The role of pigmentation in face perception

by

Richard Russell

B.A., Neuroscience
Pomona College, 1998

Submitted to the Department of Brain and Cognitive Sciences in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Cognitive Science

at the

Massachusetts Institute of Technology

September 2005

Signature of Author:____

Department of Brain and Cognitive Sciences
July 29, 2005

Certified by:____

Pawan Sinha
Associate Professor of Computational Science
Thesis Supervisor

Accepted by:____

Earl K. Miller
Picower Professor of Neuroscience
Chairman, Department Graduate Committee

The role of pigmentation in face perception

by

Richard Russell

Submitted to the Department of Brain and Cognitive Sciences in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Cognitive Science

Abstract

Faces each have distinct pigmentation as well as shape, which suggests that both cues may play a role in the perception of faces. However, there is a common implicit assumption that pigmentation cues are relatively unimportant, and so the role pigmentation plays in face perception has gone largely unexplored. This thesis is a systematic investigation of the role of pigmentation in face recognition, facial sex classification, and facial attractiveness. The present studies present evidence that pigmentation cues are in fact quite important for face perception. For face recognition, pigmentation cues are about as important as shape cues. Male and female faces differ consistently in their pigmentation, with female faces having more luminance contrast between the eyes and lips and the rest of the face than do male faces. This sex difference in pigmentation is used as a cue for judgments of facial sex classification and facial attractiveness. Together, these results implicate an important role for pigmentation, and open new avenues of research in the perception of faces.

Table of Contents

Acknowledgements

Foremost, I would like to thank my advisor, Pawan Sinha. Pawan's creativity and ability to consider problems in completely novel ways is unique. His highly visual approach to exploring and communicating ideas has greatly influenced me, and I hope one day to come close to his mastery of this difficult art. Perhaps most importantly, Pawan has trusted my abilities and has graciously allowed me follow my interests, even when they led me in directions unusual for a scientist. For that I am deeply grateful.

The other members of my committee, Nancy Kanwisher, Molly Potter and Ken Nakayama have given many comments and suggestions that have helped make this thesis more complete and rigorous.

It was a privilege and a pleasure to spend a summer at USC working with Irving Biederman and his student Marissa Nederhouser. Important parts of the work presented here were incubated during this time.

Some very capable undergraduate students helped conduct this research: Aayesha Siddiqui, Charisse Massey, Jill Jin, Nao Gamo, Omoleye Roberts, Jonathan Choi, Sarah Chu, Wendy Pierce, Mariana Recalde, Ann Breckenkamp, and Chen Zhao. I'm particularly thankful to them for having been so talented and personable, making easier my first experiences with the difficult task of mentoring.

Denise Heintze has been very helpful on numerous occasions in sorting out personal as well as administrative difficulties.

Two mentors from before graduate school, Debby Burke and Joe Devlin, were particularly influential in my decision to pursue a career in cognitive science.

I have learned a lot about vision and the mind from my colleagues and fellow students—Ben Balas, Matt Bronstad, Daniel Casasanto, Dav Clark, Jodi Davenport, Itamar Kahn, Rajesh Kasturirangan, Ethan Meyers, and Keith Thoresz, to name but a few. Many of them were good friends as well as colleagues, and helped make my time here enjoyable and meaningful.

My parents, Joe and Camille, and my older brother, Byron, helped to kindle my curiosity from a young age. I cannot express enough gratitude for their love, support, and encouragement.

I met my fiancée, Carrie Schneider, shortly after beginning this line of research. She has consistently supported and encouraged my work. More importantly, life has been better with her in it.

Chapter 1: Introduction

A central goal for the study of visual perception is determining what kinds of information are used by the visual system to perform specific tasks. An important special case of this goal arises in the domain of face perception. Here, the objective is to understand which visual features determine such attributes as the identity sex, or attractiveness of a face. This goal leads to many still open questions. What kinds of cues differ between individual faces and can also be reliably extracted by the human visual system so as to be effective for identification? What aspects of faces differ across the sexes? Can we identify subsets of facial information that are correlated with ratings of attractiveness? Despite several decades of research, we still do not have comprehensive answers to these questions. The goal of this thesis is to identify some of the important cues that the human visual system uses for various face perception tasks. To anticipate a point that will be described in greater detail below, the specific goal here is to investigate the importance of cues related to facial pigmentation for judgments of identity, gender and attractiveness. But first, let us lead up to this goal in a principled manner.

For a cue to be useful for these judgments, it must vary between individual faces (for recognition), or between the classes of faces to be distinguished (for a classification judgment such as sex classification). To determine what kinds of visual cues differ between faces, it is useful to step back and ask what kinds of cues are available to visual perception more generally. The objects of visual perception are surfaces (Gibson 1979). There are three variables that determine the visual appearance of a surface: 1) the illumination of the surface 2) the shape of the surface, and 3) the reflectance properties of the surface. The set of all cues that could possibly be useful for the visual perception of faces can be grouped into three classes: lighting cues, shape cues, and surface reflectance cues. To address the question of what cues are useful for the perception of faces, I will proceed from this top level, asking whether each of these three classes of cues is involved in performing specific judgments about faces.

The first cue, illumination, differs from the other two cues in that it is not intrinsic to a particular face, but is a product of the interaction between the face and its environment. Illumination has a large effect on the image level appearance of a face, a fact well known to artists and machine vision researchers. When people are asked to match different images of the same (unfamiliar) face, performance is worse when the two images are differently illuminated (Braje et al 1998; Hill and Bruce 1996). However, people are very accurate at naming familiar faces under different illumination (Moses et al 1994). This finding fits our informal experience, in which we are able to recognize people under a wide variety of lighting conditions, and do not experience the apparent identity or sex of a person as changing when we walk with them, say, from indoor to outdoor lighting.

There are certain special conditions when illumination can have a dramatic impact on the appearance of a face, most notably lighting from below. However, this is a rare occurrence; under natural conditions, including artificial lighting, faces are almost always lit from at or above the horizon. Consistent with the notion that the representation of facial identity includes the statistical regularity of lighting from above, faces look odd when lit from below, and several researchers have found that face recognition is impaired by bottom lighting (Enns and Shore 1997; Hill and Bruce 1996; Johnston et al 1992;

McMullen et al 2000). These findings overall are consistent with the notion that the human visual system does make some use of lighting regularities for face recognition. There have not been any studies evaluating the utility of illumination cues for classifying faces as male or female, or for making judgments of attractiveness. It would not be surprising if these judgments were impaired or affected by illumination from below, as with face recognition, because the visual system has little experience with such conditions.

Unlike illumination, the shape and surface reflectance properties of a face do not change through the course of a day. Given that the identity and sex of faces do not also change through the day, this stability of shape and surface reflectance properties makes them more useful for making judgments about these properties. For this reason, I will not discuss further the role of illumination, and will move now to consider the other two classes of cues that are available to visually distinguish faces: the shape and reflectance properties of the face surface.

Surface reflectance properties have elsewhere been referred to as 'color' or 'texture'. Unfortunately these terms also have colloquial meanings that refer to subsets of surface reflectance properties (hue and spatial variation in reflectance, respectively). Thus, for the sake of clarity as well as brevity, I will use the term 'pigmentation' here to refer to the surface reflectance properties of the face. By 'pigmentation' we refer to the way that the face surface reflects light, and by 'shape' we refer to the relative locations in space of the face surface. Shape cues include boundaries, junctions, and intersections, as well any other cue that gives information about the location of a surface in space, such as stereo disparity and motion parallax. One group of cues that have recently received a lot of attention in the face recognition literature, 'second order relations' or 'configural properties' (which consist of the distances between facial features such as the eyes and mouth), are a subclass of shape. Pigmentation cues include albedo, hue, texture, translucence, specularity, and any other property of surface reflectance. The pigmentation of a face has spatial extent and can vary across the face, and so the relative reflectance of those features or parts of features, such as the relative darkness of parts of the eye or of the entire eye region and the cheek, are also components of pigmentation.

This division of cues into two classes is not perfect. For example, should a border defined by a luminance gradient, such as the junction of the iris and sclera be classified as a shape or pigmentation cue? Because faces share a common configuration of parts, I classify such borders as shape cues when they are common to all faces (e.g. the iris-sclera border), but as pigmentation cues when they are unique to an individual (e.g. moles and freckles). Shading is another cue that does not fit neatly into this division, as it is an interaction between the illumination and shape of the face. For this reason I do not count it as either shape or pigmentation. In the studies presented here, I take steps to minimize shading cues, and in one study (Chapter 2), I use two different methods to separate shape and pigmentation cues, one of which groups shading cues with shape, and the other of which groups shading cues with pigmentation. Though the present distinction between shape and pigmentation cannot completely separate shape and pigmentation cues, it does serve to distinguish between them in the vast majority of cases.

This division between shape and pigmentation has also been used to investigate human recognition of non-face objects, although in that literature, pigmentation has typically been referred to as 'color' or 'surface properties'. Much of this work has compared human ability to recognize objects from line drawings or images with pigmentation cues, such as photographs. The line drawings contain shape cues, but not pigmentation cues, while photographs contain both kinds of cues. If pigmentation cues are important for object recognition, performance should be significantly poorer with the line drawing stimuli that lack pigmentation cues than with the continuous tone images that have pigmentation cues. Because recognition of line drawings is as good (Biederman and Ju 1988; Davidoff and Ostergaard 1988; Ostergaard and Davidoff 1985) or almost as good (Humphrey 1994; Price and Humphreys 1989; Wurm et al 1993) as recognition of photographs, there is a consensus that, pigmentation is a less important cue than shape for object recognition (Palmer 1999; Tanaka et al 2001; Ullman 1996).

While no explicit claim has been made in the face perception literature that pigmentation is less important than shape, three lines of evidence suggest that there is an implicit assumption that shape is more important. The first is that line drawings and other non-pigmented representations are commonly used as stimuli for face perception experiments. In other studies, pigmented faces are used as stimuli, but all the faces have the same pigmentation, meaning that there are not useful pigmentation cues. It would not make sense to use stimuli that throw out an entire class of cues unless one believed that those cues were relatively unimportant. Similarly, many models of face recognition use only shape cues, such as distances between features, in their representations. As with using unpigmented stimuli, this only makes sense if it is assumed that the pigmentation cues being ignored are not as important as the shape cues being retained. Finally, studies investigating aspects of shape are far more common than studies of pigmentation. Together, these observations suggest that the face research community does not believe that pigmentation is a particularly significant component of face representations.

Yet, there are reasons to believe that pigmentation may indeed be important for face perception. Unlike for objects, it is much more difficult to recognize a face from a line drawing than from a photograph (Bruce et al 1992; Davies et al 1978; Leder 1999; Rhodes et al 1987), suggesting that the pigmentation cues thrown away by such a representation are indeed important. Two seminal studies (Bruce and Langton 1994; O'Toole et al 1999) using uniformly pigmented models of faces (derived from laser scans) found evidence that pigmentation cues play an important role in recognition. For sex classification, there is evidence that pigmentation may actually be *more* important than shape information (Hill et al 1995). In this cue conflict study the shape and pigmentation representations of laser scanned faces were swapped, producing faces with the shape of a female and the pigmentation of a male (or the reverse), and subjects were more likely to classify the faces using the pigmentation cues than the shape cues. Finally, the universal and old practice of using cosmetics in the attempt to make the face more beautiful suggests that pigmentation may also play a role in facial attractiveness. Thus, for each of these three judgments about faces, there is reason to doubt the assumption that pigmentation is unimportant.

As the evidence presented above suggests that pigmentation may be an important, unexplored aspect of face perception, the goal of the current body of work is to evaluate the role of pigmentation in face perception. In this thesis, I investigate three ecologically important judgments about faces: face recognition, facial sex classification, and facial attractiveness. I have chosen to study these rather than other judgments about faces because I suspected that they would be more likely to yield positive results. The perception of facial expressions of emotion is unlikely to involve pigmentation cues, as the different emotions are produced by deformations of the face shape, without change to pigmentation (with the possible exception of embarrassment). Racial classification of faces is already presumed to involve the use of pigmentation cues. I have performed no investigations of the perception of facial age, though I believe that it may be a fruitful avenue for future studies.

Chapter 2 presents an evaluation of whether contrast negation disrupts the use of second-order relations (a subclass of shape cues) or the use of pigmentation cues. In Chapter 3, I report an investigation of the relative utility of shape and pigmentation cues for matching unfamiliar faces. The effect of vertical inversion on the perception of these cues is also considered. Chapter 4 evaluates the utility of shape and pigmentation cues for the recognition of familiar faces. In Chapter 5, I present evidence for consistent differences between the pigmentation of male and female faces, and that people use these pigmentation differences to decide whether faces are male or female. Chapter 6 describes the effect of increasing these sex differences in pigmentation on the rated attractiveness of male and female faces. Chapter 7 concludes by discussing the limitations and contributions of this work, as well as future directions of research.

