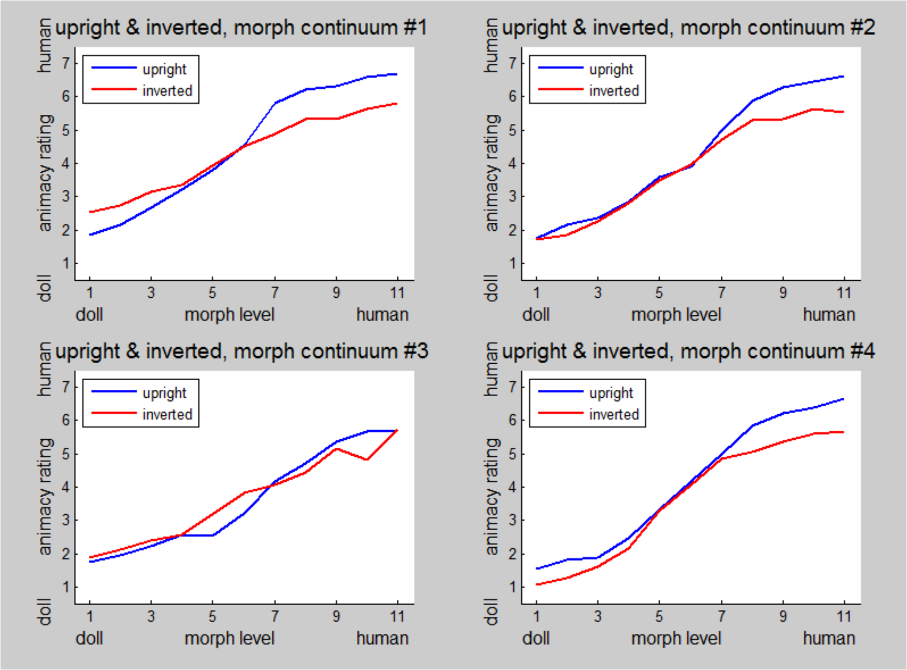


Figure 8- Graphs showing the animacy rating averages across all subjects for morph continua 1-4 in the upright (blue) and inverted (red) orientations superimposed.

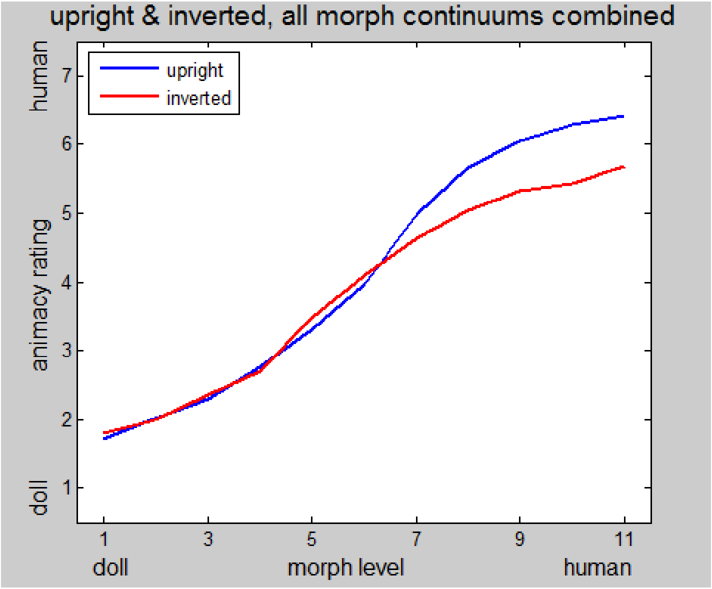


Figure 9- Graph showing the average ratings of all of the morph continua combined in the upright (blue) and inverted (red) orientations.

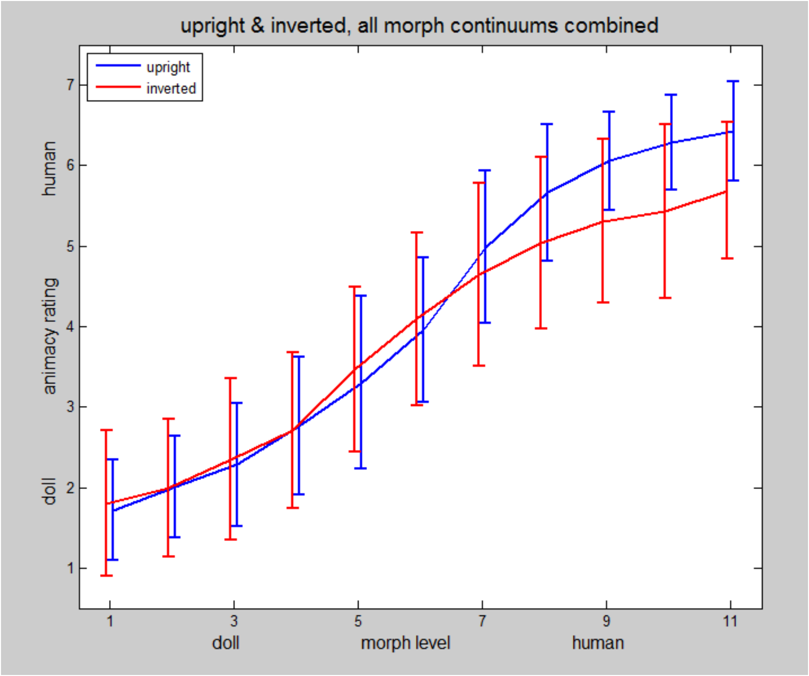


Figure 10- Graph showing the average ratings of all of the morph continua combined in the upright (blue) and inverted (red) orientations with error bars showing standard deviations. The error bars of the curves are slightly offset on the graph, so that the standard deviations for each morph level 1-11 can be compared between upright and inverted orientations.

The point of subjective equality (PSE) is the percentage (morph level) at which the subject perceives a face in the morph continuum as half human, half doll. The Palamedes toolbox, written in MATLAB, (Prins and Kingdom, 2009) was used to fit a psychometric curve onto the ratings given by the subjects for each of the face images in the continuum. This psychometric curve can be used to determine where the PSE is located and also the certainty of the subjects' ratings, given by the slope. The slopes of the animacy curves give us insight into the certainty of the subjects' ratings. When the subjects' individual PSE values for their upright and inverted animacy ratings were compared, no significant difference was observed (p= 0.55); however, it was found that the slopes for the inverted animacy ratings were significantly smaller than for the upright animacy ratings (p= 0.024), suggesting that the certainty in the subjects' ratings decreased with inversion. The averages of the ratings for the upright faces can be seen to have a steeper slope than for the inverted faces, as seen in Figures 11 and 12. The average ratings across subjects for the four morph continua are shown superimposed in different colors, for both the upright and inverted conditions. The PSEs of these curves are also plotted in Figure 13; the PSEs of these curves do not appear to differ in the upright and inverted cases, while there is a clear difference in the slopes. This trend is also observed in the psychometric curves for the averages of the upright face ratings and inverted face ratings across morph continua, where the upright curve has a slope of 6.52, and the inverted curve has a slope of 3.98 (Figures 14 and 15). The smaller slope in the curves of the data for the inverted faces suggests that the subjects were less certain when assigning animacy ratings to the inverted faces. The holistic configuration of the face may provide the subjects with more certainty when making an animacy judgment. When comparing the morph continua (Figure 1), it can be seen that morph continuum 1 appears to be generally more human-like and morph continuum 3 is more doll-like. Morph continua 2 and 4 were used in the last part of this experiment because subjects had similar average performance for these two morph continua and there was a large difference between the ratings between pure human and pure doll faces in the spectrums.

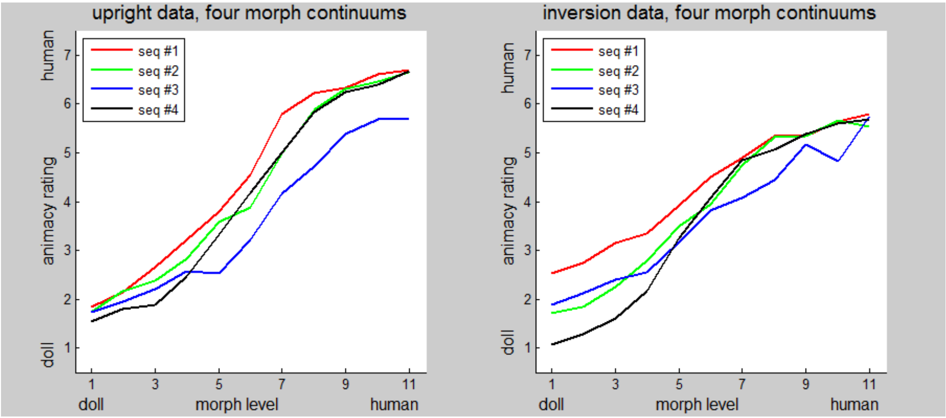


Figure 11- The average ratings across subjects for the four morph continua in the upright and inverted orientations superimposed, with different colors signifying different morph continua (continuum 1=red, continuum 2=green, continuum 3=blue, continuum 4=black).

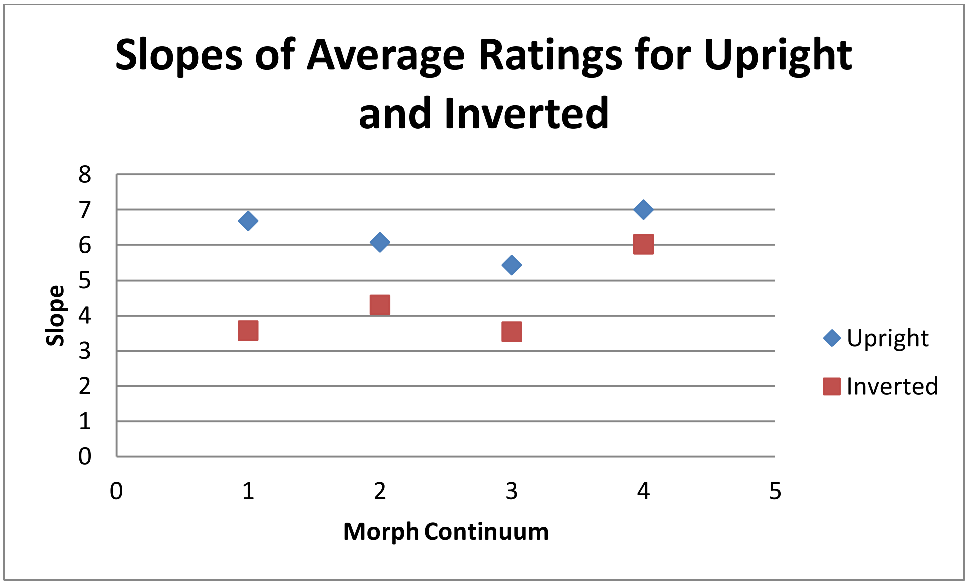


Figure 12- The slopes of the psychometric curves fit to the average animacy ratings for upright (blue) and inverted (red) orientations across subjects for each morph continuum 1-4.

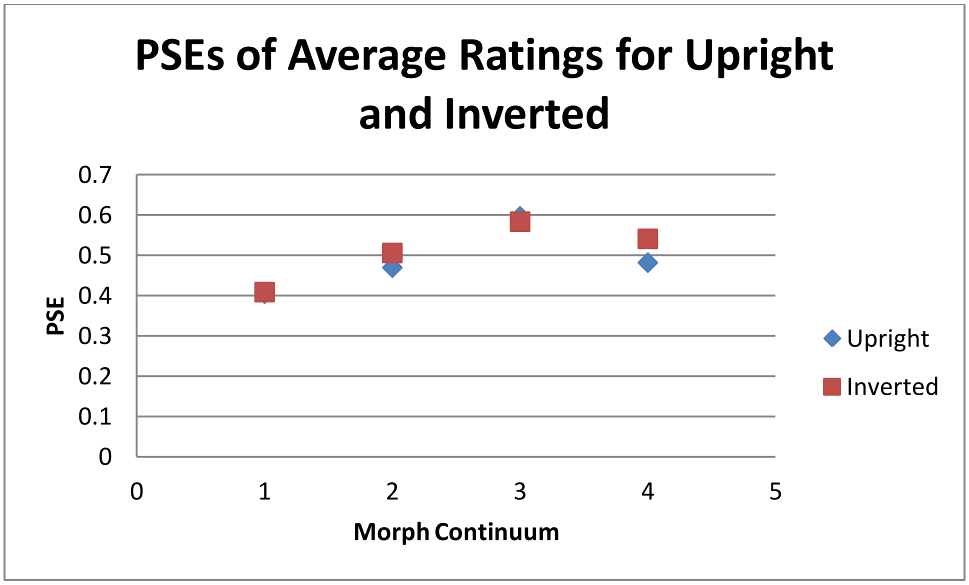


Figure 13- The PSEs of the psychometric curves fit to the average animacy ratings for upright (blue) and inverted (red) orientations across subjects for each morph continuum 1-4.

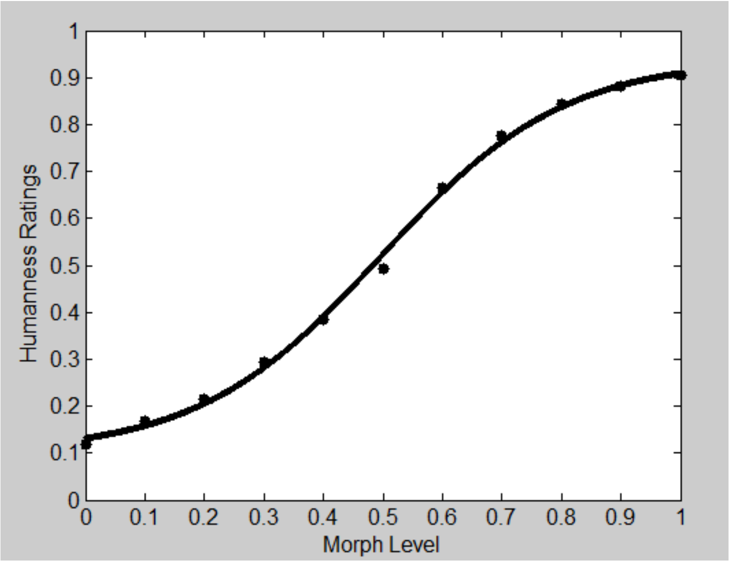


Figure 14- The psychometric curve for the average animacy ratings of 1-7 across all subjects and all of the upright morph continua normalized to a range of 0-1. The psychometric curve has a slope of 6.519, and has a PSE of 0.483.