References

Biederman I, Ju G, 1988 "Surface versus Edge-Based Determinants of Visual Recognition" *Cognitive Psychology* **20** 38-64

Braje W L, Kersten D, Tarr M J, Troje N F, 1998 "Illumination Effects in Face Recognition" *Psychobiology* **26** 371-380

Bruce V, Hanna E, Dench N, Healey P, Burton M, 1992 "The importance of 'mass' in line drawings of faces" *Applied Cognitive Psychology* **6** 619-628

Bruce V, Langton S, 1994 "The use of pigmentation and shading information in recognizing the sex and identities of faces" *Perception* **23** 803-22

Davidoff J, Ostergaard A, 1988 "The role of color in categorical judgments" *Quarterly Journal of Experimental Psychology* **40** 533-544

Davies G M, Ellis H D, Sheperd J W, 1978 "Face recognition accuracy as a function of mode of representation" *Journal of Applied Psychology* **63** 180-187

Enns J T, Shore D I, 1997 "Separate influences of orientation and lighting in the inverted-face effect" *Perception & Psychophysics* **59** 23-31

Gibson J, J., 1979 *The ecological approach to visual perception* (Hillsdale, NJ: Lawrence Erlbaum Associates)

Hill H, Bruce V, 1996 "Effects of lighting on the perception of facial surfaces" *Journal of Experimental Psychology: Human Perception and Performance* **22** 986-1004

Hill H, Bruce V, Akamatsu S, 1995 "Perceiving the sex and race of faces: the role of shape and colour" *Proceedings of the Royal Society of London B* **261** 367-73

Humphrey G, 1994 "The role of surface information in object recognition: studies of visual form agnosic and normal subjects" *Perception* **23** 1457-1481

Johnston A, Hill H, Carman N, 1992 "Recognising faces: effects of lighting direction, inversion, and brightness reversal" *Perception* **21** 365-375

Leder H, 1999 "Matching person identity from facial line drawings" *Perception* **28** 1171-1175

McMullen P A, Shore D I, Henderson R B, 2000 "Testing a two-component model of face identification: Effects of inversion, contrast reversal, and direction of lighting" *Perception* **29** 609-619

Moses Y, Adini Y, Ullman S, 1994 "Face recognition: the problem of compensating for illumination changes" *Proceedings of the European Conference on Computer Vision* 286-296

Ostergaard A, Davidoff J, 1985 "Some effects of color on naming and recognition of objects" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **11** 579-587

O'Toole A J, Vetter T, Blanz V, 1999 "Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: an application of three-dimensional morphing" *Vision Research* **39** 3145-3155

Palmer S E, 1999 *Vision Science: Photons to Phenomenology* (Cambridge, Massachusetts: MIT Press)

Price C J, Humphreys G W, 1989 "The effects of surface detail on object categorization and naming" *Quarterly Journal of Experimental Psychology* **41A** 797-828

Rhodes G, Brennan S E, Carey S, 1987 "Recognition and ratings of caricatures: Implications for mental representations of faces" *Cognitive Psychology* **19** 473-497

Tanaka J, Weiskopf D, Williams P, 2001 "The role of color in high-level vision" *Trends in Cognitive Sciences* **5** 211-215

Ullman S, 1996 *High-Level Vision: Object Recognition and Visual Cognition* (Cambridge, Massachusetts: MIT Press)

Wurm L H, Legge G E, Isenberg L M, Luebker A, 1993 "Color improves object recognition in normal and low vision" *Journal of Experimental Psychology: Human Perception and Performance* **19** 899-911

Chapter 2: Is pigmentation important for face recognition?

Evidence from contrast negation

Published as: Russell, R., Sinha, P., Biederman, I., & Nederhouser, M. (*In Press*). Is pigmentation important for face recognition? Evidence from contrast negation. *Perception*

Abstract

It is extraordinarily difficult to recognize a face in an image with negated contrast, as in a photographic negative. The variation among faces can be partitioned into two general sources: a) shape and b) surface reflectance, here termed "pigmentation.". To determine whether negation differentially affects the processing of shape or pigmentation, we made two sets of faces where the individual faces differed only in shape in one set and only in pigmentation in the other. Surprisingly, matching performance was significantly impaired by contrast negation only when the faces varied in pigmentation. This provides evidence that the perception of pigmentation, not shape, is selectively disrupted by negation, and by extension, that pigmentation contributes to the neural representation of face identity.

Introduction

Faces are particularly difficult to recognize in photographic negatives. This phenomenon was first reported in the scientific literature by Galper (Galper 1970; Galper and Hochberg 1971), and has since been studied extensively in the vision science community (Bruce and Langton 1994; Bruce and Young 1998; George et al 1999; Hayes et al 1986; Johnston et al 1992; Kemp et al 1990; Kemp et al 1996; Liu et al 1999; Phillips 1972; White 2001), with the belief that determining how recognition can be impaired helps us understand how it works under normal conditions. Negation reverses the contrast polarities of an image, making black areas white, light gray areas dark gray, and so forth. It is a reversible manipulation that does not remove any information from the image. Though no information is lost, our *ability* to use the information in the image is severely compromised. This suggests that some normally useful information is rendered unusable by negation. It cannot be the only useful information for face recognition, as negation leads to partial rather than total impairment of performance—those cues still usable after negation should also be usable under normal conditions. But whatever information is compromised by negation is likely to be among the important components of facial identity under normal conditions.

One cue that is likely disrupted by negation is the pattern of shading across a face, which is a product of the interaction between the shape of the face and the direction of lighting. Faces are normally viewed in positive contrast, with lighting from above. Both negation and lighting from below disrupt the normal patterns of shading across a face and impair recognition. Because patterns of shading can be used to estimate 3-D shape (Horn 1986; Johnston and Passmore 1994; Ramachandran 1988), abnormal patterns of shading may cause impairment in the ability to determine the 3-D surface of the face (Hill and Bruce 1996; Johnston et al 1992; Kemp et al 1996). However, faces presented with abnormal patterns of shading but veridical 3-D stereo cues are still difficult to recognize, suggesting that abnormal shading may impair 2-D pattern processing rather than 3-D processing of shapes (Liu et al 2000). The strongest evidence that abnormal shading patterns disrupt recognition comes from work with unpigmented 3-D models of faces. With images from such models, matching performance is about 10% better with positive contrast when the head model is lit from above, but is actually better with negative contrast when the lighting is from below (Liu et al 1999). When lighting was at 0° elevation (front lighting) performance was about equal with positive and negative contrast, suggesting that shading cues are minimized or eliminated by this kind of lighting. However, with normally pigmented faces, performance is significantly better in positive contrast with either top or front lighting (Johnston et al 1992). This suggests that disruption in shading is only a partial explanation of why faces are difficult to recognize in negative.

Negation has been proposed to disrupt the use of two other cues; 'second-order relations' (the distances between facial features (Diamond and Carey 1986), sometimes called 'configuration') and pigmentation. Second-order relations are widely believed to play a key role in the perceptual representation of faces. Deficits in recognition performance with vertically inverted faces have been attributed to impaired extraction of second-order relations in such stimuli (Freire et al 2000; Leder and Bruce 2000; Leder et al 2001;

LeGrand et al 2001). Several authors have raised the possibility that the perception of second-order relations may also be impaired for negated faces (Hole et al 1999; Kemp et al 1990; Lewis and Johnston 1997; White 2001). Second-order relations are a subset of the larger class of shape cues that can be used to differentiate faces. We can ask whether shape cues more generally are disrupted by negation. If shape cues are not disrupted by negation, it follows that the subordinate class of second-order relations are not disrupted by negation.

The processing of another cue—'pigmentation"--has also been proposed to be disrupted by negation. We extend the meaning of pigmentation in the current article to refer to *all* surface reflectance properties. Most prominent among these reflectance properties are albedo (the proportion of light of all wavelengths reflected by the surface, affecting how light or dark the surface appears) and hue (differential reflectance of particular wavelengths of light results in the surface appearing to have a particular hue, such as red or blue). However, there are also many other properties of surface reflectance, such as texture, (spatial variation in how light is reflected) specularity, and translucency (both are functions of the way light is scattered by the surface). Shading cues, discussed above, do not fit cleanly into the proposed distinction between shape and pigmentation. For pragmatic purposes in the current study, we will minimize shading cues and include them with shape cues in one experiment and with pigmentation cues in another. For the sake of brevity, surface reflectance properties have elsewhere been referred to as 'color' or 'texture'. Unfortunately these terms also have colloquial meanings that refer to subsets of surface reflectance properties (hue and spatial variation in reflectance, respectively). Thus we use the word 'pigmentation' for the sake of both brevity and clarity.

The evidence that the use of pigmentation is disrupted by negation derives from studies using uniformly pigmented, 3-dimensional face stimuli (Bruce and Langton 1994; Liu et al 1999). These images are of models with the 3-dimensional shape of a particular face (as derived from laser scans) that reflect light uniformly from all points on the surface. Normal faces reflect light differently at different points across the face. The disruption in recognition performance caused by negation of these images of uniformly pigmented head models is much smaller than that caused by negation of normally pigmented faces. This difference has been interpreted as evidence that the perception of pigmentation cues to identity is what is normally disrupted by negation. However, it is also possible that there is no effect of negation with the uniformly pigmented faces because they are processed by the visual system in a fundamentally different way than normal faces. In particular, it has been proposed that the pattern of dark areas (corresponding to the eyes and lips) surrounded by lighter areas that is common to faces under normal lighting conditions underlies the process of face detection (Sinha 2002; Watt 1994). If this pattern of luminance is necessary for activating normal face processing, images without this pattern, such as images produced from the uniformly pigmented face models, may not receive the same processing as normal faces. For this reason, it would be desirable to determine whether pigmentation cues of faces with normal, face-like patterns of pigmentation are disrupted by negation.

For the present study we sought to determine whether shape (including second-order relations) or pigmentation is disrupted by negation. Our strategy was to create sets of faces that differed from one another in terms of only their shape or only their pigmentation, respectively the 'Shape' and 'Pigmentation' sets. However, unlike unpigmented face stimuli, the faces in both of these sets look like normal faces, with face-like shape and pigmentation. Subjects performed a delayed match-to-sample, two-alternative forced-choice task, in which they had to distinguish between two faces from either the Shape or Pigmentation set. In this way we controlled which cues were available for distinguishing the two faces. By comparing performance with shape and pigmentation cues in recognizing faces in positive or negative contrast, we could determine whether either or both cues are disrupted by negation. If the use of shape is disrupted by negation, performance with the shape set should have been significantly worse with negated faces than with positive faces. If the use of pigmentation is disrupted by negation, performance with the Pigmentation set should have been significantly worse with negated faces than with positive faces.

Methods

We conducted two experiments that differed only in terms of the stimuli and the presentation timing. The first experiment used artificially generated faces for which shape and pigmentation could be manipulated independently. To test whether the results of the first experiment generalize to real-world facial images, the second experiment used manipulated photographs of actual faces instead of artificial stimuli. These stimuli retained potentially relevant pigmentation cues such as specularity and translucency that are lost in 3-dimensional models such as laser-scanned images or the artificial stimuli used in the first experiment. The stimuli in the two experiments differed in how shape and pigmentation were defined. For the artificial stimuli of the first experiment, shape was defined as the location of the face surface in 3-dimensions, and pigmentation was defined as the way in which light reflected off the face surface. For the photographic stimuli of the second experiment, shape was defined as the locations of face contours in the 2-dimensional image space, and pigmentation was defined as the pixel values at different locations of the image. A consequence of the different definitions of shape and pigmentation in the two experiments is that in the Experiment 1, faces in the Shape set differ in shading, while in Experiment 2, the Pigmentation set contains the differences in shading. If the shading cues play a large role, we will expect different relative performances in the two experiments. However, we selected the illumination of the faces in both experiments to minimize shading differences between the faces (described below).

Subjects

A total of 28 subjects participated in this study; 14 in experiment 1 and 14 in experiment 2. All were contacted through the MIT Brain & Cognitive Sciences subject pool, were naïve to the purpose of the study, and had normal or corrected-to-normal vision. Though the stimuli examples shown below are grayscale, both experiments used full-color stimuli.