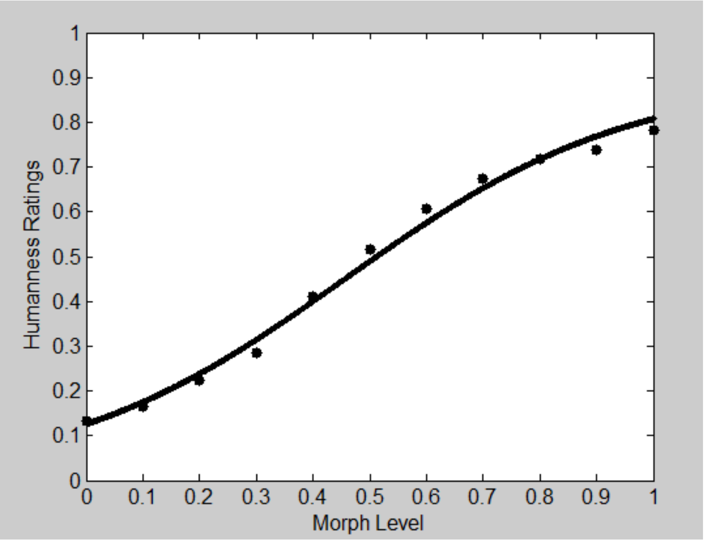


Figure 15- The psychometric curve for the average animacy ratings of 1-7 across all subjects and all of the inverted morph continua normalized to a range of 0-1. The psychometric curve has a slope of 3.983, and has a PSE of 0.5121.

The differences in both the slope and PSEs between the four morph continua observed in the first part of the inversion experiment, led us to the second part of the experiment, where the morph continua in the upright and inverted trials were kept the same, so that a direct comparison between the upright and inverted trials could be made for each subject. In the last part of the inversion experiment, the morph continua 2 and 4 for the upright and inverted experiments were used, and data was collected from four subjects. The animacy ratings by the four subjects for each of the images within morph continua 2 and 4 were averaged for each level from 0% human to 100% human. The average animacy ratings of the subjects for morph continua 2 and 4 in the upright and inverted orientations were plotted (Figures 16 and 17). A psychometric curve was fit to the averages for both the upright and inverted experiments (Figures 18 and 19). For the upright faces, it was determined that the PSE across subjects is 49.22%, and for the inverted faces there is a slightly lower PSE of 40.24% (Table 5a). The certainty of the subjects' ratings of the faces was compared between the upright and inverted faces, as well. The slope of the psychometric curve for the upright faces was found to be 19.08 while for the inverted faces it was found to be 14.21 (Table 5b). A larger slope for the upright faces suggests that the subjects were more confident in their animacy ratings of the morph continua when they were presented with upright versus inverted faces. The PSEs and slopes of the ratings were compared between the upright and inverted experiments for each of the subjects. There was a strong trend towards the PSEs being different for upright vs. inverted faces (t-test, p=0.054) and a weaker trend towards a difference between the two in slope (t-test, p=0.13). With four subjects in the analysis, we most likely did not have a lot of power to detect differences between ratings for upright and inverted faces. For the slopes, it was also found that there was no significant difference between the upright and inverted experiments. The slopes for the inverted psychometric curves were found to have an average value of 16.98 which is smaller than for the upright curves with a slope of 26.09.

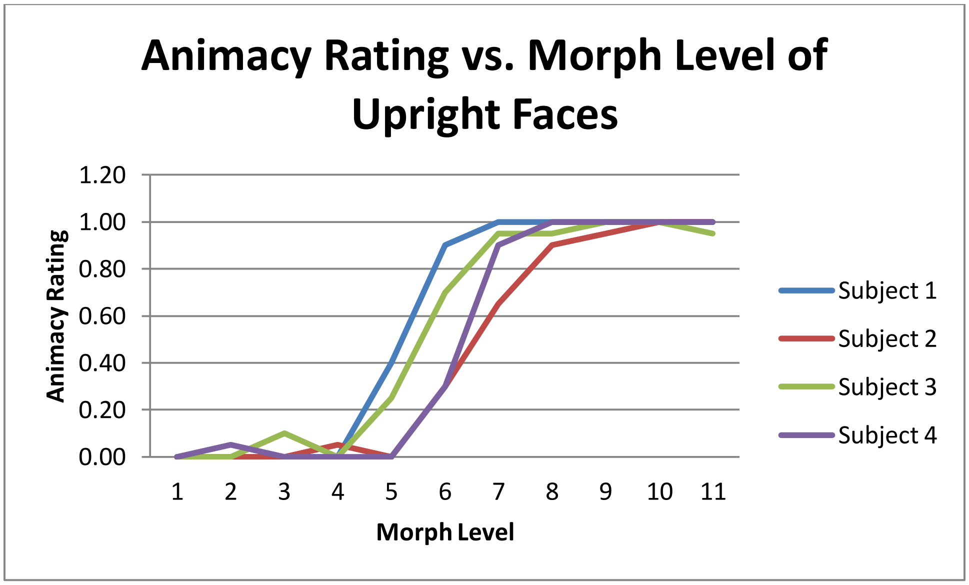


Figure 16- The average animacy ratings (0 and 1) for morph continua 2 and 4 in the upright orientation.

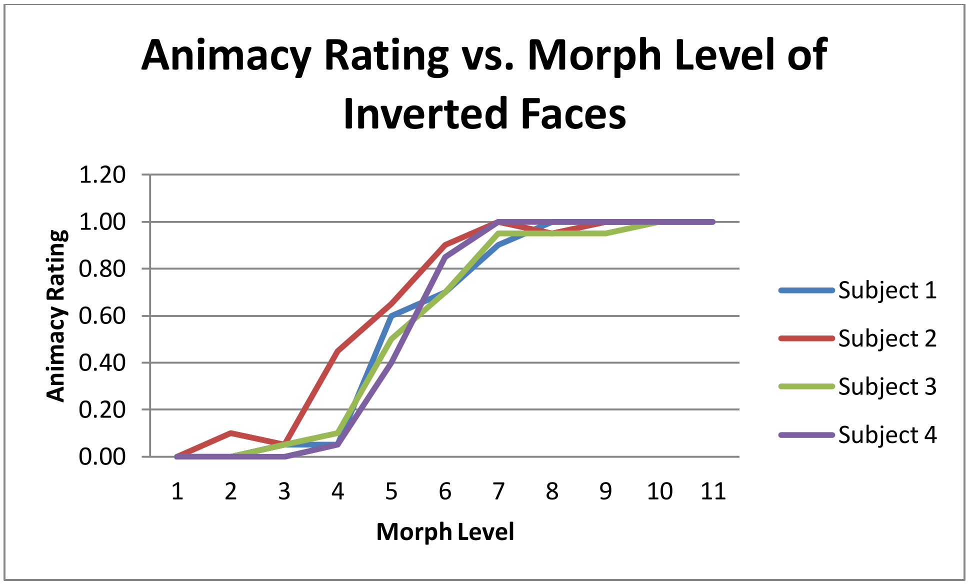


Figure 17- The average animacy ratings (0 and 1) for morph continua 2 and 4 in the inverted orientation.

A.

Upright	PSE	Slope
Subject 1	0.4192	32.013
Subject 2	0.565	15.6217
Subject 3	0.4585	22.4149
Subject 4	0.5292	34.3109
Average	0.492975	26.09013

B.

Inverted	PSE	Slope
Subject 1	0.42	15.1542
Subject 2	0.3395	12.6382
Subject 3	0.4212	15.1061
Subject 4	0.4198	25.0285
Average	0.400125	16.98175

Table 5- The PSEs and slopes of the psychometric functions fit to the average ratings of morph continua 2 and 4 in the upright and inverted orientations.

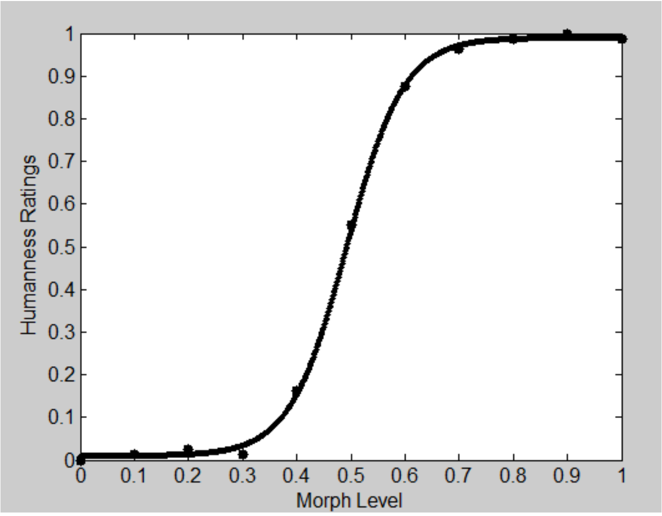


Figure 18- Psychometric curve fit to the average animacy ratings across all four subjects for morph continua 2 and 4 in the upright orientation. The curve has a slope of 19.08 and a PSE of 0.4922.

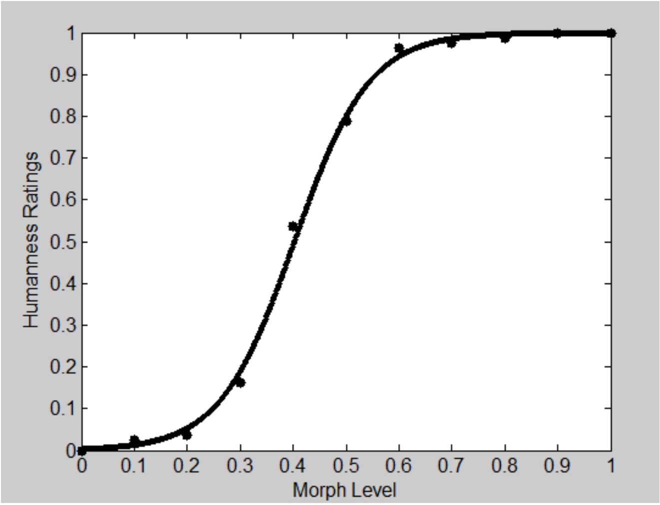


Figure 19- Psychometric curve fit to the average animacy ratings across all four subjects for morph continua 2 and 4 in the inverted orientation. The curve has a slope of 14.21 and a PSE of 0.4024.

3. Experiment II: Temporal

In previous studies, it was observed that face recognition can occur by 90 ms (Veres-Injac and Schwaninger, 2009). In this experiment we limited the exposure duration of morph face images between dolls and real humans to gain insight into the amount of time necessary to form a perception of animacy. We also compared the time-course of face recognition with the time-course of animacy perception. It was predicted that simpler cues such as eye size and skin texture may be used in animacy judgment, which may require less time for face-processing than the cues used in face recognition. It may also be possible that animacy judgment relies more on initial holistic information than detailed featural information in comparison to face recognition, which may require finer details. It is expected that more detailed analysis of features such as the fine lines of the face require more time than holistic processing. Pilot experiments were run with a preliminary experimental design using the durations of 33 ms, 120 ms, 220 ms, and self-paced, in order to choose reasonable exposure durations for testing. The exposure durations that were used in the final experimental design were 33 ms, 84 ms, and 2.02 seconds.

3.1. Method

3.1.1. Subjects

Thirty-seven subjects (male and female between the ages of 19-60 years) were selected from the subject pool at the Brain and Cognitive Sciences Department at MIT. The participants that were selected had normal or corrected vision during the experiment. The subjects were paid for their participation in the research after completion of the experiments. In the preliminary face recognition and temporal experiments, data was collected from twenty-seven subjects, and in the final experimental design, data was recorded from a total of eleven subjects.

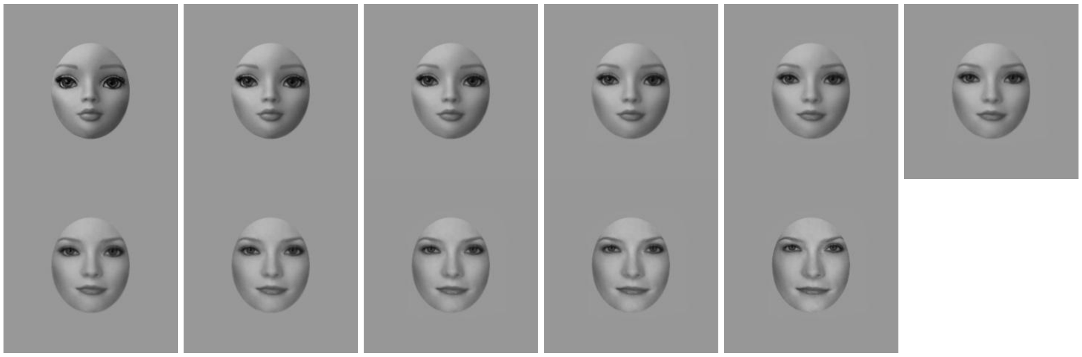
3.1.2. Stimuli

Ten adult faces with neutral expressions were used in the face recognition task. Microsoft Photoshop was used to standardize the external shape of the faces using a template. A gray color for the background was used for all of the images for the recognition task and in the morph continuum for the animacy task. The images were presented in grayscale. The temporal animacy perception trials used the same faces in the four morph continua that were used in the inversion experiments. The only difference was that the external shape of the face stimuli in the continuum was normalized by using an oval face template.