Experimental design

Both experiments used the same delayed match-to-sample, two-alternative forced-choice task, with the same experimental design (Figure 1). In each trial of this task, the observer saw a fixation dot followed by a sample face, then a visual noise mask, then a blank screen. Next, two faces were presented side by side in the center of the screen. One of the two faces, the sample, had been presented just previously, and the other was a distractor face that differed from the target in shape only (Shape condition), in pigmentation only (Pigmentation condition), or in both shape and pigmentation (Shape + Pigmentation condition). The task was to decide which of the two faces matched the sample, and to press the corresponding key as quickly as possible. Trials from the different conditions were intermixed , and the left-right ordering of target and distractor was counterbalanced. There were eight male and eight female faces in each condition. All the faces in a trial were presented in positive contrast or in negative contrast. Each of the eight faces was paired with every other face of its same sex for each condition, resulting in 28 (pairs) x 2 (sex) x 3 (condition) x 2 (contrast) x 2 (left-right counterbalance) = 672 trials per observer. Stimuli were presented in Matlab (The MathWorks, Inc.) using the Psychophysics Toolbox extensions (Brainard 1997; Pelli 1997).

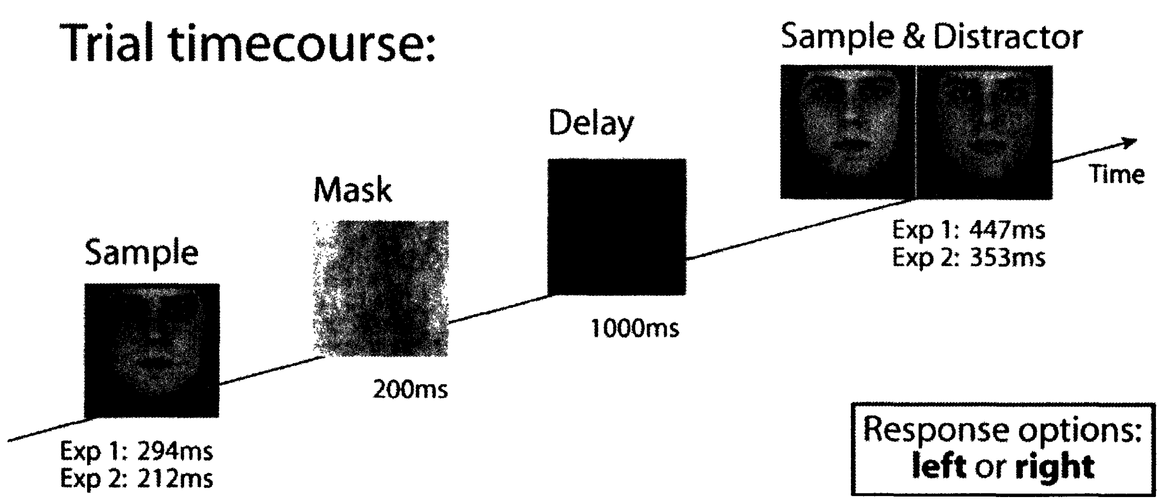

Figure 1. Experimental design. On each trial, subjects viewed a face, followed by a visual noise mask, then a delay. Two faces were then presented; one face (the target) was the identical image as the sample, and the other face (the distractor) was drawn from the same category of faces. For example, if the face in the upper left of Figure 2 was the target, the face to the right could be the distractor, but not one of the faces below. In this way, the information distinguishing the target and distractor faces defined the condition: Shape only, Pigmentation only, and Shape + Pigmentation.

Experiment 1 stimuli and presentation

Stimuli for the first experiment were produced using FaceGen Modeller 3.0 (Singular Inversions Inc.) software that creates 3-dimensional models of artificial faces, with separate representations for shape and for pigmentation (referred to as 'texture' by the software). To make the faces appear as naturalistic as possible, "photofits" were used, in which frontal and profile photographs of actual faces are used to determine the shape and pigmentation settings of the models. The photofit system is proprietary, and thus we do not know exactly how the 2-dimensional photographic images were used to create separate pigmentation and 3-dimensional shape. However, these details are irrelevant for the present purposes; the important point is face representations could be created that varied in terms of both their 3-dimensional shape and texture (pigmentation) models, and that these separate models could subsequently be altered independently. Photofits of eight male and eight female faces were created. The people on whom the photofits were based ranged in age from 18 to 25, with a mean age of 21. Because it would be less surprising if pigmentation were used to distinguish among faces of different races, all the faces were of the same race (Caucasian). The models generated directly from the photofits differed from one another in terms of both their shape and their pigmentation, and so they formed the stimuli for the Shape + Pigmentation condition. To create the stimuli for the Pigmentation condition, the shape models for each of the photofit heads were given the same settings (the software's 20 year old male or female Caucasian average), producing a set of faces each with distinct pigmentation, but the same (average) shape. To create the stimuli for the Shape condition, the texture models for each of the photofit heads were given the same settings (the software's 20 year old male or female Caucasian average), producing a set of faces each with distinct shape, but the same (average) pigmentation. The models were rendered with a single light source at 0^{o}, with the lighting source set to 60% with 40% ambient brightness. These lighting conditions were selected to minimize effects of shading cues, and were applied consistently in order to eliminate effects of varied lighting direction (Braje et al 1998; Moses et al 1994). A 7/8 viewpoint was selected to make both shape and pigmentation cues readily apparent. Presentation times for experiment 1 were: probe 294msec, mask 200msec, delay 1000msec, target and distractor 447msec. These times were selected to elicit accuracy levels of approximately 75%, in the center of the range between chance and perfect performance.

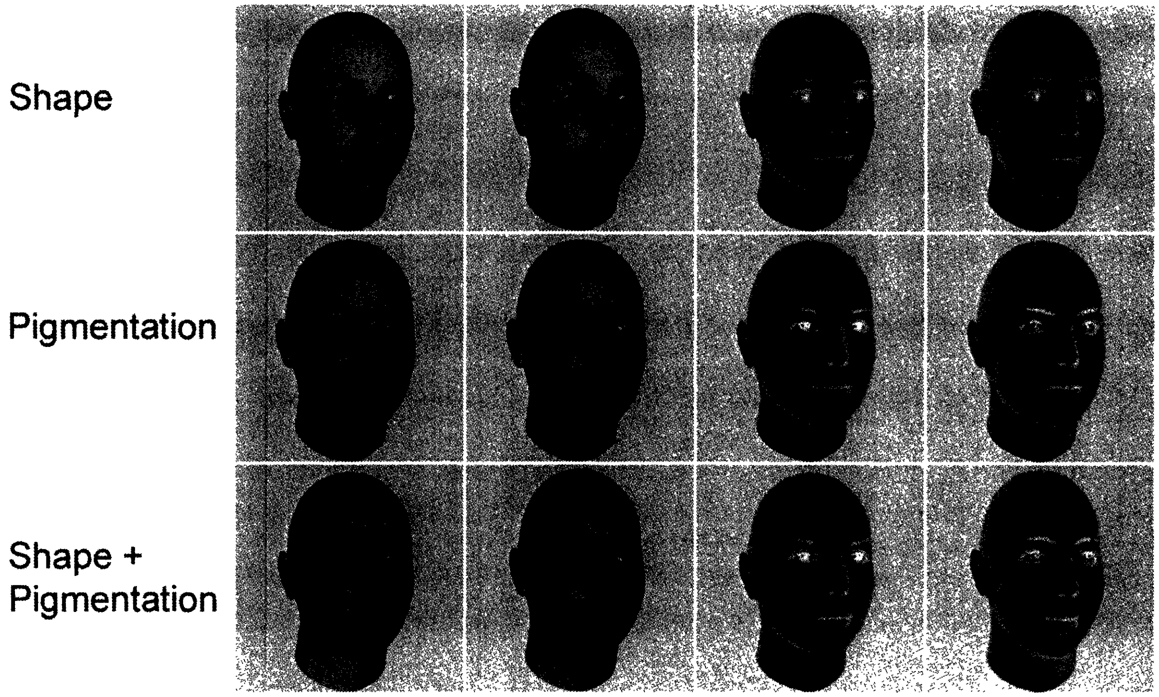

Figure 2. Examples of artificial stimuli. The Shape condition faces along the top rows of each set have the same pigmentation as one another and are distinguishable only by shape. The Pigmentation condition faces along the middle rows have the same shape and are distinguishable only by pigmentation. The Shape + Pigmentation faces along the bottom rows, like normal faces, are distinguishable by both shape and pigmentation. In both sets, the left two columns show faces with positive contrast, and the right two columns show the same faces with negative contrast.

Experiment 2 stimuli and presentation

We took frontal photographs of eight male and eight female Caucasian faces, ranging in age from 18 to 25, with a mean age of 20. Large, diffuse lighting centered at 0° elevation was used to eliminate cast shadows and to minimize the effects of shading cues. The heads of the models were placed in the same location relative to the lighting sources and the rest of the room, to ensure equivalent lighting the faces. These photographs were then manipulated using Morph Man 3.0 (Stoik Imaging). For each sex, the eight original faces were morphed together to produce an average face. We created the stimuli for the Shape condition by warping (moving pixels in the image plane without averaging pixel values, as done in morphing) this average face into the shape of each of the original eight faces, producing eight new faces, each with the same (average) pigmentation, but with distinct shape. Similarly, we created the stimuli for the Pigmentation condition by warping each of the eight original faces into the shape of the average face, which produced eight faces, each with the same (average) shape, but with distinct pigmentation. The original photographs of the eight faces of each sex, differing from one another in terms of both their shape and pigmentation, formed the stimuli for the Shape + Pigmentation condition. Presentation times for experiment 2 were: probe 212msec, mask 200msec, delay 1000msec, target and distractor 353msec. The probe and target were presented more briefly in experiment 2 than in experiment 1 in order to maintain accuracy levels around 75%.

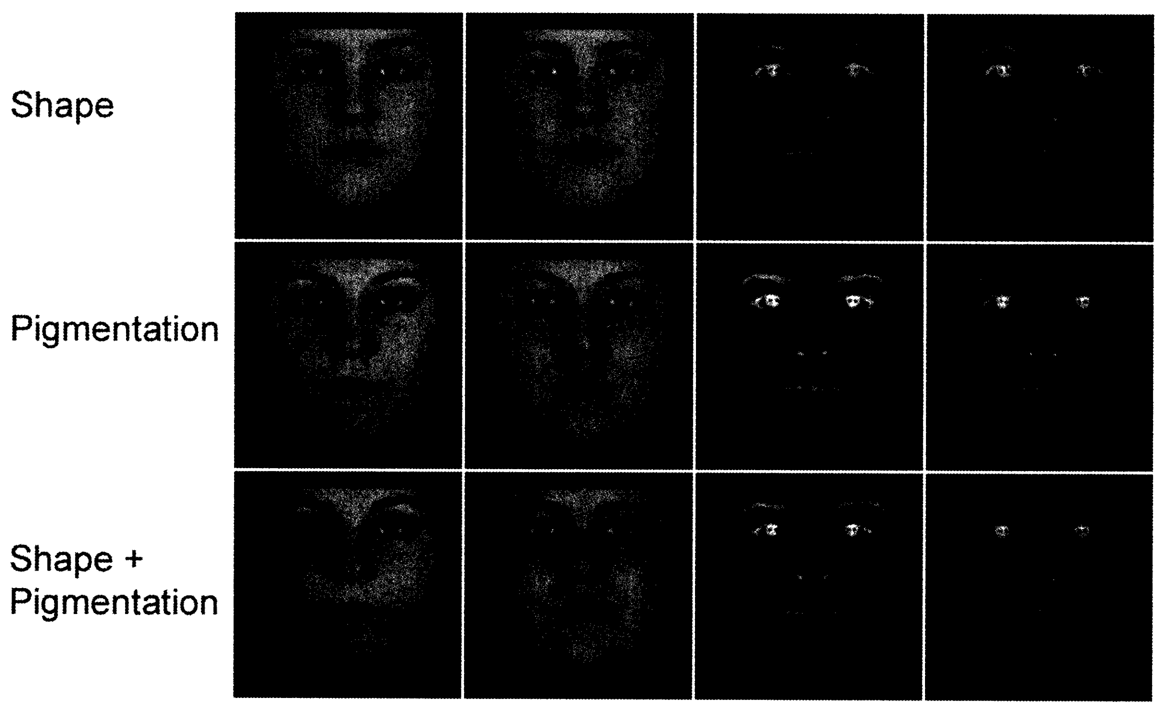

Figure 3. Examples of photographic stimuli. The faces are arrayed the same way as in Figure 2.

Results

Experiment 1

The first experiment used artificial stimuli (Fig. 2 and Methods). We recorded observers' performance on a delayed match-to-sample, two-alternative, forced-choice task (Fig. 1 and Methods). Figure 4 shows that negation impaired performance (producing higher error rates) in the Pigmentation and Shape + Pigmentation conditions, but not the Shape condition. That is, when only shape cues were available for the discrimination of the faces, performance was not significantly worse with negated faces than positive faces.