A mask was presented immediately after each face stimulus, because retinal cells continue to respond approximately 60 ms after the visual stimulus is no longer present on the screen (Keysers and Perrett, 2002). A 256 by 256 matrix of random values between 0 (black) and 1 (white) was generated to create a mask stimulus.

A.

B.



C.

D.

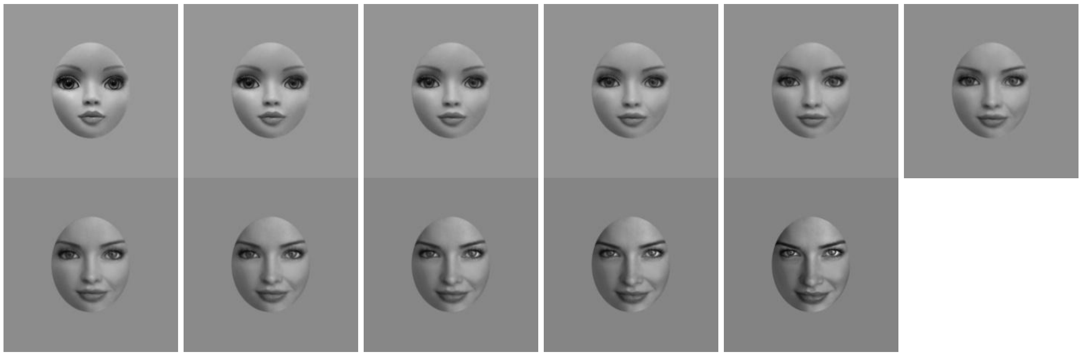


Figure 20- (A) Morph continuum 1 with eleven faces 0% human to 100% human in increments of 10% with face shape standardized with a template. (B) Morph continuum 2 (C) Morph continuum 3 (D) Morph continuum 4

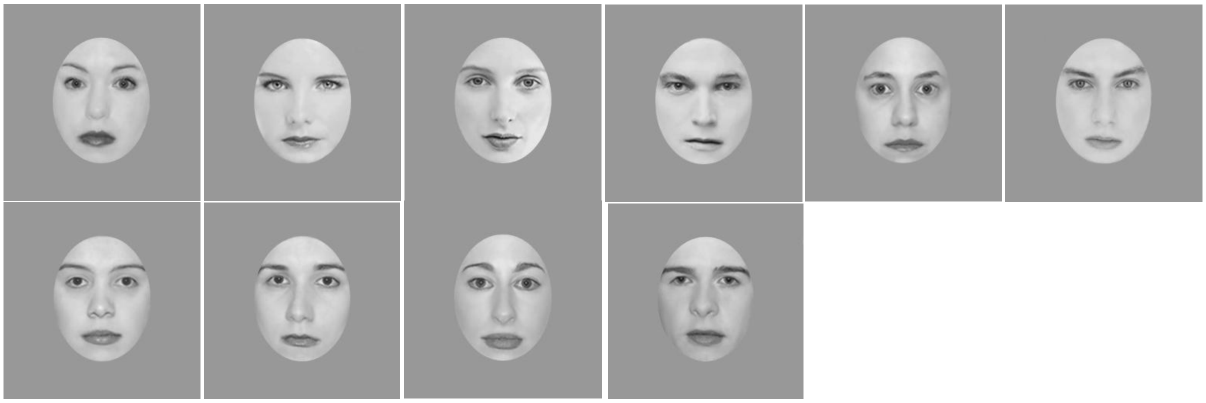


Figure 21- Male and female real adult human faces used in the face recognition task. The face shape was standardized using an oval template.

3.1.3. Procedure

A. Preliminary Experimental Design

Preliminary face recognition and animacy perception temporal trials were conducted using exposure durations of 33 ms, 120 ms, 220 ms, and self-paced. In the face recognition task, the subject was presented with a face on the screen for 33 ms, 120 ms, or 220 ms. Subjects were first presented with a face from Figure 21 followed immediately by a mask image before being shown two images side by side on the screen, with either the left or right image being the previously presented image and the remaining image a randomly chosen face from a pool of faces being used in the experiment (Figure 22). Figure 22 shows an example of a trial, with the left image as the correct answer. The face images were repeated four times for each of the exposure durations (33 ms, 120 ms, and 220 ms), making a total of 120 trials. For the temporal animacy perception experiment, the same four morph continua that were used in the inversion experiment were presented to the subjects. In the base trials, we presented the morph images and asked the subjects to rate the images as 'real' or 'doll' on the keyboard (0= doll, 1= real) without restricting viewing time. Each of the face images in two of the morph continua were repeated three times, making a total of 66 trials. In the temporal trials, the remaining two morph continua were presented. Each image was shown four times for each of the three exposure durations (33 ms, 120 ms, and 220 ms), making a total of 264 trials. It was determined that there was no significant difference in animacy perception between 120 ms, 220 ms, and self-paced trials. It was found, however, that there was a significant difference between 33 ms and self-paced. From these results, it was decided to choose an exposure duration between 33 ms and 120 ms for testing animacy perception and face recognition. Because the ratings at 120 ms and 220 ms showed no significant difference from the self-paced trials, the new experimental design removed these two exposure durations, and added the exposure duration of 84 ms. Instead of self-paced, in order to save time, we limited the exposure duration to 2.02 seconds and removed any time in between the presentation of the different frames. The number of trials per subject was also increased to reduce the noise in the data. All four morph continua were presented at the three different exposures of 33 ms, 84 ms, and 2.02 seconds.

B. Final Experimental Design

In the final temporal experimental design the exposure durations of 33 ms, 84 ms, and 2.02 seconds were used. The face recognition task in the preliminary experiment was conducted using the new exposure times. The subjects completed a practice face recognition task before the actual trials, using six faces. In the practice trials, three of the images were presented at 84 ms, and the other three at 2.02 seconds; the 33 ms exposure time was not used in the practice trials, in order to ensure that the subjects understood the task. In the new face recognition experimental design, the subjects indicated when they were ready to begin a trial by pressing the spacebar, and the stimulus appeared a fixed time after the spacebar was hit. The verbal instructions were the following: "This task will test your ability to recognize faces. To begin a trial, you should press the spacebar when you see a red cross on the screen. You will then see a face on the screen, sometimes for a very short period of time, and then you will be presented with two faces afterwards. I want you to pick the face that you think you had just seen on the screen, by either pressing 1 for left or 0 for right." If the participant received a score of lower than fifty percent correct in face recognition, they had to complete the task once again until they reached fifty percent correct or higher in accuracy.

For the temporal animacy perception experiment, the same four morph continuum images were presented. Practice trials were also given so that the subject could familiarize themselves with the experimental task. Six practice trials were run using different faces than the ones found in the four morph continua. In the new experimental design for animacy perception, the subjects were allowed to begin the trials when they were ready by pressing the spacebar. If the subjects received below a fifty percent correct on detecting completely human or completely doll faces, the subjects were required to repeat the practice trials using a set of six new face images. The subjects were given the task to rate the face image on the screen as either ‘doll’ or ‘real’ using the keyboard (0= doll, 1= real), with the verbal instructions: “We morphed faces between real human and doll faces to create a continuum from doll to human. When you see a red cross on the screen you can begin the next trial by pressing the spacebar. You will be seeing a face on the screen, sometimes for a very short period of time, but I would like you to try your best to rate the face as either ‘doll’, where you press 0, or ‘real’, where you press 1 when you see a red square on the screen.” The images were shown for 33 ms, 84 ms, or 2.02 seconds. Each of the face images in the four morph continua were repeated four times for each of the exposure durations 33 ms, 84 ms, and 2.02 seconds making a total of 528 trials. Both the order of the face images and the durations they were shown for were randomized during the experiment.

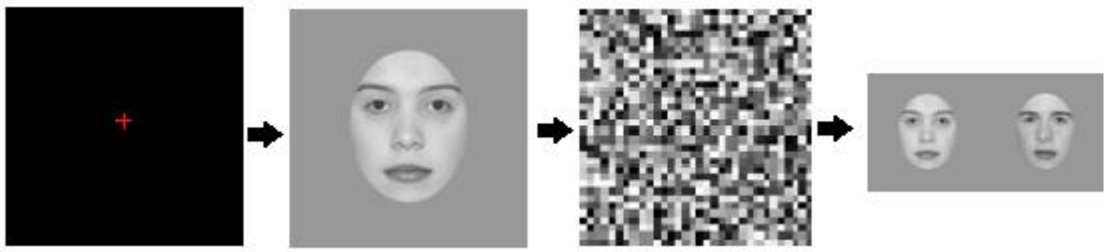


Figure 22- Example of a sequence of images used in one trial of the face recognition task. The subjects began a trial by first pressing the spacebar when a red cross appeared on the screen. The first face image was presented for a time of 33 ms, 84 ms, or 2.02 s. Immediately afterwards, a mask image was shown on the screen for 1 second, and then two faces were placed side by side on the screen for the subject to pick the face they think they had seen. In this trial, the correct answer would be the image to the left.

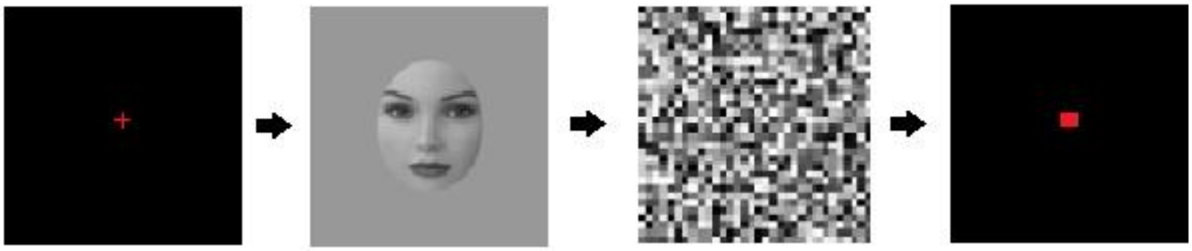


Figure 23- Example of a sequence of images from a trial in the animacy task. The subject began a trial by pressing the spacebar when a red cross appeared on the screen. The morph images from continua 1-4 were shown on the screen for 33 ms, 84 ms, and 2.02 s. Immediately after the subject was presented with

an image for a certain amount of exposure time, a mask image was shown on the screen for 1 second. The subjects were asked to rate the image as either human by pressing 1, or doll by pressing 0 when a red square appeared on the screen.

3.2. Results

In this experiment, face recognition and animacy perception were measured for each of the subjects at 33 ms, 84 ms, and 2.02 second exposure times. Both of these tasks were conducted in order to compare the timing of face recognition and animacy. In the face recognition task, the percentage of the faces correctly recognized by each of the subjects was calculated for 33 ms, 84 ms, and 2.02 s (Table 6). As was expected, it was found that as exposure time increased, the accuracy on the face recognition task increased. It was determined that face recognition accuracy significantly improved from 33 ms to 84 ms exposure time (t-test, p=0.0002). The subjects also showed a significant improvement in the task from 84 ms to 2.02 s exposure (t-test, p=0.0004). When the percentage correct was averaged across subjects, it was found that at 33 ms, the accuracy was approximately 47.0%, or near chance (Figure 24). At 84 ms, the average was 68.2% and at 2.02 s it was 92.0% (Figure 24). From these results it can be inferred that at 33 ms exposure time face recognition does not occur, while for 84 ms it appears that face recognition ability improves, but is still not at its optimum level of performance. From these results it can also be seen that 2.02 seconds is a sufficient amount of exposure time for the recognition of faces.

Face Recognition	**33 ms**	**84 ms**	**2.02 s**
Subject 1	65	55	92.5
Subject 2	32.5	75	100
Subject 3	50	85	97.5
Subject 4	40	70	97.5
Subject 5	50	65	92.5
Subject 6	52.5	67.5	92.5
Subject 7	47.5	82.5	92.5
Subject 8	40	55	100
Subject 9	35	70	100
Subject 10	52.5	80	95
Subject 11	52.5	45	52.5

Table 6- The percentage correct at 33 ms, 84 ms, and 2.02 s exposure duration for each of the subjects.

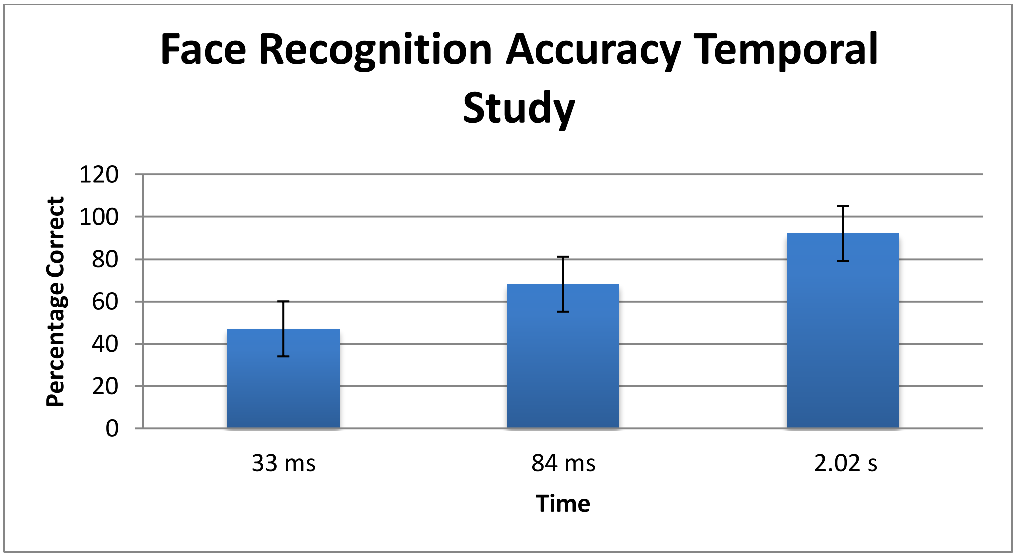


Figure 24- The average percent correct across subjects on the face recognition task for exposure durations of 33 ms, 84 ms, and 2.02 s with standard error bars.

The next task that subjects were tested on was their animacy perception at the same exposure times of 33 ms, 84 ms, and 2.02 s. The animacy ratings of the four morph continua were averaged for each level in the continuum for each of the eleven subjects. The variability in the average ratings for the 33 ms exposure time was observed to be very large between subjects (Figure 25). This result was expected due to the very limited time for analysis of the face. The subjects' responses also appear to be flat or have very small slopes, meaning that the perceived animacy was not markedly greater for the human face than for the doll; in other words, subjects had trouble differentiating between human and doll faces, even those at the extreme ends of the continuum. For the 84 ms exposure time, however, curvature appears in the animacy ratings of the subjects, with decreased variability between subjects' responses (Figure 26). There is further improvement in the curve shape and further increase in the slope of the function with the 2.02 second exposure time (Figure 27). The variability in the responses for the 2.02 second exposure time also appears to be smaller than the 84 ms exposure.

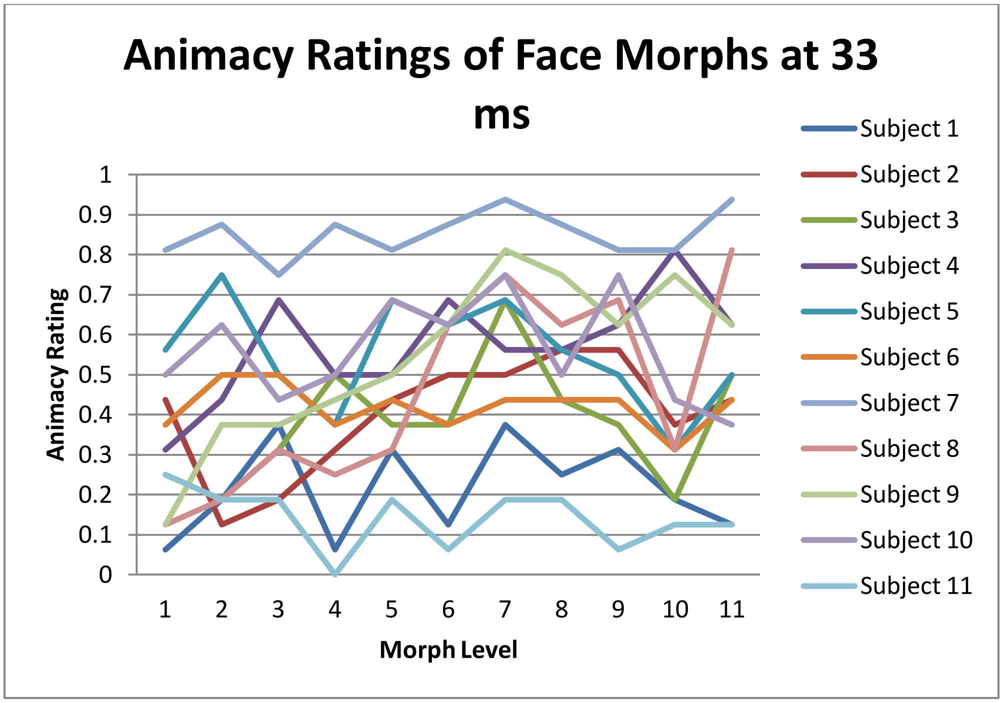


Figure 25- The average ratings of 0 and 1 of morph continua 1-4 at 33 ms exposure duration.

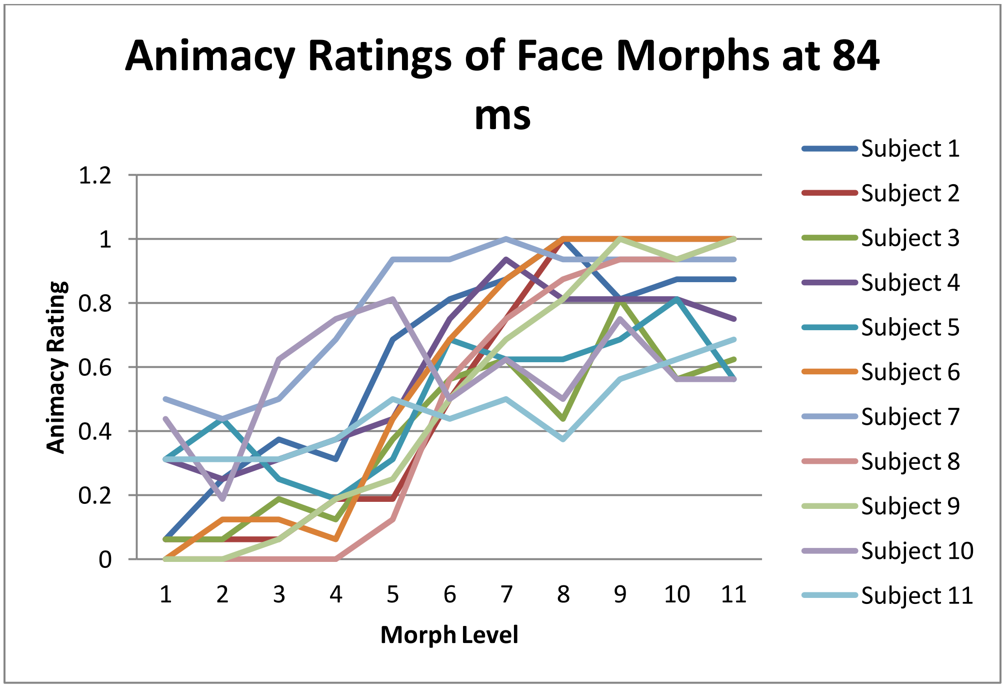


Figure 26- The average ratings of 0 and 1 of morph continua 1-4 at 84 ms exposure duration.

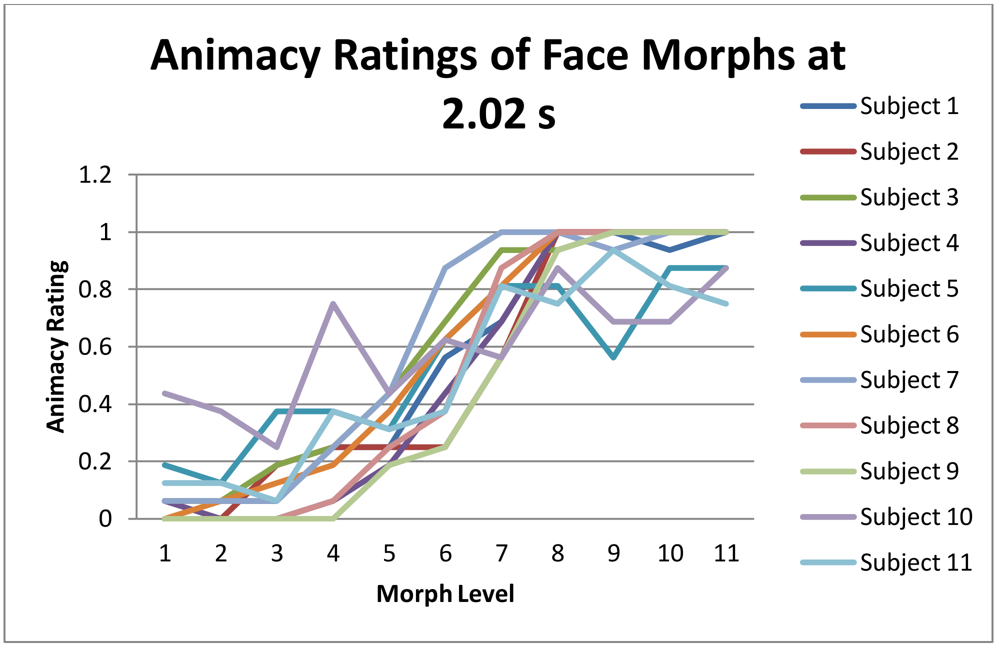


Figure 27- The average ratings of 0 and 1 of morph continua 1-4 at 2.02 s exposure duration.

The ratings of all of the subjects were averaged for each of the three exposure times, and the standard deviations of the ratings for each morph level were calculated in order to quantify the variability in the responses. For the average animacy ratings at 33 ms, it can be observed that the error bars are very large, and there is very little difference in variability along the curve (Figure 28). The standard deviations at the human end and doll end are not smaller than the value in the center, suggesting that rating completely doll and completely human faces was not noticeably easier for the subjects. At 84 ms, the standard deviations are smaller than what was observed at 33 ms (Figure 29). At the human and doll ends of the continuum, however, the variability is not markedly different from the values closer to the center. At 2.02 seconds, the standard deviations along the curve are even smaller than what was found at 84 ms, and the spread at the human and doll ends is generally smaller than the values in the center of the curve as was expected; this is most noticeable on the human end (Figure 30). When all of the standard deviations along the average curves are superimposed, it can be seen that the spread for the 2.02 second exposure time is smaller than what is found for 33 ms and 84 ms (Figure 31). At the human and doll ends of the continua, it can also be seen that the variability in ratings is greatest at 33 ms, then 84 ms, and then 2.02 seconds (Figure 31).

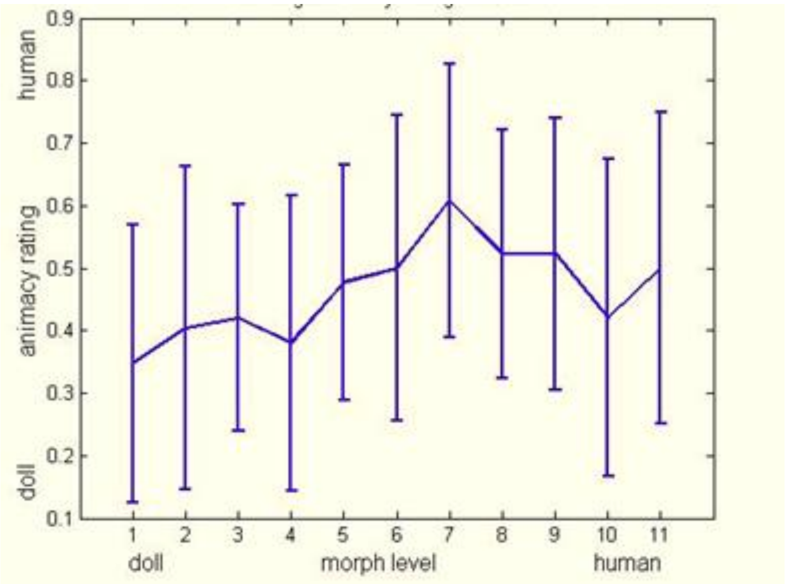


Figure 28- The average animacy ratings across subjects at 33 ms exposure duration with standard deviations.

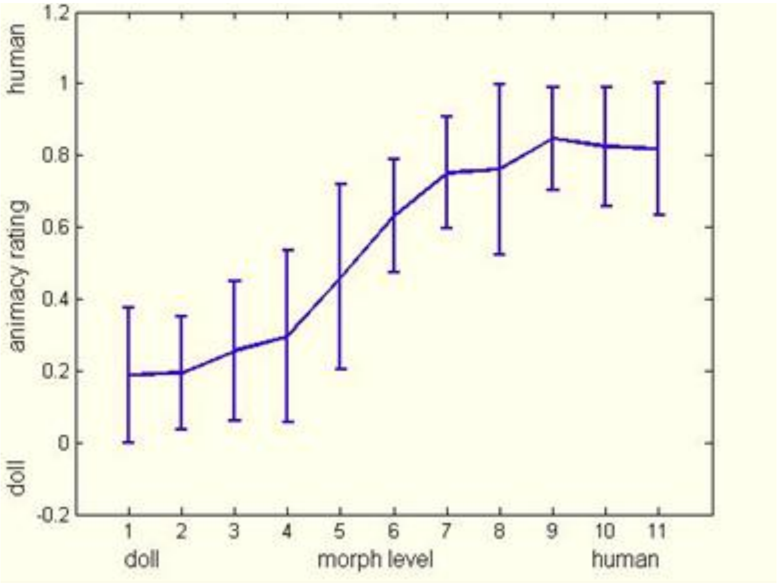


Figure 29- The average animacy ratings across subjects at 84 ms exposure duration with standard deviations.

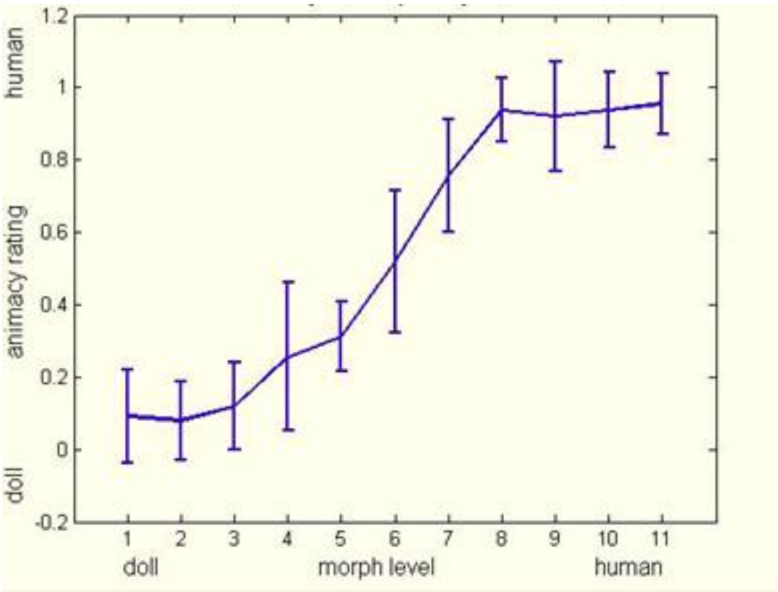


Figure 30- The average animacy ratings across subjects at 2.02 s exposure duration with standard deviations.