There was a main effect of condition, $F(2,26)$ 14.1, $p < 0.001$, with the best performance when both shape and pigmentation cues were available. There was also a main effect of contrast, $F(1,13) = 62.2$, $p < 0.001$, with performance worse on negative than positive faces, consistent with all other studies on the effect of negation. Almost all this cost of negation was on pigmentation cues, producing a significant interaction between condition and contrast, $F(2,16) = 3.7$, $p < 0.05$. Post-hoc comparisons of performance on positive and negative contrast faces within each condition found significant differences ($p<0.05$, Bonferroni corrected) between the positive and negative versions only for the two conditions with pigmentation cues (Pigmentation and Shape + Pigmentation), but not for the Shape condition.

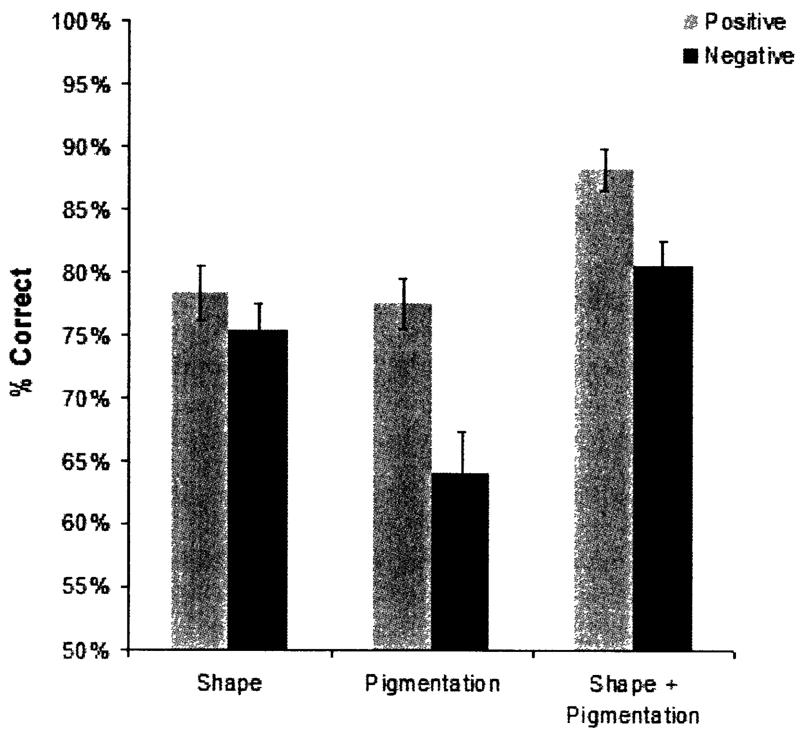

Figure 4. Results of Experiment 1, with artificial faces, expressed as percentage of correct responses. Error bars indicate standard errors.

These were not floor or ceiling effects in that the Shape condition, which showed no effect of negation, was midway between the Pigmentation and Shape + Pigmentation conditions.

Experiment 2

Because the artificial stimuli in Experiment 1 were not entirely naturalistic in appearance, the second experiment used stimuli created from photographs of actual faces (Fig. 3 and Methods) to assess whether the results of Experiment 1 do indeed generalize to more naturalistic face images. The same task and experimental design was used as in Experiment 1 (Fig. 1 and Methods). The pattern of results from Experiment 2 was the same as that from the first experiment (Fig. 5). There were main effects of condition,

$F(2,26) = 25.1$, $p < 0.001$, with performance best in the Shape + Pigmentation condition), and contrast, $F(1,13) = 30.5$, $p < 0.001$, with performance better on positive than negative faces). There was also a significant interaction between condition and contrast, $F = 5.0$, $p < 0.05$, and post-hoc comparisons of performance on positive and negative contrast faces within each condition found significant differences ($p<0.05$, Bonferroni corrected) between positive and negative conditions only for the two conditions with pigmentation cues (Pigmentation and Shape + Pigmentation), but not for the Shape condition.

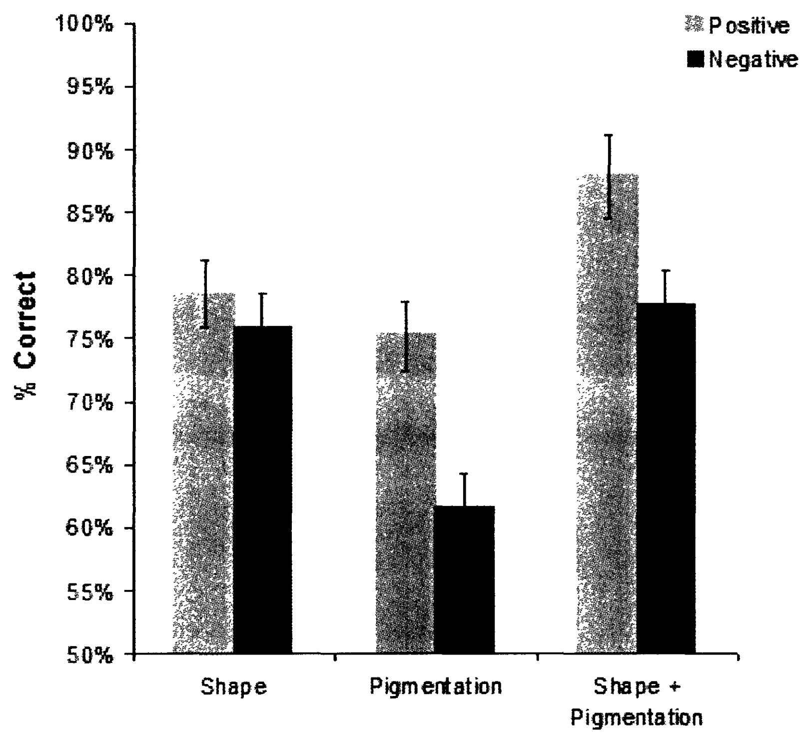

Figure 5. Results of Experiment 2, with photographic faces. expressed as percentage of correct responses. Error bars indicate standard errors.

Data from a more naturalistic set of photographic stimuli thus provide further evidence that faces are difficult to recognize in contrast negative because pigmentation cues are rendered unusable. An alternative hypothesis, that perception of second-order relations is disrupted by negation, was not supported.

Discussion

Why is it difficult to recognize faces in photographic negative? In two experiments, face recognition was significantly disrupted by negation when pigmentation cues provided the discriminable information, but when only shape cues were discriminable, negation had no significant effect. These results are consistent with the notion that negation disrupts our ability to use pigmentation cues. Because negation severely disrupts our ability to recognize faces, an implication of this finding is that pigmentation is one of the main components of the representation of face identity. Our results are consistent with those from other studies finding a role for pigmentation in face recognition (Bruce and Langton 1994; O'Toole et al 1999), as well as facial sex classification (Hill et al 1995; Tarr et al 2001) and attractiveness (Jones et al 2004; Russell 2003). Additional work from our labs suggests that the patterns of pigmentation around the eye regions play a particularly important role in judgments of identity (Sinha and Gilad Under review).

The findings also provide evidence against an alternative hypothesis of why negation impairs recognition performance: that the perception of second-order relations of features is disrupted by negation. In each of two experiments, there was no significant effect of negation when only shape cues were available. Discontinuous contours such as line drawings and high-pass filtered images are not disrupted by negation (Hayes et al 1986). Only the faces in the Shape condition differed in terms of discontinuous contours, and it is possible that subjects adopted a strategy in the Shape condition of focusing only on discontinuous contours, for example the outline of the face. Two lines of evidence argue against this possibility. The experiments used randomized rather than blocked presentation, such that subjects could not anticipate which cues would be useful for subsequent matching. Secondly, while there are readily distinguishable differences in the discontinuous contours of the photographic stimuli of Experiment 2, there are not such readily distinguishable differences among the artificial faces of Experiment 1 (compare Figures 2 and 3).

This finding that the perception of shape cues, including second-order relations of features, is not disrupted by negation is somewhat at odds with two previous studies. Kemp et al. (1990) found that when viewing negative images, subjects were less sensitive to changes to the distances between features, which would suggest that negation disrupts second-order relations. However, the stimuli for those experiments consisted of two-tone (black or white) images of a single face. A subsequent study (Liu and Chaudhuri 1997) found that performance on an old-new task (without manipulation of distances between features) was more impaired by negation with two-tone than continuous tone images. This suggests that the Kemp et al. 1990 finding may have been an artifact of the stimulus type rather than the manipulation of second-order relations. Lewis & Johnston (1997) reported that with faces presented in negative contrast, subjects were slower to determine whether a pair of faces differed in terms of being 'thatcherised' or not. This finding was interpreted as evidence that negation disrupts the perception of second-order relations ('configuration' in their terminology). Given that thatcherisation does not change the second-order relations of a face, it is not clear why the slower performance with negated faces should be interpreted in this way. We believe that the present findings constitute a more direct test of the hypothesis that shape (including second-order relations) is disrupted by negation.

Because negation is so detrimental to recognition performance, the observation that negation specifically disrupts pigmentation cues implies that pigmentation is an important component of identity for normal recognition (i.e. with positive contrast). However, negation is not the sole litmus test of importance for recognition, and so the results do not imply that shape is not also an important component of identity for normal recognition. In addition to evidence from many other studies, two findings from the present investigation suggest that shape is indeed important. The first is that performance with positive contrast in the Shape condition was well above chance, which indicates that shape alone can be used to recognize faces. The second is that performance with negative contrast in the Shape + Pigmentation condition was also well above chance, illustrating that faces can still be recognized after negation. Indeed, performance with negative contrast in the Shape + Pigmentation condition was almost the same as performance with positive contrast in the Shape condition, consistent with the idea that negation selectively disrupted pigmentation cues, leaving shape cues unaffected.

We have used a task that involved the matching of unfamiliar faces, and so we cannot be certain that these results would generalize to recognition memory for familiar faces. However, several reasons suggest that such generalization is likely. Many studies of face recognition, including a majority of those cited here, use face matching rather than familiar recognition tasks, and most problems associated with familiar recognition, such as inversion and negation, are also found with face matching, suggesting important commonalities between matching and recognition. Our task also had a memory component—subjects performed the matching task after viewing a visual noise mask and more than a second of delay—indicating that perceptual comparison alone was not sufficient for performance. In any case, matching unfamiliar faces is an important task worth investigating in its own right (Hancock et al 2000), for its application to eyewitness testimony and for the construction of automatic systems of face recognition.

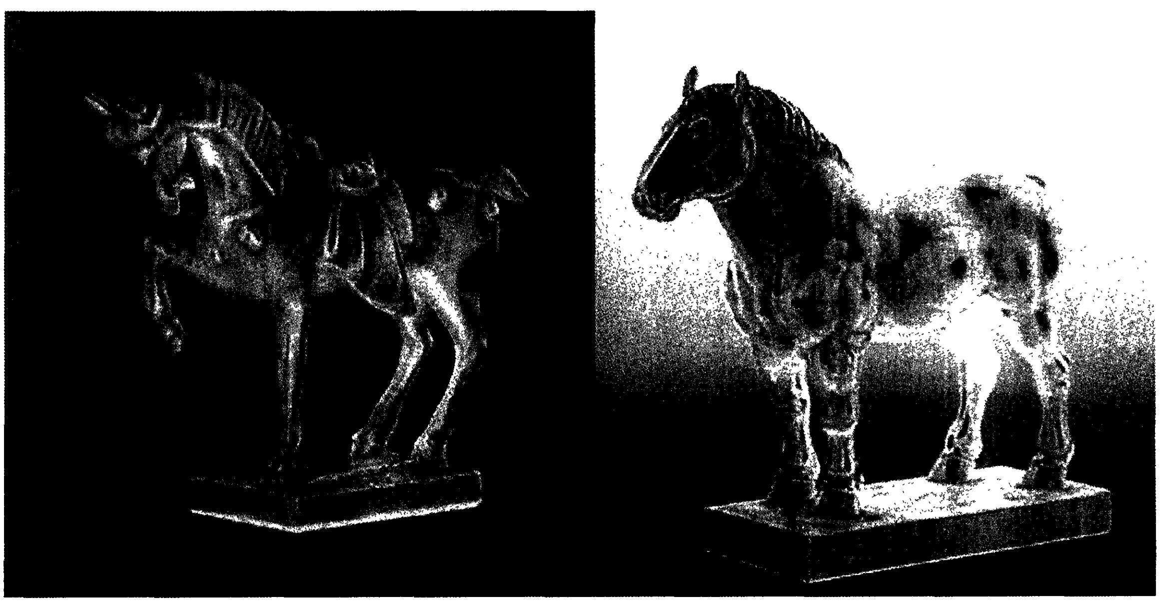

Figure 6. Two negative contrast images. It is readily apparent that both images depict statues of horses. However, it is much more difficult to determine that the statue on the left is constructed of bronze, and the statue on the right is of wood. In positive contrast, recognizing these materials is easy.