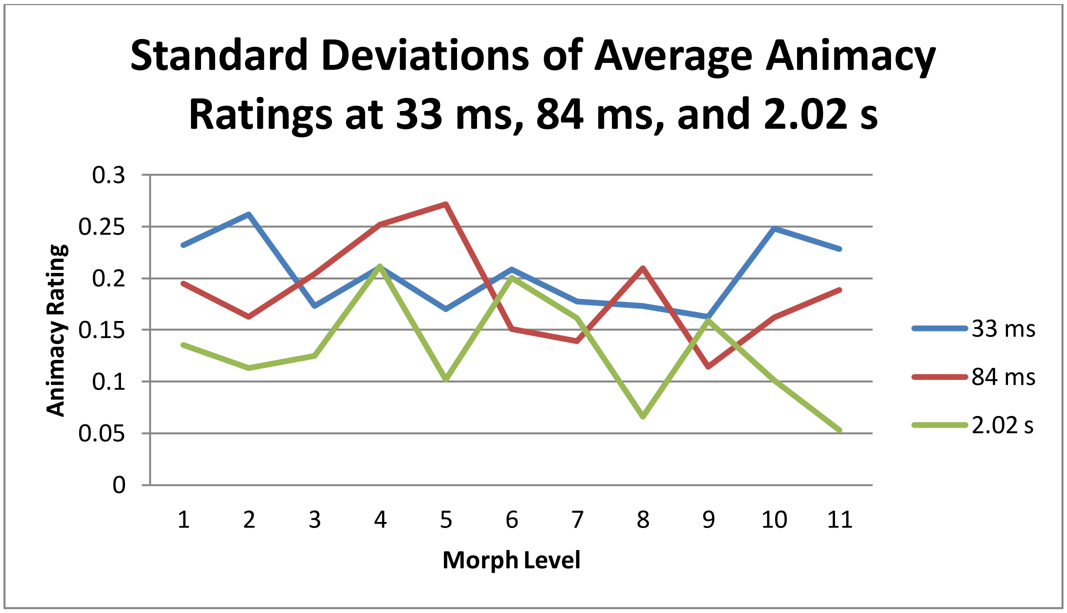


Figure 31- The standard deviations of the average animacy ratings across levels 1-11 at 33 ms (blue), 84 ms (red), and 2.02 s (green) exposure duration.

When the animacy ratings were averaged across subjects for 33 ms, 84 ms, and 2.02 s, it was observed that with an increase in exposure time, the slope or curvature of the plots increased (Figure 32). It can also be seen that with an increase in time, the accuracy of the animacy ratings for completely human and completely doll faces increased. The average humanness rating for a 100% human face was found to be 97.5% at 2.02 seconds, 83.1% at 67 milliseconds, and 53.8%

at 20 milliseconds. The average humanness ratings for a 100% doll face were found to be 8.8% at 2.02 seconds, 17.5% at 84 ms, and 35.6% at 33 ms. In order to approximate the slopes of the animacy curves, psychometric curves were fit to the average ratings for each subject for 84 ms and 2.02 second exposure times; a psychometric curve was not used to fit the data values collected at 33 ms, because this data could not be fit with a psychometric function (Figure 33). A linear function was instead used to fit the values at 33 ms. It can be seen from the values at 33 ms that the slopes are close to 0, and most of the subjects showed slight positive slopes (Table 7). This suggests that there may have been some perception of an increase in humanness as one moved from doll to human in the continuum at 33 ms exposure time. The slopes of the 33 ms curves, however, were not significantly different from 0 (t-test, $p>0.05$). The humanness ratings on the doll side and human side were averaged (five values to the left, and five values to the right) in order to see if there is an increase in rating values as one moves from the doll end to human end on the continuum (Table 8). The difference between the human end average values and doll end values were calculated (Table 8). It was determined that all of the values, except for two, are positive. These differences were found to be statistically significant (t-test, $p<0.05$), suggesting that subjects were picking up some, albeit little, information about face animacy even in 33 ms. These results tell us that very limited animacy perception can occur at 33 ms, which is contrary to what was found in face recognition, where performance was near chance.

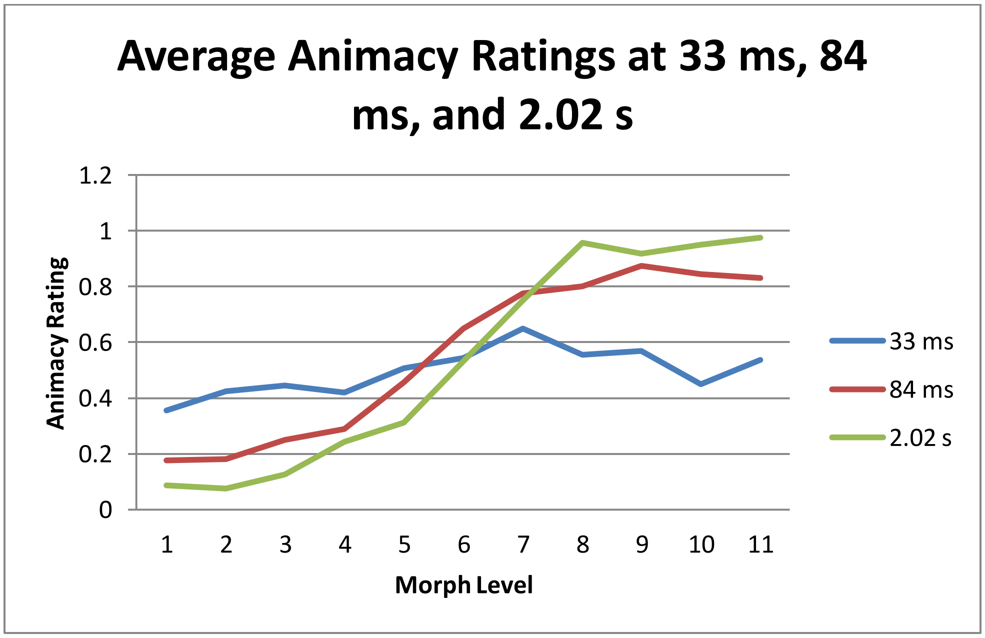


Figure 32- The animacy ratings averaged across subjects for 33 ms (blue), 84 ms (red), and 2.02 s (green) exposure duration.

Slopes	33 ms	84 ms	2.02 s
Subject 1	-0.008	1.7236	8.3331
Subject 2	0.0051	10.1542	11.7764
Subject 3	0.0244	16.8865	23.7469
Subject 4	0.0148	3.8553	11.4885
Subject 5	0.0278	5.4413	16.1863
Subject 6	-0.0153	2.7123	4.8283
Subject 7	-0.0045	15.4391	11.5264
Subject 8	0.0063	7.1553	19.3716
Subject 9	0.0568	17.4933	16.8129
Subject 10	0.0517	9.2654	15.5876
Subject 11	-0.0034	0.8817	2.2544

Table 7- The slopes of the trend line and curves of the average animacy ratings for each of the subjects at 33 ms, 84 ms, and 2.02 s exposure duration.

Average	Left	Right	Diff
Subject 1	0.1625	0.1375	-0.025
Subject 2	0.2	0.25	0.05
Subject 3	0.3	0.4875	0.1875
Subject 4	0.325	0.4375	0.1125
Subject 5	0.4875	0.6375	0.15
Subject 6	0.575	0.5125	-0.0625
Subject 8	0.825	0.875	0.05
Subject 9	0.2375	0.6375	0.4
Subject 10	0.3625	0.7125	0.35
Subject 11	0.55	0.5625	0.0125

Table 8- The average animacy ratings from 0%-40% human (left) and from 60%-100% human (right) for each of the subjects, and the difference between the left and right values.

The PSEs of the psychometric functions for each of the subjects' ratings were also calculated for 84 ms and 2.02 seconds (Table 9). The PSE could not be determined for the responses at 33 ms, due to the high level of uncertainty in the ratings. The PSE values at 84 ms were not found to be significantly different from what was found at 2.02 second exposure time (t-test, p=0.45).

PSE	84 ms	2.02 s
Subject 1	0.6391	0.4902
Subject 2	0.314	0.4804
Subject 3	0.5038	0.5698
Subject 4	0.6244	0.4112
Subject 5	0.3412	0.5164
Subject 6	0.4993	0.4263
Subject 7	0.4411	0.4376
Subject 8	0.1089	0.3961
Subject 9	0.5098	0.4938
Subject 10	0.5063	0.5563
Subject 11	0.0653	0.2977

Table 9- The PSEs of the psychometric curves fit to the average ratings of each of the subjects at 84 ms, and 2.02 s exposure duration.

The animacy ratings of all of the subjects were averaged for 33 ms, 84 ms, and 2.02 s exposure times, and psychometric curves were fit, in order to calculate the PSE and slope of the averages across subjects. The trendline for the average ratings at 33 ms appears to have a positive slope (Figure 33). The slope of the line was determined to be 0.0142 with a PSE of 0.85. For 84 ms, there is a noticeably larger slope and curvature in the psychometric function (Figure 34). The curve for the average ratings at 84 ms exposure had a slope of 6.1 and a PSE of 0.42. The psychometric function for the 2.02 second exposure time appears to have an increase in slope relative to the curves obtained for 84 ms with a slope of 10.6 and a PSE of 0.47 (Figure 35).

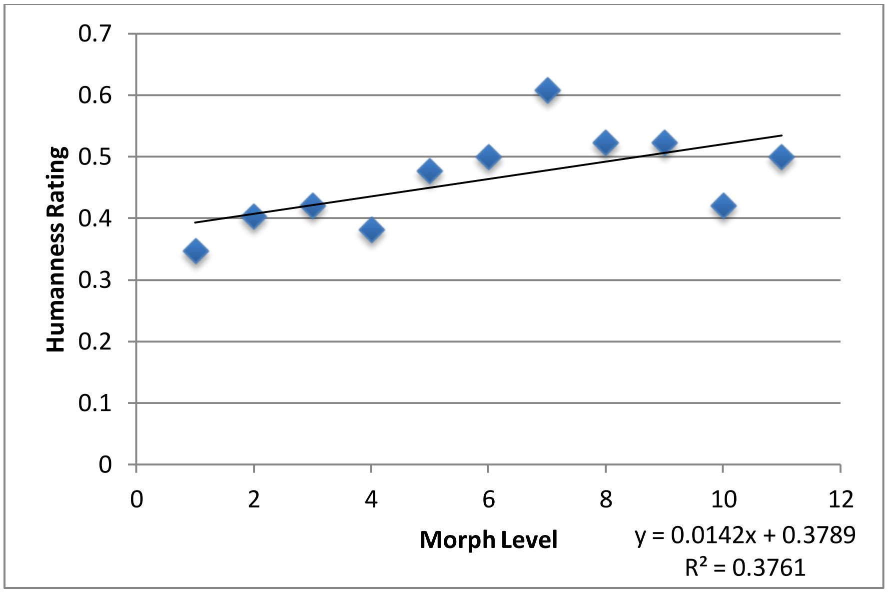


Figure 33- The trendline for the average ratings across all subjects at 33 ms exposure time. The trendline has a slope of 0.0142 and a PSE of 0.85.

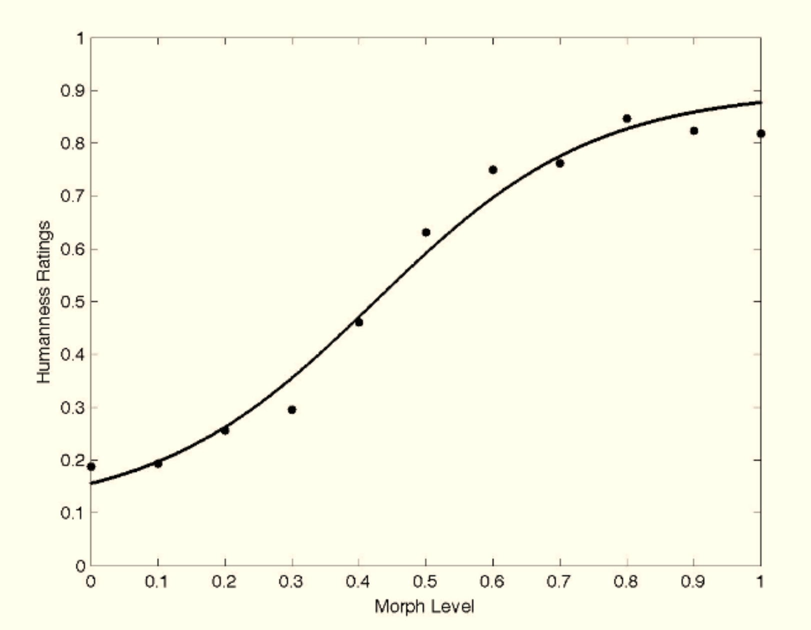


Figure 34- A psychometric curve fit to the average animacy ratings across all subjects at 84 ms exposure duration. The curve has a slope of 6.1 and a PSE value of 0.42.

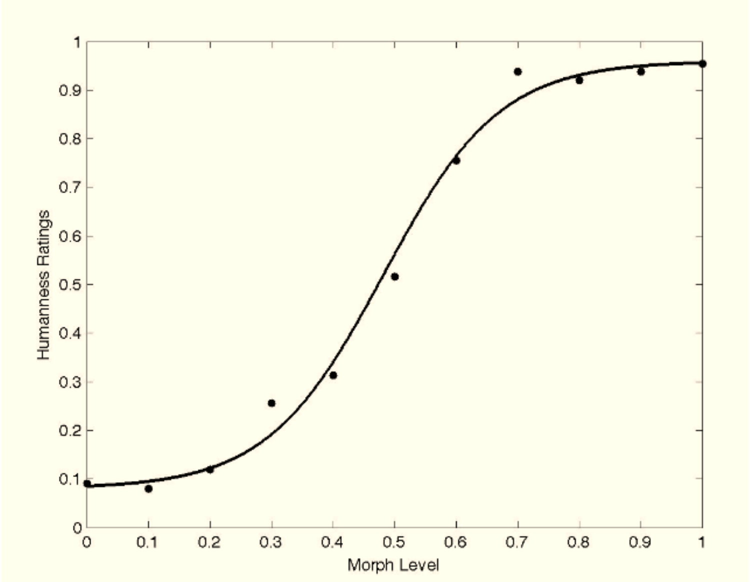


Figure 35- A psychometric curve fit to the average animacy ratings across all subjects at 2.02 s exposure duration. The curve has a slope of 10.6 and a PSE value of 0.47.