How specific are these findings to faces? Negation has been found not to disrupt subordinate-level recognition of at least one class of object (chairs (Subramaniam and Biederman 1997)), and of artificial "blobs" (asymmetrical harmonics of a sphere with smooth surfaces (Nederhouser et al 2003)). Material perception, however, is disrupted by negation (Fleming et al 2004) (Fleming et al 2001). In the negated images of Figure 6, as an example, we can see what kinds of objects are represented, but it is very difficult to recognize the materials of which they are composed. This difference in susceptibility to negation is likely due to the different kinds of information available for recognition of objects and materials. There is agreement that shape is the primary cue for recognition of non-face objects—what are generally referred to with 'count nouns', such as horse, house, and hammer—and that pigmentation usually plays a secondary role(Biederman and Ju 1988; Tanaka et al 2001). In contrast, materials—what are commonly referred to with 'mass nouns', such as wax, wood, and water—are recognized almost entirely on the basis of their pigmentation (surface reflectance properties) (Biederman 1987). This bolsters the argument that negation disrupts the use of pigmentation cues, and suggests that the phenomenon may not be specific to faces, but rather a general property of recognition. Although face recognition is typically compared with object recognition, these findings suggest that another fruitful comparison would be to material perception, which also depends critically on surface reflectance properties.

For humans, the task of recognizing faces is both difficult and extremely important. It would not make sense for our recognition system to throw away useful information. We have presented evidence from contrast negation that pigmentation is an important part of face identity. This finding cautions against the use of stimuli for recognition experiments that do not retain pigmentation cues, and is not consistent with models of face recognition that throw away pigmentation cues to rely entirely on facial-metric cues such as the shape, size, or distance between features.

Acknowledgments. We thank P. Bronstad, J. Davenport, K. Nakayama, and M. Potter for comments on earlier versions of the manuscript, and the members of the Sinha and Biederman labs for their helpful comments and technical assistance. This research was supported by a grant to P.S. from the DARPA Human ID Program administered by the Information Awareness Office, and by grants ARO DAAG55-98-1-0293, NSF EEC-9529152, HFSP 99-44, JSMF No. 96-44 to I.B.. R.R. was supported by an NIH training grant.

References

Biederman I, 1987 "Recognition-by-components: a theory of human image understanding" *Psychological Review* **94** 115-147

Biederman I, Ju G, 1988 "Surface versus Edge-Based Determinants of Visual Recognition" *Cognitive Psychology* **20** 38-64

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

Braje W L, Kersten D, Tarr M J, Troje N F, 1998 "Illumination Effects in Face Recognition" *Psychobiology* **26** 371-380

Bruce V, Langton S, 1994 "The use of pigmentation and shading information in recognizing the sex and identities of faces" *Perception* **23** 803-22

Bruce V, Young A, 1998 *In the Eye of the Beholder: The science of face perception* (Oxford: Oxford University Press)

Diamond R, Carey S, 1986 "Why Faces Are and Are Not Special: An Effect of Expertise" *Journal of Experimental Psychology: General* **115** 107-117

Fleming R W, Dror R O, Adelson E H, 2001 "How do Humans Determine Reflectance Properties under Unknown Illumination?", in *CVPR* (Kauai, Hawaii: Kauai, Hawaii)

Fleming R W, Jensen H W, Bulthoff H H, 2004 "Perceiving translucent materials", in *ACM SIGGRAPH* (Los Angeles, CA: Los Angeles, CA)

Freire A, Lee K, Symons L A, 2000 "The face-inversion effect as a deficit in the encoding of configural information" *Perception* **29** 159-170

Galper R E, 1970 "Recognition of faces in photographic negative" *Psychonomic Science* **19** 207-208

Galper R E, Hochberg J, 1971 "Recognition memory for photographs of faces" *American Journal of Psychology* **84** 351-354

George N, Dolan R J, Fink G R, Baylis G C, Russell C, Driver J, 1999 "Contrast polarity and face recognition in the human fusiform gyrus" *Nature Neuroscience* **2** 574-580

Hancock P J B, Bruce V, Burton A M, 2000 "Recognition of unfamiliar faces" *Trends in Cognitive Sciences* **4** 330-337

Hayes T, Morrone M C, Burr D C, 1986 "Recognition of positive and negative bandpass-filtered images" *Perception* **15** 595-602

Hill H, Bruce V, 1996 "Effects of lighting on the perception of facial surfaces" *Journal of Experimental Psychology: Human Perception and Performance* **22** 986-1004

Hill H, Bruce V, Akamatsu S, 1995 "Perceiving the sex and race of faces: the role of shape and colour" *Proceedings of the Royal Society of London B* **261** 367-73

Hole G, George P A, Dunsmore V, 1999 "Evidence for holistic processing of faces viewed as photographic negatives" *Perception* **28** 341-359

Horn B K H, 1986 *Robot Vision* (Cambridge, MA: MIT Press)

Johnston A, Hill H, Carman N, 1992 "Recognising faces: effects of lighting direction, inversion, and brightness reversal" *Perception* **21** 365-375

Johnston A, Passmore P J, 1994 "Shape from shading I: Surface curvature and orientation" *Perception* **23** 169-189

Jones B C, Little A C, Burt D M, Perret D I, 2004 "When facial attractiveness is only skin deep" *Perception* **33** 569-576

Kemp R, McManus C, Pigott T, 1990 "Sensitivity to the displacement of facial features in negative and inverted images" *Perception* **19** 531-543

Kemp R, Pike G, White P, Musselman A, 1996 "Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues" *Perception* **25** 37-52

Leder H, Bruce V, 2000 "When inverted faces are recognized: the role of configural information in face recognition" *Quarterly Journal of Experimental Psychology Hum. Exp. Psychol* **53A**

Leder H, Candrian G, Huber O, Bruce V, 2001 "Configural features in the context of upright and inverted faces" *Perception* **30** 73-83

LeGrand R, Mondloch C J, Maurer D, Brent H P, 2001 "Early visual experience and face processing" *Nature* **410** 890

Lewis M B, Johnston R A, 1997 "The Thatcher Illusion as a test of configural disruption" *Perception* **26** 225-227

Liu C H, Chaudhuri A, 1997 "Face recognition with multi-tone and two-tone photographic negatives" *Perception* **26** 1289-1296

Liu C H, Collin C A, Burton A M, Chaurdhuri A, 1999 "Lighting direction affects recognition of untextured faces in photographic positive and negative" *Vision Research* **39** 4003-4009

Liu C H, Collin C A, Chaudhuri A, 2000 "Does face recognition rely on encoding of 3-D surface? Examining the role of shape-from-shading and shape-from-stereo" *Perception* **29** 729-743

Moses Y, Adini Y, Ullman S, 1994 "Face recognition: the problem of compensating for illumination changes" *Proceedings of the European Conference on Computer Vision* 286-296

Nederhouser M, Mangini M C, Biederman I, Okada K, 2003 "Invariance to contrast inversion when matching objects with face-like surface structure and pigmentation" *Journal of Vision* **3** 93

O'Toole A J, Vetter T, Blanz V, 1999 "Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: an application of three-dimensional morphing" *Vision Research* **39** 3145-3155

Pelli D G, 1997 "The VideoToolbox software for visual psychophysics: Transforming numbers into movies" *Spatial Vision* **10** 437-442

Phillips R J, 1972 "Why are faces hard to recognize in photographic negative" *Perception and Psychophysics* **12** 425-426

Ramachandran V S, 1988 "Perception of Shape from Shading" *Nature* **331** 133-166

Russell R, 2003 "Sex, beauty, and the relative luminance of facial features" *Perception* **32** 1093-1107

Sinha P, 2002 "Qualitative representations for recognition" *Lecture Notes in Computer Science* **LNCS 2525** 249-262

Sinha P, Gilad S, Under review "Face Recognition with Contrast Chimeras" *Nature Neuroscience*

Subramaniam S, Biederman I, 1997 "Does contrast reversal affect object identification" *Investigative Opthalmology & Visual Science* **38** 998

Tanaka J, Weiskopf D, Williams P, 2001 "The role of color in high-level vision" *Trends in Cognitive Sciences* **5** 211-215

Tarr M J, Kersten D, Cheng Y, Rossion B, 2001 "It's Pat! Sexing Faces Using Only Red and Green" *Journal of Vision* **1** 337

Watt R J, 1994 "A computational examination of image segmentation and the initial stages of human vision" *Perception* **23** 383-98

White M, 2001 “Effect of photographic negation on matching the expressions and identities of faces” *Perception* **30** 969-981

Chapter 3: The role of shape and pigmentation cues in the recognition of upright and inverted faces

Abstract

The variation among faces can be partitioned into two general sources: a) shape and b) surface reflectance, here termed "pigmentation". It has been implicitly assumed that pigmentation is less important than shape, though there is little empirical evidence to support that notion. To compare the utility of shape or pigmentation for face recognition, we created two sets of faces, wherein individual exemplars differed only by shape in one set and only by pigmentation in the other set. Matching performance was about equal whether pigmentation or shape cues were available for the task. This provides evidence that pigmentation cues, as well as shape cues, play an important role in face recognition.

Introduction

From a cognitive perspective, a central goal for the study of recognition is determining the nature of the internal representations that allow a match between percept and memory. In face recognition, the specific goal is to understand which visual features are used to recognize faces. To better understand which visual cues are important for face recognition we take the approach here of dividing all possibly useful cues into two broad classes, 'shape' and 'pigmentation', and then asking whether one or the other class is more important for recognition.

Visual perception consists of the perception of surfaces, and together, 'shape' and 'pigmentation' cues capture all the visual information that is available about these surfaces. Shape refers to the relative locations of parts of the surface in space, and Pigmentation refers to the reflectance properties of the surface. While surface reflectance properties are sometimes referred to as 'color' or 'texture', these terms are problematic because they invite the misperception that it is specifically hue or variegated reflectance that is being referred to, rather than all surface reflectance properties. We chose the term 'pigmentation' because it avoids that problem, and has the advantage of being intuitive in the context of faces. Cues to shape include boundaries, junctions, and intersections, as well any other cue that gives information about the locations of parts of the surface in space, such as stereo disparity, shape from shading, and motion parallax. Cues to pigmentation include albedo, hue, texture, translucence, specularity, and any other property of surface reflectance. 'Second order relations', or the distances between facial features such as the eyes and mouth, are a subclass of shape. However, the relative reflectance of those features or parts of features, such as the relative darkness of the entire eye region and the cheek, are a subclass of pigmentation. Some other cues, most notably illumination, are not captured by this classification. The direction of lighting has been found to play a role in recognition (particularly unfamiliar recognition), with lighting from below causing particular difficulties for recognition (Braje et al 1998; Enns and Shore 1997; Hill and Bruce 1996; Johnston et al 1992; McMullen et al 2000), but we will not consider this further here.

This division of cues for recognition into two classes is admittedly not perfect. For example, texture is an attribute of pigmentation, but combined with motion, can be a clue to depth and hence, shape. More relevant for faces, however, is the question of whether a border defined by a luminance gradient, such as the joining of the iris and sclera should be classified as a shape or pigmentation cue. Because faces share a common configuration of parts, we classify such borders as shape cues when they are common to all faces (e.g. the iris-sclera border), but as pigmentation cues when they are unique to an individual (e.g. a mole or freckles). Though the classification cannot completely separate shape and pigmentation cues, it does separate the vast majority of cues. There are many such distinctions that are even less complete, but nonetheless useful, such as describing the cortical flow of visual information as consisting of two streams. We believe that dividing cues for recognition into shape and pigmentation is in fact mostly complete, and a useful distinction to draw.

Indeed, this very same division has been used to investigate recognition of non-face objects. Much of this work has compared human ability to recognize, at the basic level, objects from line drawings or continuous tone images, such as photographs. The assumption here is that line drawings contain shape cues, but not pigmentation cues, and hence ability to recognize an object from a line drawing indicates reliance on shape cues. In particular, these studies have found recognition of line drawings to be as good ((Biederman and Ju 1988; Davidoff and Ostergaard 1988; Ostergaard and Davidoff 1985)) or almost as good ((Humphrey 1994; Price and Humphreys 1989; Wurm et al 1993)) as recognition of photographs. On the basis of these and similar studies, there is a consensus that, in most cases, pigmentation is less important than shape for object recognition (Palmer 1999; Tanaka et al 2001; Ullman 1996).

In the face recognition literature there is not an explicit assumption that shape is more important than pigmentation, as there is in the object recognition literature. However, three lines of evidence suggest that there is an implicit assumption that shape is more important. The first is that line drawings and other non-pigmented representations are commonly used as stimuli for face recognition experiments. It would not make sense to stimuli that throw out an entire class of cues unless you believed that those cues were relatively unimportant. Similarly, many models of face recognition use only shape cues, such as distances between features, as the featural inputs. As with using unpigmented stimuli, this only makes sense if it is assumed that the pigmentation cues being ignored are not as important as the shape cues being retained. Finally, studies investigating aspects of shape are far more common than studies of pigmentation. Together, these practices suggest that the face research community does not believe that pigmentation is a particularly relevant component of face representations.

Yet, there are reasons to believe that pigmentation may indeed be important for face recognition. Unlike for objects, it is much more difficult to recognize a face from a line drawing than from a photograph (Bruce et al 1992; Davies et al 1978; Leder 1999; Rhodes et al 1987), suggesting that the pigmentation cues thrown away by such a representation are indeed important. Two seminal studies (Bruce and Langton 1994; O'Toole et al 1999) using uniformly pigmented models of faces (derived from laser scans) found evidence that pigmentation cues play an important role in recognition. Recognizing faces is immensely important, yet, in addition to the many problems of visual recognition generally, face recognition is particularly difficult because the distinguishing features of faces are relatively subtle and few. Because face recognition is both important and difficult, it is likely that any cues that could be used to distinguish faces can potentially be used by people to recognize faces. In light of these reasons, we believe that it is not justifiable to assume without investigation that pigmentation is unimportant for face recognition. We present here such an investigation.

Our approach is to force subjects to use only shape or only pigmentation cues when recognizing faces. Under these circumstances, we can ask whether the observers exhibit differential abilities to make use of the two kinds of cues. Toward this end, we created sets of faces in which the exemplars differ from one another in terms of only their shape or only their pigmentation. The Shape set consists of faces with the same pigmentation, but different shape from one another, while the Pigmentation set consists of faces with the same shape as each other, but different pigmentation. The faces were created from photographs of real people, such that the variation between faces in, for example the Shape set, is proportional to the shape variation between the original faces. We used a a delayed match-to-sample task in which both the sample and the distractor are drawn from the same set. Thus, when the faces are drawn from the Shape set, the subjects must use shape cues to perform the task, and when the faces are drawn from the Pigmentation set, the subjects must use pigmentation cues. We then compared performance between the Shape and Pigmentation sets, to determine whether using only one or the other kind of cue results in better recognition.

In comparing the utility of shape and pigmentation cues, it is important to know whether better performance is caused by one cue intrinsically containing more information than the other, or by the visual system processing those cues more efficiently. We can address that question with the faces we created for this study by equating the similarity of the Shape and Pigmentation sets. To determine the similarity of these faces, we selected the similarity metric developed by Christoph von der Malsburg and colleagues that is loosely based on the response properties of neurons in area V1 (Lades et al 1993). This metric was chosen because of it has been found to do a good job of predicting human similarity judgments of faces (Biederman and Kalocsai 1997; Hancock et al 1998; Kalocsai et al 1994). We will refer here to similarity values derived using this metric as "Malsburg similarity". With the Malsburg similarity ratings and the behavioral performance, we will have both a measure of the stimulus energy available for discrimination, and the ability of humans to use these cues for recognition. Another benefit of equating the Shape and Pigmentation sets for Malsburg similarity is that it is a way of controlling some of the arbitrariness that is inherent in selecting a particular set of faces as stimuli to investigate the recognition of faces in general. Though the collection of faces is necessarily arbitrary, the matched similarity of the sets is not.

The literature on object recognition has often referred to pigmentation as 'color', and this assumption that pigmentation consists of primarily hue information has penetrated the face recognition literature (Kemp et al 1996). To evaluate whether pigmentation cues exist only or even primarily in hue channels, we used grayscale and full color versions of all of the stimuli. If performance in the Pigmentation condition is better than chance with grayscale faces, it will suggest that there are significant components of pigmentation other than hue, which would imply that hue should not be confounded with pigmentation more generally. Using both color and grayscale stimuli also allows us to investigate whether hue plays an important role in normal (upright) face recognition, a question that has remained unresolved (Bruce and Young 1998). If color does play a role in face recognition, we should find better performance with full color stimuli than grayscale stimuli, though not in the Shape condition, because even when those faces are colored, they do not differ in terms of color.

Using sets of faces that differ from one another in either shape or pigmentation, we can also evaluate whether shape or pigmentation cues are selectively disrupted by inversion (turning a face upside-down). The decline in recognition performance that occurs with face inversion is typically attributed to impairment in perceiving second-order relations (the distances between named features, such as the eyes and mouth). Consistent with this hypothesis, larger inversion effects have been found in face matching tasks when the faces differed in terms of second-order relations than in terms of 'features' (the color or shape of the eyes and mouth) (Barton et al 2001; Freire et al 2000; Leder and Bruce 2000; LeGrand et al 2001). Because second-order relations are a subclass of shape cue, we can use the present sets of faces to evaluate this claim. If second-order relations drive the inversion effect, we should find a greater decline in performance with inverted faces in the shape set (which differ from one another in terms of their second-order relations) than the pigmentation set (which do NOT differ from one another in terms of their second-order relations). We can also test the stronger hypothesis that shape cues, but not pigmentation cues are disrupted by inversion. This hypothesis will be supported if we find that performance with the shape set, but not with the pigmentation set, is impaired by inversion.

For the current study we used carefully matched, naturalistic stimuli to determine whether shape or pigmentation is more important for face recognition. We also addressed whether the perception of shape or pigmentation cues is selectively disrupted by face inversion, and whether color (hue) cues are used in face recognition.

Methods

We ran two experiments using the same experimental design. In Experiment 1 stimuli were presented in grayscale, and in Experiment 2 stimuli were presented in full color.

Subjects

Thirty subjects participated in the experiments—fifteen in Experiment 1, and fifteen in experiment 2. All were contacted through the MIT Brain & Cognitive Sciences subject pool, were naïve to the purpose of the study, and had normal or corrected-to-normal vision.

Stimuli

Full frontal photographs were taken of eight male and eight female faces. The people photographed were all Caucasian, and ranged in age from 18 to 25, with a mean age of 20. All faces were photographed under the same illumination, in the same room, in which all light came from two studio lamps with diffusing heads. These two lamps, the camera and tripod, and the chair on which the photographic subjects sat, were kept in locations that were fixed with respect to each other and the room. Similarly, the heights of the camera and lamps were fixed, and the chair used by the photographic subjects was adjusted such that their heads were all placed at the same height. The lights were centered at 0° elevation, to eliminate cast shadows and minimize the effects of shading cues (Liu et al 1999).

These photographs were then manipulated using Morph Man 3.0 (Stoik Imaging). For each sex, the eight original faces were morphed together to produce an average face. We created the stimuli for the Shape condition by distorting this average face into the shape of each of the original eight faces, producing eight new faces with the same (average) pigmentation, but distinct shape. Similarly, we created the stimuli for the Pigmentation condition by distorting each of the eight original faces into the shape of the average face, which produced eight faces with the same (average) shape, but distinct pigmentation. The original eight faces, which differed from one another in terms of both their shape and their pigmentation, were the stimuli for the Shape + Pigmentation condition. Further description of these stimuli can be found in Russell et al. (In press). In Experiment 1, these stimuli were presented in grayscale, and in Experiment 2 the same stimuli were presented in full color (Figure 1).

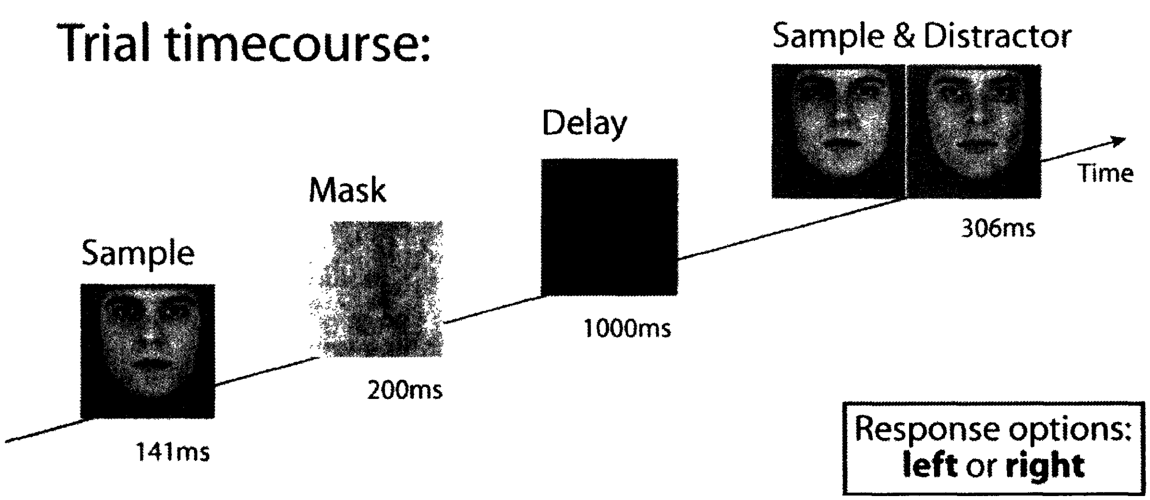

Figure 1. Experimental design. On each trial, subjects viewed a face, followed by a visual noise mask, then a delay. Two faces were then presented; one face (the target) was the identical image as the sample, and the other face (the distractor) was drawn from the same category of faces. For example, if the face in the upper left of Figure 2 was the target, the face to the right could be the distractor, but not one of the faces below. In this way, the information distinguishing the target and distractor faces defined the condition: Shape only, Pigmentation only, and Shape + Pigmentation.

The grayscale versions of the Shape and Pigmenation sets were matched with one another for Malsburg similarity. Specifically, the distributions of similarities between faces within the Shape and Pigmentation sets had the same mean and median. We did this by carefully choosing the gray level of the background. Stronger luminance contrast between the background and the faces increased similarity among the Pigmentation faces and decreased similarity among the Shape faces. This is because stronger contrast decreases perception of subtle differences between the faces in overall pigmentation, but increases perception of the shape of the outline of the face. The actual faces were not altered to match the sets for Malsburg similiarity. Images in the Shape + Pigmentation set were less similar to one another than were the images in the other two sets, because they differed from one another along more dimensions. For this reason we expected that performance with these images would be as good or better than performance with the Shape or Pigmentation sets. The Malsburg system is designed to operate on grayscale (single channel) stimuli, and so the color versions of the stimuli could not be matched for similarity.

Procedure

We used a delayed match-to-sample, two-alternative forced-choice task (Figure 2). In each trial of this task, the observer saw a fixation dot followed by a sample face for 141msec, then a visual noise mask for 200msec, then and a blank screen for 1000msec. Next, two faces were presented side by side in the center of the screen for 306msec. One of the two faces, the sample, had been presented just previously, and the other was a distracter face that differed from the target in shape only (Shape condition), in pigmentation only (Pigmentation condition), or in both shape and pigmentation (Shape + Pigmentation condition). The task was to decide which of the two faces matched the sample, and to press the corresponding key as quickly as possible. Trials from the different conditions were randomized, and the left-right ordering of target and distracter was counterbalanced. There were eight male and eight female faces in each condition, and each of the eight faces was paired with every other face of its same sex for each condition. All the faces in a trial were presented upright or inverted. Stimuli were presented in Matlab (The MathWorks, Inc.) using the Psychophysics Toolbox extensions (Brainard 1997; Pelli 1997).

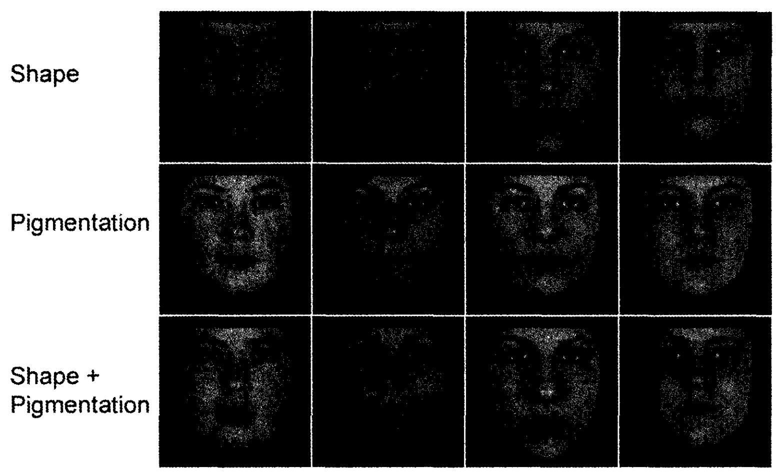

Figure 2. Some examples of stimuli. The Shape condition faces along the top row have the same pigmentation as one another and are distinguishable only by shape. The Pigmentation condition faces along the middle row have the same shape and are distinguishable only by pigmentation. The Shape + Pigmentation faces along the bottom row are photographs of actual people, and are distinguishable by both shape and pigmentation.