4. Experiment III: Spatial Frequency

The spatial frequency ranges necessary for determining the identity, gender, and expression of a face have been explored in previous studies (Halit et al., 2006). The purpose of this experiment was to further investigate the importance of configural cues in animacy perception by manipulating the spatial frequency content of the face images. It has been previously shown that more of the holistic cues in the face are contained in the lower spatial frequency information, while the high spatial frequencies aid more in featural analysis. In this experiment we analyzed which range of spatial frequencies is more critical for animacy perception. The results of this experiment may increase our understanding of whether configural processing or featural processing is more important for animacy perception. From the results of the Looser and Wheatley experiments, we predicted that the low spatial frequencies would be more important in animacy perception because holistic cues are mostly encoded in the low spatial frequencies of the face. This study is still in the piloting phase with a current total of four subjects. The morph images that were presented either contained all of the spatial frequencies, only high spatial frequencies, or only low spatial frequencies. The ratings from the high spatial frequency images and low spatial frequency images were compared to the ratings for images containing the entire range of spatial frequencies, in order to observe the effects of reducing certain types of facial information.

4.1. Method

4.1.1. Subjects

Four subjects (male and female between the ages of 19-60 years) were selected from the subject pool at the Brain and Cognitive Sciences Department at MIT. The participants that were selected had normal or corrected-to-normal vision. The subjects were paid for their participation in the research after completion of the tasks.

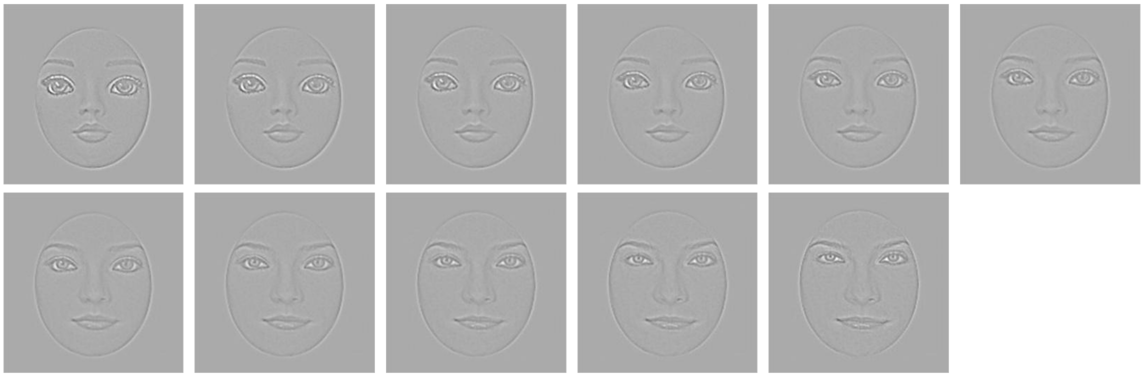
4.1.2. Stimuli

The faces in morph continua 2 and 4 from the previous two experiments were used in this experiment. The size of the human and doll faces were enlarged and were morphed using the Fantamorph software. The external shape of the face stimuli was normalized by using an oval face template. The morph continua of doll and human faces ranged from 0% human to 100% human in increments of 10%, resulting in a total of eleven images for each morph continuum (Figure 36). The morph images were 256 by 256 pixels with standard gray backgrounds. The low frequency images were created by cutting off the frequencies above 8 cycles per face. The high frequency images were formed by removing the frequencies below 24 cycles per face.

A.

B.

C.



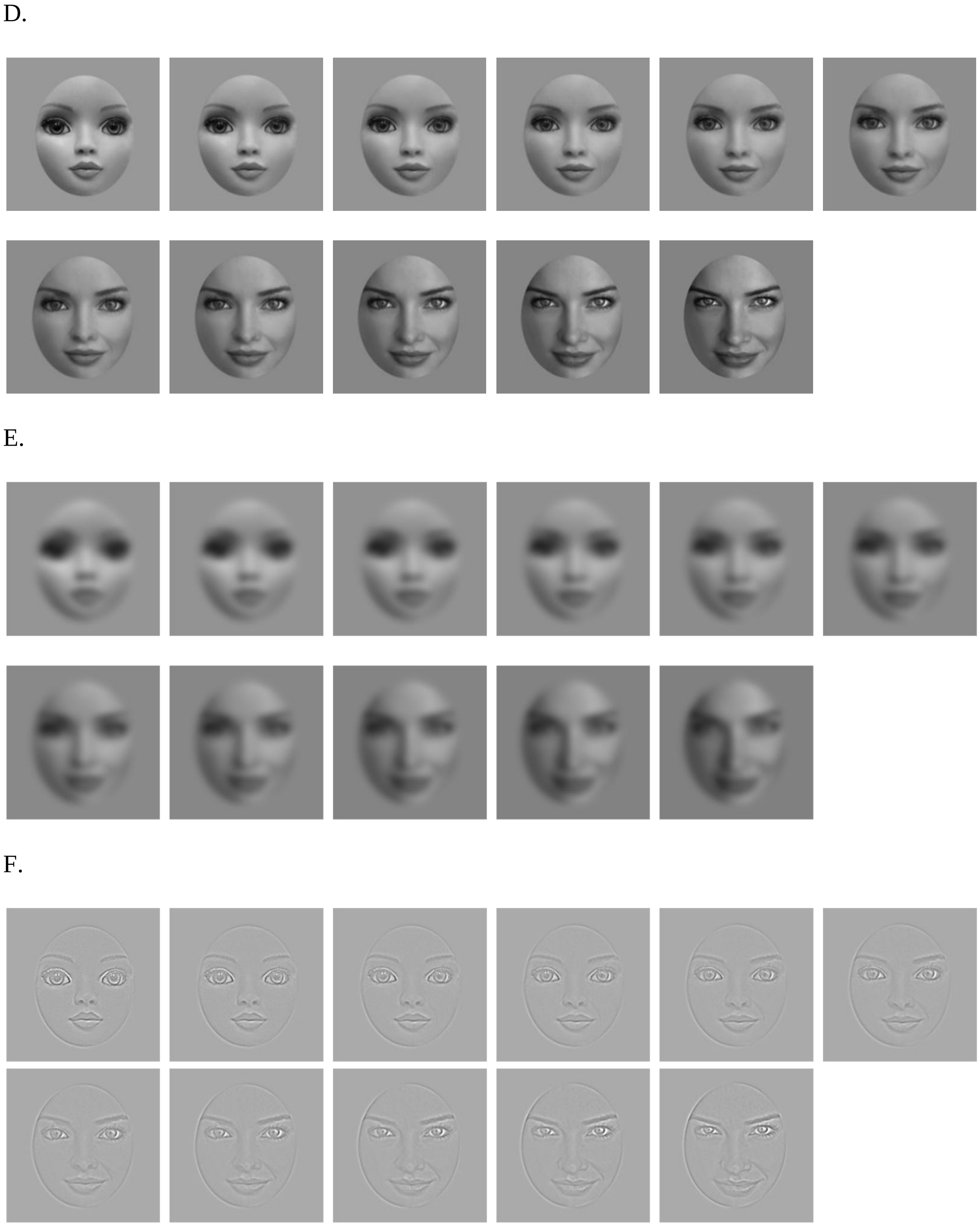
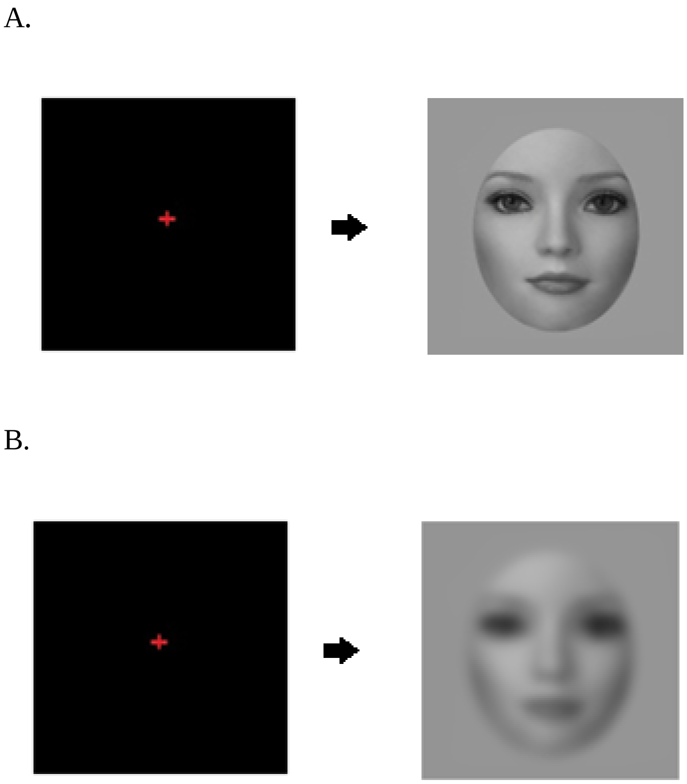


Figure 36- (A) Morph continuum 2 with eleven faces 0% human to 100% human in increments of 10% with face shape standardized with a template. All spatial frequencies are present in the images. (B) Morph continuum 2 with only the low spatial frequencies present in the images. (C) Morph continuum 2 with only the high spatial frequencies present in the images. (D) Morph continuum 4 with all spatial frequencies present in the images. (E) Morph continuum 4 with only the low spatial frequencies present in the images. (F) Morph continuum 4 with only the high spatial frequencies present in the images.

4.1.3. Procedure

The three image groups that were tested in this experiment were the low spatial frequency, high spatial frequency, and all spatial frequency images (Figure 37). The exposure times for all the images were kept constant at 2.02 seconds, which is an exposure duration that gives the subject a sufficient amount of time to process the face. The subjects were given the task to rate the face images on the screen as either ‘doll’ or ‘real’ using the keyboard (0=doll, 1=real). The verbal instructions that were presented to the subjects were the following: “We morphed faces between real human and doll faces to create a continuum from doll to human. You will see these morph images on the screen and I would like you to press 1 if you think the face looks more human-like or 0 if you think the face looks more doll-like. You will see these faces sometimes blurred, and sometimes just the outlines of the features of the face, and other times you will see the normal image.” Each of the face images in the four morph continua was repeated seven times making a total of 462 trials. The order of the face images in the spectrum and the spatial frequency ranges present were randomized during the experiment.



C.

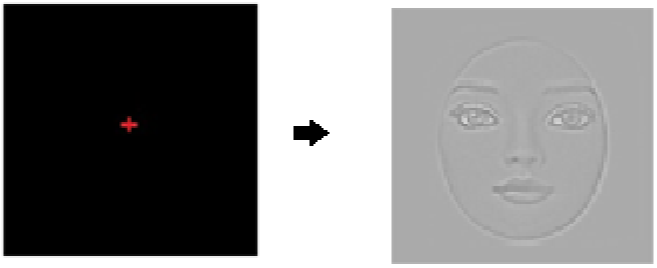


Figure 37- (A) An example of a trial where a subject is shown a face morph with all spatial frequencies and asked to rate the face as either human or doll. (B) An example of a trial where a face morph is presented with only low spatial frequencies. (C) An example of a trial where a face morph is presented with high spatial frequencies.

4.2. Results

A pilot experiment with four subjects was run for the spatial frequency experiment. The subjects rated faces from morph continua 2 and 4 with all spatial frequencies present, only the low spatial frequencies, and only high spatial frequencies. The animacy ratings of the two morph continua for the different spatial frequency groups were averaged (Figures 38-40). Psychometric curves were fit to the average ratings for the three spatial frequency ranges. No significant difference between the PSEs obtained for the low spatial frequency images and the images containing all of the spatial frequencies was found (t-test, p=0.293). There was also no significant difference between the PSE values obtained from the data for the high spatial frequency images and images with the entire range of spatial frequencies (t-test, p=0.321). When the slopes of the subjects' animacy rating curves were compared for the low spatial frequency images and the normal images, a significant difference between the slopes was not observed (p=0.353). Also, no significant difference in the slopes, between the high spatial frequency images and the normal images was found (p=0.0583), however this p value could potentially become significant if more subjects were tested. The curves with the ratings for the high spatial frequency images have slightly smaller slopes than for the normal images, which may also be due to the low contrast images that were used; it may have been more difficult for subjects to see the features. The animacy ratings were averaged across subjects for each of the three groups of images, and psychometric curves were fit to the ratings (Figures 41-43). The PSE for the average ratings was 0.514 for the images with all spatial frequencies, 0.478 for the images with only the low spatial frequencies, and 0.574 for the images with the high spatial frequencies. The slope of the curves was 16.56 for the images with all spatial frequencies, 17.44 for the low spatial frequency images, and 9.27 for the high spatial frequency images.

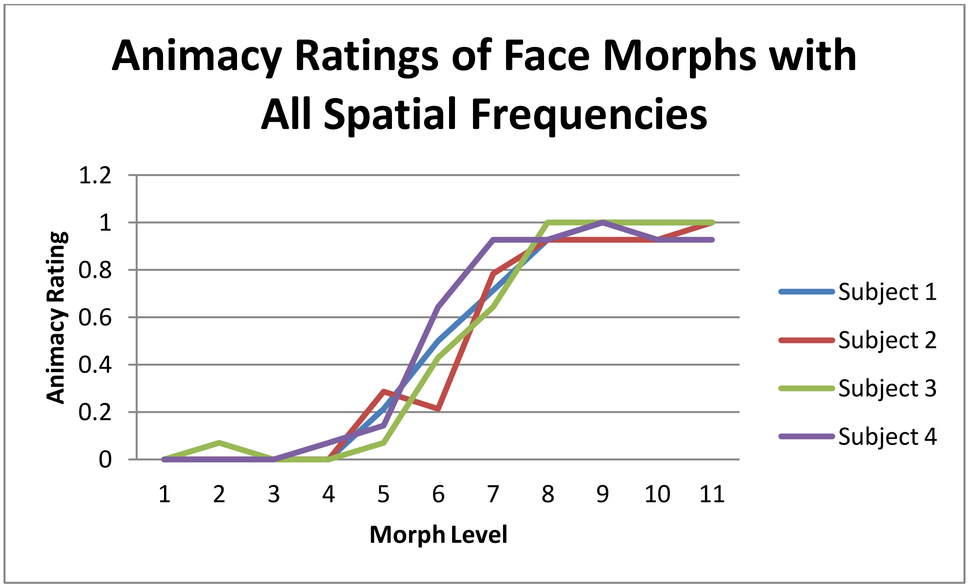


Figure 38- The average animacy ratings of 0 and 1 for morph continua 2 and 4 with all spatial frequencies present.

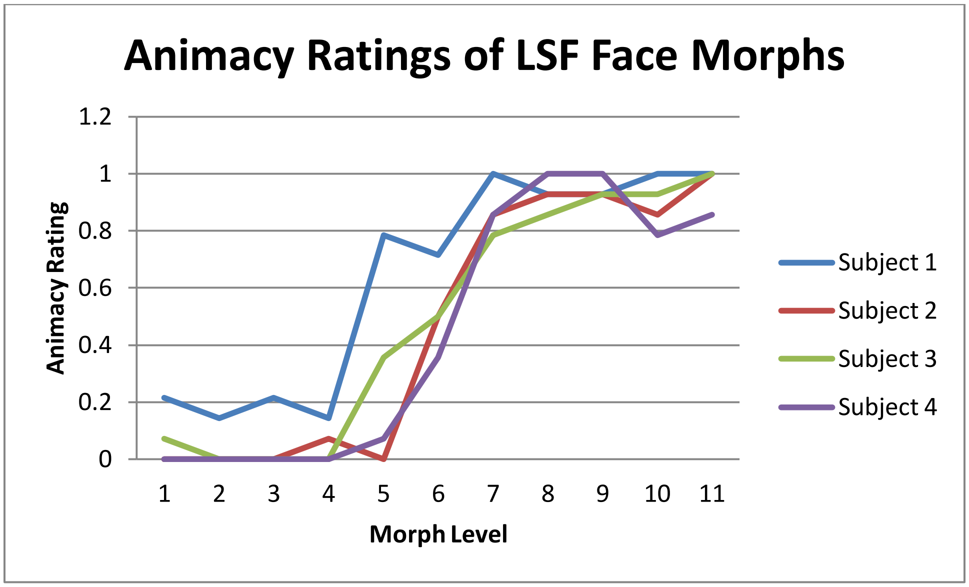


Figure 39- The average animacy ratings of 0 and 1 for morph continua 2 and 4 with only the low spatial frequencies present.

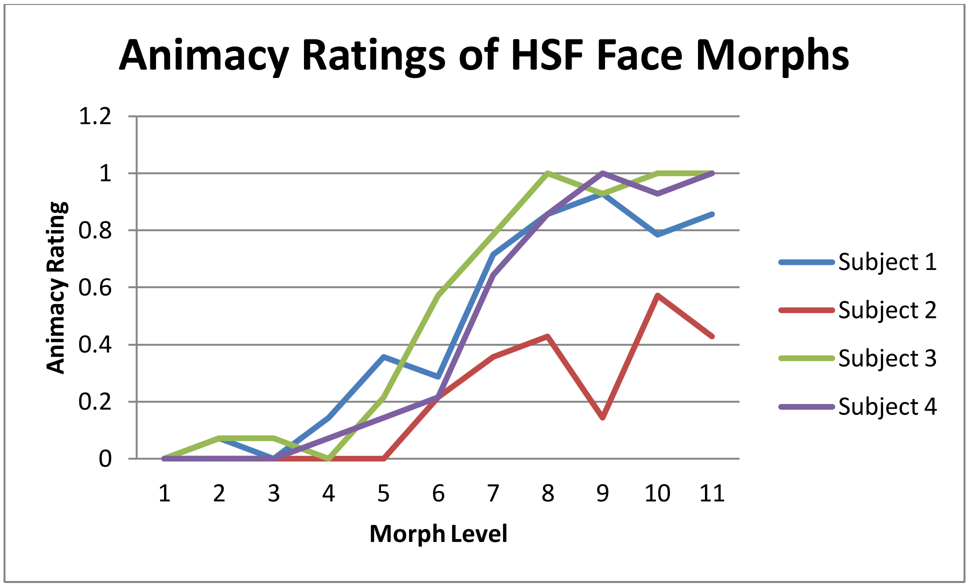


Figure 40- The average animacy ratings of 0 and 1 for morph continua 2 and 4 with only high frequencies present.

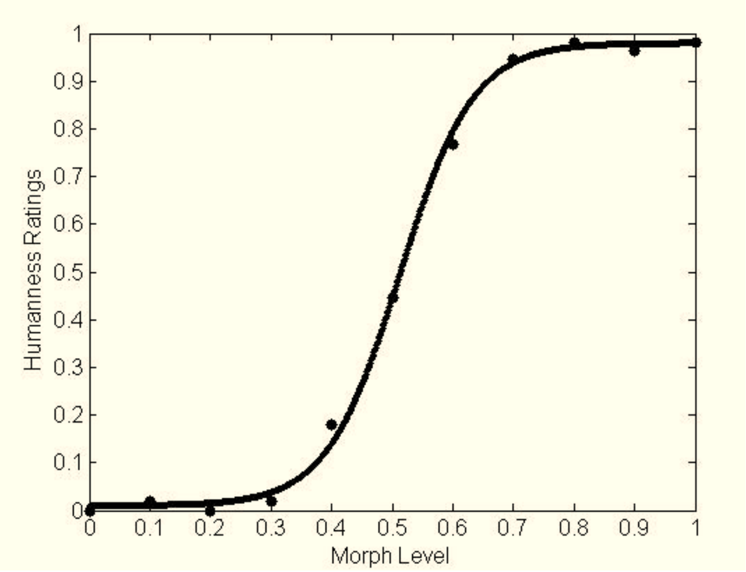


Figure 41- A psychometric curve fit to the average ratings across subjects for morph continua 2 and 4 with all spatial frequencies present.

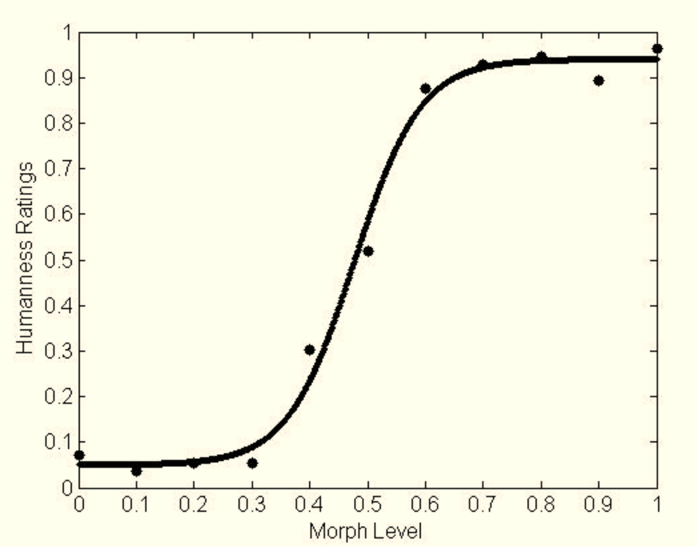


Figure 42- A psychometric curve fit to the average ratings across subjects for morph continua 2 and 4 with only the low spatial frequencies present.

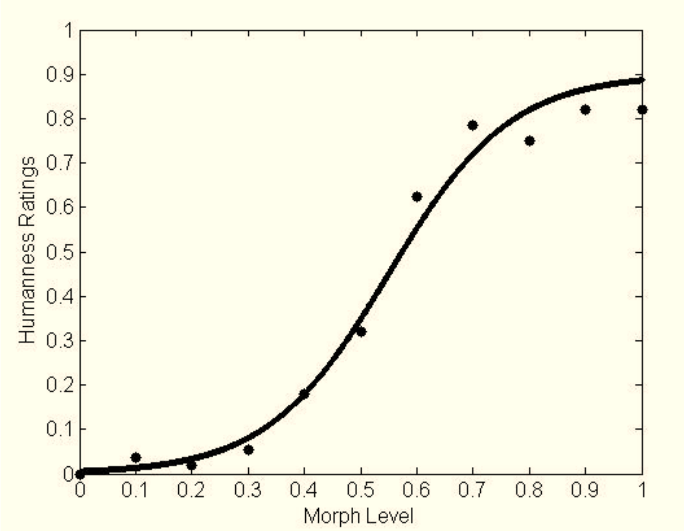


Figure 43- A psychometric curve fit to the average ratings across subjects for morph continua 2 and 4 with only the high spatial frequencies present.

5. Discussion

Three experiments were conducted for this thesis in order to gain insight both on the time-course of animacy perception and the extent to which animacy perception relies on configural information. This included the inversion experiment, temporal experiment, and spatial frequency experiment. Previous studies have shown that there is an 'inversion effect' on face perception most likely due to reduced configural information when analyzing inverted faces (Tanaka and Gauthier, 1997). This configural information in a face is generally believed to be processed more quickly than the finer lines and details of the individual features of the face (Richler et al., 2009). In the inversion experiment we decreased the holistic processing of the morph images by inverting the faces and compared the ratings to upright faces, in order to see if this manipulation has an effect on animacy perception. Our prediction for this experiment was that animacy perception would change due to inversion, which was based on the findings of Looser and Wheatley (2010) that configural information is necessary in order to form an accurate perception of animacy.

Previous experiments have tested the time-course of human face recognition ability, and in the second experiment we limited the exposure durations of the face images, to better understand the time-course of animacy perception in faces. A face recognition experiment was also conducted in order to compare the time-course across tasks. The hypothesis for the temporal experiment was that 33 ms exposure time would not be enough time to either process animacy nor recognize a face, while with 84 ms it would be possible to form a perception of animacy while not being quite enough time to recognize a face.

For the spatial frequency pilot experiment we examined which range of spatial frequencies is more important for animacy perception in faces. The spatial frequency ranges may be linked to configural and featural information in a face (Goffaux et al., 2005). Low spatial frequencies may be more important for configural processing, while high spatial frequencies may be more important for featural processing (Goffaux et al., 2005). It is important to note, however, that high spatial frequency images can still be used to derive configural information about the face. Based on the results from Looser and Wheatley (2010) we predicted that the lower spatial frequencies of the face would be more important because in their study they had found that holistic information is an important part of forming this judgment.

From the inversion experiment, variability across subjects was a factor that was used to determine the availability of important cues used to form animacy perception. It was observed that variability across subjects' responses increased when the faces in three of the morph continua were inverted. The averages of the standard deviations for levels 4-8 of three of the morph continua were larger for the inverted faces than for the faces in the upright orientation. However, even with this variability across subjects, people were able to perform the task in both the upright and inverted cases, which can be seen from the large difference in the ratings on the human and doll ends of the spectrum. The increase in variability for the inverted faces suggests that it may have been more difficult for the subjects to assign a rating to the faces in this orientation. The lack of increase in variability for the fourth morph continuum in the inverted orientation may be due to a single featural characteristic or multiple featural characteristics that were not affected by inversion. An example of such a factor could be that the difference in eye-size between the 100% doll and 100% human face was greater for this morph continuum than for the others. Inversion of the face may not have an impact on the size perception of this individual featural cue. In the animacy ratings for all of the morph continua, it can be seen that the variability is generally larger for images closer to the center of the continuum than at the human and doll ends. Another observation made from the data is that on the human end of all of the rating curves, the 100% human upright faces were rated as being significantly more human-like than the 100% human inverted faces. This pattern was not observed for the upright or inverted faces located at the doll end of the morph spectrums. This suggests that there was an 'inversion effect' on the perception of animacy of real human faces but not for objects that resemble human faces very closely. This result also implies that holistic processing may be needed to make a more accurate judgment on whether a face is human or doll. The PSEs of the animacy ratings for the upright and inverted faces were not found to be significantly different, however, the slopes of the animacy ratings for the inverted faces were significantly smaller than that obtained for the upright faces.

In the last part of the inversion experiment, where the morph continua were kept the same in the upright and inverted trials, it was determined that the PSE of the average ratings across subjects is 49.22% for upright faces, and only slightly smaller in the inverted orientation with a PSE of 40.24%. No significant difference between the upright and inverted PSE values between subjects was found. The slope of the average ratings in the upright orientation was determined to be 19.08 and 14.21 in the inverted orientation, however a significant difference was not observed in the slopes. A comparison between the PSE values may not be the best determiner of accuracy in the animacy ratings, because the differences in variability of the ratings are not accounted for. A subject pool of four participants is also not large enough for accurate statistical analysis. From these inversion experiments it can be inferred that holistic processing improves the certainty of the subjects' animacy ratings, and significantly improves the perception of animacy of real human faces.