Results

Experiment 1—grayscale images matched for Malsburg similarity

Percent correct responses are shown in Figure 3. Performance was best in the Shape + Pigmentation condition, better in the Shape condition, and worst in the Pigmentation condition. There was a main effect of condition $F(2,28) = 41.4$, $p < 0.001$ with the best performance when both shape and pigmentation cues were available. There was also a main effect of orientation $F(1,14) = 48.4$, $p < 0.001$ with performance worse with inverted than upright faces, consistent with all other studies on the effect of orientation. There was not a significant interaction between condition and orientation, $F(2,28) = 0.3$, $p > 0.05$, indicating that orientation did not disrupt one cue more than the other. Subsequent Tukey's HSD post-hoc comparisons indicated that performance in each of the conditions was significantly different at the 0.05 level. Thus performance in the Shape condition was significantly better than performance in the Pigmentation condition.

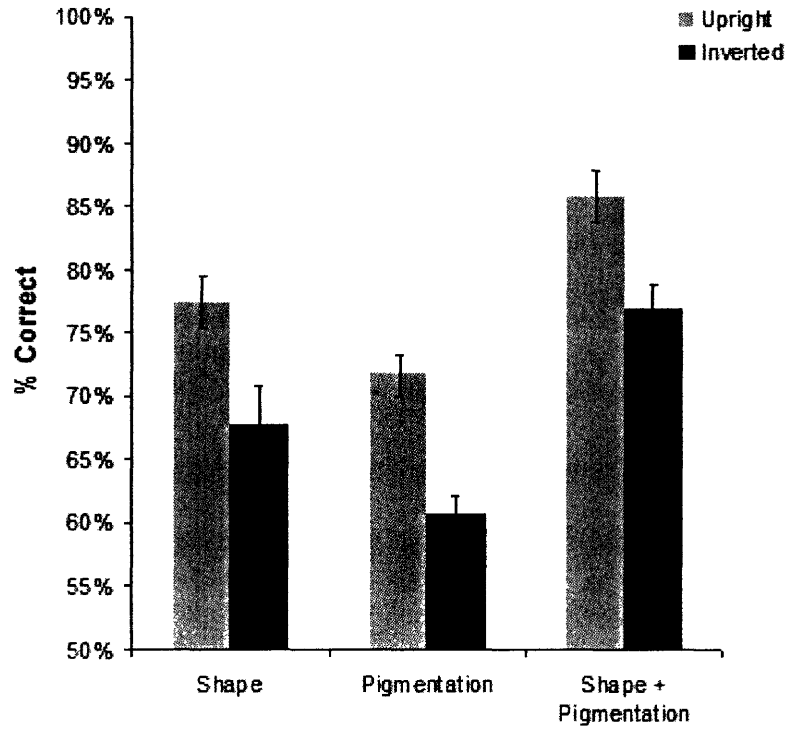

Figure 3. Results of Experiment 1, with grayscale stimuli, expressed as percentage of correct responses. Error bars indicate standard errors.

3.2 *Experiment 2—full color images*

The pattern of results in Experiment 2 differed somewhat from that of Experiment 1 (Figure 4), with better performance in the Pigmentation condition than in the Shape condition, and smaller declines in performance with inverted faces. As in Experiment 1, there was a main effect of condition $F(2,28) = 34.7$, $p < 0.001$ with the best performance when both shape and pigmentation cues were available, and a main effect of orientation $F(1,14) = 15.7$, $p < 0.005$, with performance worse with inverted than upright faces. Again, there was not a significant interaction between condition and orientation, $F(2,28) = 0.02$, $p > 0.05$, indicating that orientation did not disrupt one cue more than the other. Subsequent Tukey's HSD post-hoc comparisons indicated that performance in each of the conditions was significantly different at the 0.05 level. Thus performance in the

Pigmentation condition was significantly better than performance in the Shape condition, the opposite relationship from that found in Experiment 1.

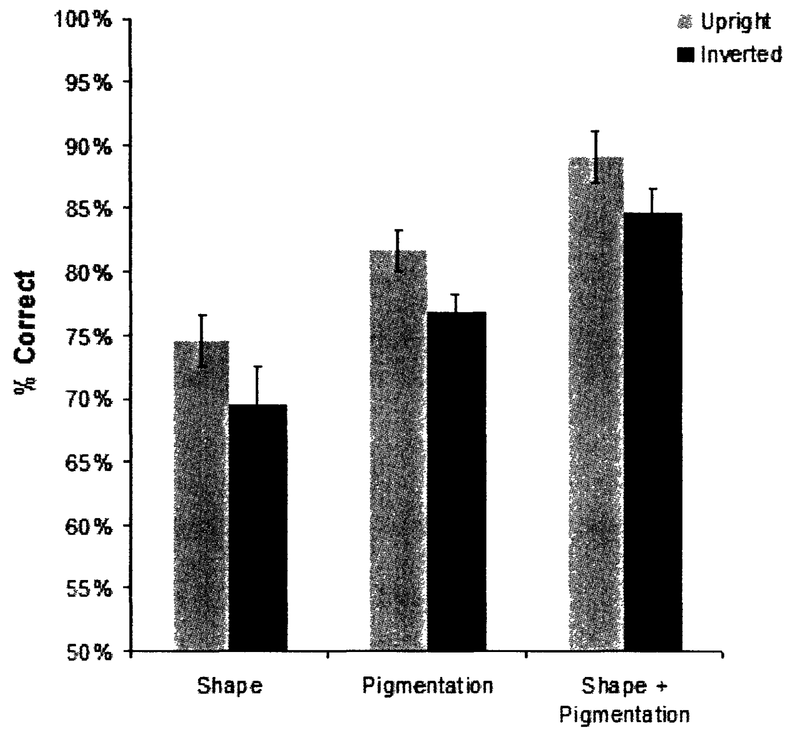

Figure 4. Results of Experiment 2, with full color stimuli, expressed as percentage of correct responses. Error bars indicate standard errors.

To evaluate the utility of color cues on normal (upright) face recognition, we ran an ANOVA with upright faces only, on the data from both experiments with color (full color or grayscale experiment) treated as a between subjects factor. Figure 5 presents the results with upright faces across the three conditions for full color (from Experiment 1) and grayscale faces (from Experiment 2). There were not significant main effects of color $F(1,2.5) = 0.8, p > 0.05$ or condition $F(1,1.6) = 2.7, p > 0.05$, but there was a significant interaction between color and condition $F(13,26) = 6.3, p < 0.01$. T-tests comparing performance with grayscale and full color faces in each of the conditions found significant difference in the Pigmentation condition $t(58) = 4.9$, $p < 0.001$, but not in the Shape $t(58) = 0.3$, $p > 0.05$ or Shape + Pigmentation $t(58) = 1.6$, $p > 0.05$ conditions, confirming that the interaction between color and condition was driven by the Pigmentation condition, in which performance was much better with full color than with grayscale.

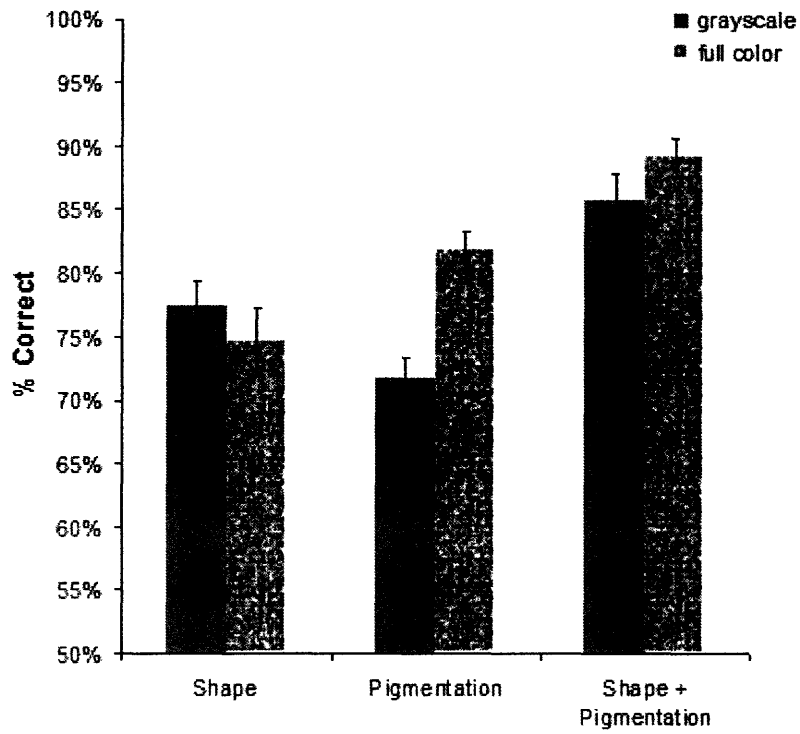

Figure 5. Results with upright faces only, of Experiments 1 (grayscale stimuli) and 2 (full color stimuli).

Inversion had a much smaller effect on performance with color faces than with grayscale ones. This is a surprising finding, as most studies of face inversion have used color images, and they consistently find large effects of inversion. Investigating this further, we found that, while there was no main effect of the sex of the face being matched $F(1,14) = 0.3$, $p > 0.05$), there was a significant interaction between orientation and the sex of the face $F(1,14) = 5.3, p < 0.05$). Figure 6 shows the results of Experiment 2 with performance at matching male and female separated. While there is a sizeable cost of inversion on the female faces, there is almost no effect of inversion on the male faces. Thus, the smaller effect of inversion on performance in Experiment 2 is due to the negligible effect of inversion on matching the male faces. There was no such interaction between orientation and the sex of the face in Experiment 1, with grayscale faces.

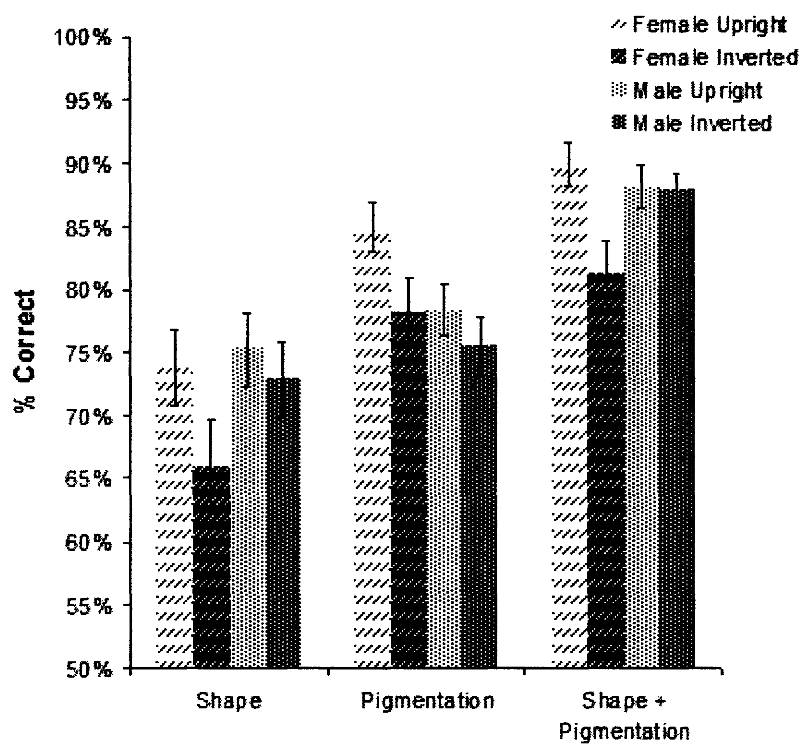

Figure 6. Results of Experiment 2, with responses to male and female versions of the stimuli presented separately. Inversion did not have a significant effect on the male faces.

Discussion

Three findings were in evidence in our experiments: 1) Subjects could use pigmentation cues alone to match faces, performing significantly better than chance, and indeed at about the same level as with shape cues alone. 2) Performance was equally impaired by inversion with faces differing in terms of second-order relations and with faces not differing in terms of second-order relations. 3) In the Pigmentation condition, when shape cues were not available, performance was better when subjects could make use of color as well as grayscale cues.