For the temporal experiment in the face recognition task, the percentages of face images recognized correctly at 33 ms, 84 ms, and 2.02 s were calculated to compare performance at these three durations. It was determined that with an increase in exposure duration, there was an increase in performance on the face recognition task. Face recognition performance was found to significantly improve when exposure was increased from 33 ms to 84 ms, and also when exposure time increased from 84 ms to 2.02 s. The average accuracy at 33 ms was found to be near chance at 47.0% correct, while accuracy was 68.2% at 84 ms, and 92.0% at 2.02 s. These results suggest that at 33 ms, face recognition does not occur, at 84 ms there is limited face recognition ability, and at 2.02 s face recognition can occur at an optimal level. The variability in the ratings was found to be large between subjects at 33 ms. The slopes of the animacy curves at this exposure time also were observed to be relatively flat, suggesting that the human faces were not perceived to be more human-like than the doll faces. With 84 ms exposure time, the variability in the ratings between subjects decreased, and the slopes of the rating curves increased. Further improvement in decreased variability and increase in slope was found at 2.02 s exposure time. The spread between the subjects' ratings was generally found to be the least at 2.02 second exposure, and greatest at 33 ms. The accuracy in the ratings for 100% human and 100% doll faces also increased with an increase in exposure time. The average perceived humanness rating for a 100% human face was 97.5% at 2.02 s, 83.1% at 84 ms, and 53.8% at 33 ms. The average humanness ratings for a 100% doll face were found to be 8.8% at 2.02 s, 17.5% at 84 ms, and 35.6% at 33 ms. The difference between the average ratings at the human and doll ends was determined to be statistically significant, suggesting that there is limited animacy perception at 33 ms. This is contrary to what was found for face recognition, where performance was at chance. The PSE values at 84 ms exposure time were not significantly different from the PSE values at 2.02 s. From these results, it can be inferred that animacy perception can occur with less time than is needed for face recognition, suggesting that the detailed aspects of the face may be more important for face recognition than for animacy judgment.

In the spatial frequency pilot study, there was no significant difference between the PSEs for the animacy ratings obtained for the low spatial frequency and the images with all of the spatial frequencies. A significant difference between the PSEs for the ratings of the high spatial frequency images and the all spatial frequency images also was not found. The results from this experiment suggest that subjects were able to form similar animacy perceptions of the morph images even when the high spatial frequencies or low spatial frequencies were eliminated. A strong conclusion, however, cannot be made with these results, because four subjects is not a large enough subject pool. In order to improve this experiment, a number of factors should be changed in future studies. The high spatial frequency images were very low contrast images, which may have made it difficult for subjects to extract information from the faces. In a future study, the contrasts of these images should be increased. Data should also be collected from more subjects in order to decrease the probability that the null hypothesis will be incorrectly retained.

The results from the first experiment in this thesis suggest that a holistic processing of a face is important for perceiving whether a face is real. How this configural information is encoded in the brain, however, has yet to be determined. From the second experiment it can be inferred that face recognition requires more time and detailed analysis of the face than animacy perception. The stages in forming a perception of animacy may require more holistic or simpler cues that require less time to extract than more detailed characteristics of the face. The results from the spatial frequency pilot experiment show us that removing the high spatial frequencies, or low spatial frequencies in the image may have no significant effect on animacy perception; a strong conclusion, however, cannot be made from these results because the experimental design needs to be improved and a larger subject pool is necessary.

The subjects in the first inversion experiment were able to successfully form an animacy judgment even with the faces in the inverted orientation. Their animacy judgments, however, were significantly less human-like for the completely human inverted faces. Taking in to mind the theory that featural and configural processing of faces are parallel pathways, it is possible that using the pathway for featural processing allows for limited animacy perception, and that the configural pathway aids in reinforcing this animacy judgment. This reasoning may provide an answer as to why the subjects were able to perform the task, but were less certain in their judgments with inversion. From the results of the temporal experiment there is evidence that more processing pathways may be necessary for face recognition than for animacy perception. If the spatial frequency pilot study is expanded, it would be interesting to see whether the high spatial frequency or low spatial frequency pathway has greater weight in forming an animacy judgment.

The Theory of Mind is an intrinsic ability that humans have to understand the beliefs, desires, and intentions of other human beings (Saxe, 2006). This ability is lacking, however, in people diagnosed with autism spectrum disorders (Simon Baron-Cohen et al., 1985). Animacy judgment of an object that closely resembles a human face may precede and lead to the theory of mind processing in the human brain. Forming a judgment of whether an object is animate or inanimate may be a critical prerequisite for the theory of mind, because humans are not interested in understanding the thoughts and desires of dolls or simple smiley face drawings. It is possible that people with autism spectrum disorders have an impairment in their animacy perception, which may as a result lead to deficits in theory of mind. Future experiments can explore this question by conducting animacy perception studies on people with autism spectrum disorders to gain more insight on the processing pathways that are impaired in people with this disorder.

6. Bibliography

1. Allison, Truett, Aina Puce, and Gregory McCarthy. "Social perception from visual cues: role of the STS region." *Trends in cognitive sciences* 4.7 (2000): 267-278.
2. Baron-Cohen, Simon, Alan M. Leslie, and Uta Frith. "Does the autistic child have a “theory of mind”?" *Cognition* 21.1 (1985): 37-46.
3. Bartlett, James C., and Jean Searcy. "Inversion and configuration of faces."*Cognitive psychology* 25 (1993): 281.
4. Brainard, D. H. The Psychophysics Toolbox, *Spatial Vision 10* (1997):433-436.
5. Bruce, Vicki, Tony Doyle, Neal Dench, and Mike Burton. "Remembering facial configurations." *Cognition* 38.2 (1991): 109-144.
6. Diamond, Rhea, and Susan Carey. "Why faces are and are not special: an effect of expertise." *Journal of Experimental Psychology: General; Journal of Experimental Psychology: General* 115.2 (1986): 107.
7. Ellis, Hadyn D., Jan B. Deregowski, and John W. Shepherd. "Description of white and black faces by white and black subjects." *International Journal of Psychology* 10.2 (1975): 119-123.
8. *FantaMorph*. *Abrosoft FantaMorph*. N.p., 2002. Web. 20 Apr. 2013. <http://www.fantamorph.com/index.html>.
9. Farah, Martha J., Kevin D. Wilson, Maxwell Drain, James N. Tanaka. "What is "special" about face perception?."*Psychological Review-New York-* 105 (1998): 482-498.
10. Farah, Martha J., James W. Tanaka, and H. Maxwell Drain. "What causes the face inversion effect?." *Journal of Experimental Psychology: Human perception and performance* 21.3 (1995): 628.
11. Freiwald, Winrich A., Doris Y. Tsao, and Margaret S. Livingstone. "A face feature space in the macaque temporal lobe." *Nature neuroscience* 12.9 (2009): 1187-1196.
12. Goffaux, Valerie, Barbara Hault, Caroline Michel. "The respective role of low and high spatial frequencies in supporting configural and featural processing of faces." *Perception-London* 34.1 (2005): 77-86.
13. Goldstein, Alvin G. "Learning of inverted and normally oriented faces in children and adults." *Psychonomic Science* (1965).
14. Halit, H., M. De Haan, P. G. Schyns, and M. H. Johnson. "Is high-spatial frequency information used in the early stages of face detection?" *Brain research* 1117 (2006): 154-161.
15. Harmon, Leon D., and Bela Julesz. "Masking in visual recognition: effects of two-dimensional filtered noise." *Science* (1973).

16. Haxby, James V., Elizabeth A. Hoffman, and M. Ida Gobbini. "The distributed human neural system for face perception." *Trends in cognitive sciences* 4.6 (2000): 223-233.
17. Hoeger, Rainer. "Speed of processing and stimulus complexity in low-frequency and high-frequency channels." *Perception-London-* 26 (1997): 1039-1046.
18. Johnson, Mark H. "Subcortical face processing." *Nature Reviews Neuroscience* 6.10 (2005): 766-774.
19. Kanwisher, Nancy, Frank Tong, and Ken Nakayama. "The effect of face inversion on the human fusiform face area." *Cognition* 68.1 (1998): B1-B11.
20. Kanwisher, Nancy, Josh McDermott, and Marvin M. Chun. "The fusiform face area: a module in human extrastriate cortex specialized for face perception."*The Journal of Neuroscience* 17.11 (1997): 4302-4311.
21. Kanwisher, Nancy. "Domain specificity in face perception." *Nature neuroscience* 3 (2000): 759-763.
22. Keysers, Christian, D-K. Xiao, P. Földiák, and D. I. Perrett. "The speed of sight." *Journal of cognitive neuroscience* 13.1 (2001): 90-101.
23. Keysers, Christian, and David I. Perrett. "Visual masking and RSVP reveal neural competition." *Trends in cognitive sciences* 6.3 (2002): 120-125.
24. Kleiner M, Brainard D, Pelli D, 2007, "What's new in Psychtoolbox-3?" Perception 36 ECVP Abstract Supplement.
25. Le Grand, Richard, Catherine J. Mondloch, Daphne Maurer, and Henry P. Brent. "Neuroperception: Early visual experience and face processing." *Nature* 410.6831 (2001a): 890-890.
26. Le Grand, Richard, Catherine J. Mondloch, Daphne Maurer, and Henry P. Brent. "Correction: Early visual experience and face processing." *Nature* (2001b).
27. Le Grand, Richard, Catherine J. Mondloch, Daphne Maurer, and Henry P. Brent. "Impairment in holistic face processing following early visual deprivation." *Psychological Science* 15.11 (2004): 762-768.
28. Bruce, Helmut Leder Vicki. "Local and relational aspects of face distinctiveness." *The Quarterly Journal of Experimental Psychology: Section A*51.3 (1998): 449-473.
29. Lee, Kieran, Graham Byatt, and Gillian Rhodes. "Caricature effects, distinctiveness, and identification: Testing the face-space framework."*Psychological Science* 11.5 (2000): 379-385.
30. Liu, Chang Hong, Charles A. Collin, Stephane JM Rainville, and Avi Chaudhuri. "The effects of spatial frequency overlap on face recognition." *Journal of Experimental Psychology: Human Perception and Performance* 26.3 (2000): 956.
31. Livingstone, Margaret, and David Hubel. "Segregation of form, color, movement, and depth: anatomy, physiology, and perception." *Science* 240.4853 (1988): 740-749.
32. Looser, Christine E., and Thalia Wheatley. "The Tipping Point of Animacy How, When, and Where We Perceive Life in a Face." *Psychological science* 21.12 (2010): 1854-1862.
33. Maurer, Daphne, Richard Le Grand, and Catherine J. Mondloch. "The many faces of configural processing." *Trends in cognitive sciences* 6.6 (2002): 255-260.

34. McKone, Elinor. "Isolating the special component of face recognition: peripheral identification and a Mooney face." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30.1 (2004): 181.
35. Merigan, William H., and John HR Maunsell. "How parallel are the primate visual pathways?." *Annual review of neuroscience* 16.1 (1993): 369-402.
36. Moscovitch, Morris, Gordon Winocur, and Marlene Behrmann. "What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition." *Journal of Cognitive Neuroscience* 9.5 (1997): 555-604.
37. Özbek, Müge, and Markus Bindemann. "Exploring the time course of face matching: Temporal constraints impair unfamiliar face identification under temporally unconstrained viewing." *Vision research* 51.19 (2011): 2145-2155.
38. Pelli, D. G. The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision 10 (1997)*:437-442.
39. Prins, N & Kingdom, F. A. A. (2009) Palamedes: Matlab routines for analyzing psychophysical data. http://www.palamedestoolbox.org
40. Rhodes, Gillian. "Looking at faces: first-order and second-order features as determinants of facial appearance." *Perception* (1988).
41. Richler, Jennifer J., Michael L. Mack, Isabel Gauthier, and Thomas J. Palmeri. "Holistic processing of faces happens at a glance."*Vision research* 49.23 (2009): 2856-2861.
42. Ro, Tony, Charlotte Russell, and Nilli Lavie. "Changing faces: A detection advantage in the flicker paradigm." *Psychological Science* 12.1 (2001): 94-99.
43. Ruiz-Soler, Marcos, and Francesc S. Beltran. "Face perception: An integrative review of the role of spatial frequencies." *Psychological research* 70.4 (2006): 273-292.
44. Saxe, Rebecca. "Uniquely human social cognition." *Current opinion in neurobiology* 16.2 (2006): 235-239.
45. Schiller, P. H., J. G. Malpeli, and S. J. Schein. "Composition of geniculostriate input ot superior colliculus of the rhesus monkey." *Journal of Neurophysiology* 42.4 (1979): 1124-1133.
46. Seeck, M., C. M. Michel, N. Mainwaring, R. Cosgrove, H. Blume, J. Ives, T. Landis, and D. L. Schomer. "Evidence for rapid face recognition from human scalp and intracranial electrodes." *Neuroreport* 8.12 (1997): 2749-2754.
47. Sergent, Justine. "Influence of task and input factors on hemispheric involvement in face processing." *Journal of Experimental Psychology: Human Perception and Performance* 11.6 (1985): 846.
48. Tanaka, James W., and Joseph A. Sengco. "Features and their configuration in face recognition." *Memory & Cognition* 25.5 (1997): 583-592.
49. Tanaka, James W., and Martha J. Farah. "Parts and wholes in face recognition." *The Quarterly Journal of Experimental Psychology* 46.2 (1993): 225-245.
50. Tanaka, James, and Isabel Gauthier. "Expertise in object and face recognition."*Psychology of learning and motivation* 36 (1997): 83-125.
51. Thompson, Peter. "Margaret Thatcher: A new illusion." *Perception* 9.4 (1980): 483-484.

52. Torralba, Antonio, and Pawan Sinha. "Detecting faces in impoverished images." (2001).
53. Tsao, Doris Y., and Margaret S. Livingstone. "Mechanisms of face perception."*Annual review of neuroscience* 31 (2008): 411.
54. Veres-Injac, Bozana, and Adrian Schwaninger. "The time course of processing external and internal features of unfamiliar faces." *Psychological research* 73.1 (2009): 43-53.
55. Yin, Robert K. "Face recognition by brain-injured patients: a dissociable ability?." *Neuropsychologia* 8.4 (1970): 395-402.
56. Yin, Robert K. "Looking at upside-down faces." *Journal of experimental psychology* 81.1 (1969): 141.