These data provide direct evidence that pigmentation is an important cue for face recognition. First, performance using pigmentation cues alone was significantly better than chance, and second, performance using shape cues alone was significantly worse than performance using both shape and pigmentation cues. The evidence is less clear for whether shape or pigmentation cues are more useful for recognition. With grayscale images that were matched for Malsburg similarity, performance with shape cues was slightly but significantly better than performance with pigmentation cues. However, when the same faces were presented in full color (and not matched for similarity), we found better performance using pigmentation cues. A natural temptation would be to assume that the results with color images are more relevant because faces are normally recognized in color, and presenting the faces in grayscale selectively removes pigmentation cues but not shape cues. There is reason however, to question this assumption. Faces appear in a variety of contexts, under different lighting, and often behind occlusions such as hair or clothing. Any of these factors can selectively impair specific cues. In the case of our stimuli, they are all viewed frontally, with frontal lighting, with the same cropping and the same artificial background. Any of these choices could play a role in making shape or pigmentation more useful. It is also possible that the sixteen Caucasian faces we used as stimuli are unusually similar to one another in terms of either their shape or pigmentation. Setting the grayscale Shape and Pigmentation sets to be equal in similarity is a means of controlling for this arbitrariness, and this might suggest that the grayscale results are more relevant. Because of these issues, we do not believe that the present data are sufficient for a quantitative comparison of the relative utility of shape and pigmentation cues for face recognition under general conditions. However, these results do offer evidence for a more qualitative argument that both pigmentation and shape are important cues for face recognition.

Overall, this finding demonstrates that pigmentation is an important cue for face recognition. This is consistent with the findings from another study using a very similar method with images of laser-scanned faces that partly inspired our current work (O'Toole et al 1999), and with two studies finding that the impairment of recognition by contrast negation is at least partly caused by the disruption of pigmentation cues (Bruce and Langton 1994; Russell et al (In Press)). The finding that pigmentation is important for recognition places key constraints on models of face recognition, and on the kinds of stimuli that are suitable for experiments investigating face recognition. A possible reason why the study of face recognition has been biased toward shape cues is that we have better metrics for shape than pigmentation cues. Pigmentation is typically conceived as a one dimensional signal, with an entire face being darker or lighter. To better understand how pigmentation cues are used to perform face recognition, appropriate metrics for characterizing facial pigmentation as a two dimensional signal must be developed.

These results also bear on the question of whether inversion selectively impairs the use of second-order relations in face recognition. In both experiments, we found that the use of shape and pigmentation cues was equally impaired by inverting the faces. Because the faces in the Pigmentation condition did not differ in terms of their second-order relations, this provides evidence against the assertion that it is specifically the impairment in the perception of second-order relations that causes poorer recognition with inverted than upright faces. This joins two other recent studies in finding that performance with configural cues does not decline more with inversion than does performance with featural cues (Riesenhuber et al 2004; Yovel and Kanwisher 2004). The current study extends these findings with the discovery that not only are featural and relational cues not differently disrupted by inversion, neither are the much broader categories of shape and pigmentation cues. Why there was no inversion effect for male faces when presented in color remains unclear. Future studies will be required to determine whether this finding is robust, or a result our stimuli.

Performance with grayscale faces that differed only in terms of pigmentation was significantly above chance. This provides direct evidence that color (hue) is not the only pigmentation cue useful for recognition. Yet performance with these same faces was significantly better when color cues were available. However, performance with normal (Shape + Pigmentation) faces was not significantly better with color cues than without, though there was a trend in that direction. This is consistent with another study from our laboratories finding that color celebrity faces were not recognized better in full color than in grayscale when the images were not degraded. However, when the images were blurred, performance was better with full color (Yip and Sinha 2002). Together, the two studies suggest that color cues are useful for face recognition under conditions in which some other cues are degraded—shape cues in the case of the present study or high frequency spatial information in the case of the 2002 study. Several other researchers, using different methods, have also reported small but significant improvements to recognition performance from color cues (Kemp et al 1996; Lee and Perrett 1997; Lee and Perrett 2000).

We have shown that pigmentation cues are about as useful as shape cues for face recognition, discounting the notion that pigmentation is unimportant for face recognition. This finding places an important constraint on theories of face recognition, and argues against the use of stimuli without pigmentation cues in experiments investigating face recognition.

References

Barton J J S, Keenan J P, Bass T, 2001 "Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration" *British Journal of Psychology* **92** 527-549

Biederman I, Ju G, 1988 "Surface versus Edge-Based Determinants of Visual Recognition" *Cognitive Psychology* **20** 38-64

Biederman I, Kalocsai P, 1997 "Neurocomputational bases of object and face recognition" *Philosophical Transactions of the Royal Society of London B* **352** 1203-1219

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

Braje W L, Kersten D, Tarr M J, Troje N F, 1998 "Illumination Effects in Face Recognition" *Psychobiology* **26** 371-380

Bruce V, Hanna E, Dench N, Healey P, Burton M, 1992 "The importance of 'mass' in line drawings of faces" *Applied Cognitive Psychology* **6** 619-628

Bruce V, Langton S, 1994 "The use of pigmentation and shading information in recognizing the sex and identities of faces" *Perception* **23** 803-22

Bruce V, Young A, 1998 *In the Eye of the Beholder: The science of face perception* (Oxford: Oxford University Press)

Davidoff J, Ostergaard A, 1988 "The role of color in categorical judgments" *Quarterly Journal of Experimental Psychology* **40** 533-544

Davies G M, Ellis H D, Sheperd J W, 1978 "Face recognition accuracy as a function of mode of representation" *Journal of Applied Psychology* **63** 180-187

Enns J T, Shore D I, 1997 "Separate influences of orientation and lighting in the inverted-face effect" *Perception & Psychophysics* **59** 23-31

Freire A, Lee K, Symons L A, 2000 "The face-inversion effect as a deficit in the encoding of configural information" *Perception* **29** 159-170

Hancock P J B, Bruce V, Burton M A, 1998 "A comparison of two computer-based face identification systems with human perceptions of faces" *Vision Research* **38** 2277-2288

Hill H, Bruce V, 1996 "Effects of lighting on the perception of facial surfaces" *Journal of Experimental Psychology: Human Perception and Performance* **22** 986-1004

Humphrey G, 1994 "The role of surface information in object recognition: studies of visual form agnosic and normal subjects" *Perception* **23** 1457-1481

Johnston A, Hill H, Carman N, 1992 "Recognising faces: effects of lighting direction, inversion, and brightness reversal" *Perception* **21** 365-375

Kalocsai P, Biederman I, Cooper E E, 1994 "To what extent can the recognition of unfamiliar faces be accounted for by a representation of the direct output of simple cells" *Investigative Opthalmol. Visual Sci.* **35** 1626

Kemp R, Pike G, White P, Musselman A, 1996 "Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues" *Perception* **25** 37-52

Lades M, Vortbruggen J C, Buhmann J, Lange J, von der Malsburg C, Wurtz R P, Konen W, 1993 "Distortion invariant object recognition in the dynamic link architecture" *IEEE Transactions on Computers* **42** 300-311

Leder H, 1999 "Matching person identity from facial line drawings" *Perception* **28** 1171-1175

Leder H, Bruce V, 2000 "When inverted faces are recognized: the role of configural information in face recognition" *Quarterly Journal of Experimental Psychology Hum. Exp. Psychol* **53A**

Lee K J, Perrett D, 1997 "Presentation-time measures of the effects of manipulations in colour space on discrimination of famous faces" *Perception* **26** 733-752

Lee K J, Perrett D I, 2000 "Manipulation of colour and shape information and its consequences upon recognition and best-likeness judgments" *Perception* **29** 1291-1312

LeGrand R, Mondloch C J, Maurer D, Brent H P, 2001 "Early visual experience and face processing" *Nature* **410** 890

Liu C H, Collin C A, Burton A M, Chaurdhuri A, 1999 "Lighting direction affects recognition of untextured faces in photographic positive and negative" *Vision Research* **39** 4003-4009

McMullen P A, Shore D I, Henderson R B, 2000 "Testing a two-component model of face identification: Effects of inversion, contrast reversal, and direction of lighting" *Perception* **29** 609-619

Ostergaard A, Davidoff J, 1985 "Some effects of color on naming and recognition of objects" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **11** 579-587

O'Toole A J, Vetter T, Blanz V, 1999 "Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: an application of three-dimensional morphing" *Vision Research* **39** 3145-3155

Palmer S E, 1999 *Vision Science: Photons to Phenomenology* (Cambridge, Massachusetts: MIT Press)

Pelli D G, 1997 "The VideoToolbox software for visual psychophysics: Transforming numbers into movies" *Spatial Vision* **10** 437-442

Price C J, Humphreys G W, 1989 "The effects of surface detail on object categorization and naming" *Quarterly Journal of Experimental Psychology* **41A** 797-828

Rhodes G, Brennan S E, Carey S, 1987 "Recognition and ratings of caricatures: Implications for mental representations of faces" *Cognitive Psychology* **19** 473-497

Riesenhuber M, Jarudi I, Gilad S, Sinha P, 2004 "Face processing in humans is compatible with a simple shape-based model of vision" *Proceedings of the Royal Society of London, B - Biological Sciences* **271** S448-S450

Russell R, Sinha P, Biederman I, Nederhouser M, (In Press) "Is pigmentation important for face recognition? Evidence from contrast negation" *Perception*

Tanaka J, Weiskopf D, Williams P, 2001 "The role of color in high-level vision" *Trends in Cognitive Sciences* **5** 211-215

Ullman S, 1996 *High-Level Vision: Object Recognition and Visual Cognition* (Cambridge, Massachusetts: MIT Press)

Wurm L H, Legge G E, Isenberg L M, Luebker A, 1993 "Color improves object recognition in normal and low vision" *Journal of Experimental Psychology: Human Perception and Performance* **19** 899-911

Yip A, Sinha P, 2002 "Contribution of color to face recognition" *Perception* **31** 995-1005

Yovel G, Kanwisher N, 2004 "Face Perception: Domain Specific, Not Process Specific" *Neuron* **44** 889-898

Chapter 4: Is pigmentation important for familiar face recognition?

Abstract

Recent evidence has suggested that pigmentation cues are about as important as shape cues for face recognition. However, this evidence has come only from experiments using unfamiliar face recognition tasks, such as face matching. Here we performed a familiar recognition experiment, in which subjects were asked to recognize images of the faces of their friends. The images were manipulated such that only pigmentation cues or only shape cues were useful for recognizing any particular face. Relative performance differed by the kind of face being recognized—Asian female and Caucasian male faces were recognized equally well with shape or pigmentation cues, while Caucasian females were recognized significantly better with pigmentation cues than shape cues. Most importantly, performance with either shape or pigmentation cues alone for all face types was far above chance, indicating that each cue alone could be used to perform familiar recognition.

Introduction

As described in the previous two chapters, evidence has accumulated over the past decade to suggest that pigmentation plays an important role in face recognition (Bruce and Langton 1994; O'Toole et al 1999; Russell et al 2004; Russell et al (In Press); Vuong et al 2005). A primary reason for the importance of face recognition as a subject of research is the ecological importance of the task. Recognizing another person allows the perceiver to bring to bear their memory of how that person has acted in the past, including how that person interacted with the perceiver and with others. As highly social animals, this kind of knowledge about others and their interpersonal relations is of critically important for our survival.

But despite the fact that face recognition is considered an important research topic in large part because of this ecological importance, it is typically studied in the laboratory using methods that are far from ecological. In particular, the vast majority of recognition experiments use tasks that can be described as "unfamiliar recognition", in which the subject does not actually know the people whose faces are presented. The studies that provided evidence that pigmentation is important for recognition also used unfamiliar recognition tasks. In contrast, "familiar recognition" tasks involve faces of people who are known to the subjects. Perhaps the canonical example of a familiar recognition task is naming who is represented in a photograph.

How unfamiliar recognition works is, in fact, an important question in it's own right, particularly in applied settings such as the construction of automated systems for face recognition or in forensic settings (Hancock et al 2000). But the majority of face recognition studies use unfamiliar recognition tasks not because of a specific interest in unfamiliar recognition, but because it is much easier for the experimenter to conduct an unfamiliar recognition task than a familiar recognition task. This is because performing a familiar recognition task requires obtaining a set of faces of people who are known to a group of people who can serve as subjects in the experiment. By contrast, it is easy to find a set of photos of people who are *not* known to a group of subjects.

There is reason to feel somewhat confident about the validity of the results from unfamiliar recognition tasks, because there is evidence that both tasks call on similar mechanisms, and both can be impaired by similar manipulations. For example, both familiar and unfamiliar recognition are impaired by vertical inversion or contrast negation of images of faces. Yet there is always the possibility that novel effects found with unfamiliar recognition tasks could be driven by image-level processing rather than face specific processing. If the subjects' task is to match or remember *specific images* (that happen to include a face), the subject is bound to make use of properties of those images. Indeed, the only information that the subject has to perform an unfamiliar recognition task are those images. This differs from familiar recognition, in which the subject has a much broader range of experiences with the actual physical object of the face. In this sense, the two kinds of tasks are quite different, and ultimately, results found using unfamiliar recognition tasks are interesting primarily to the extent to which they predict results from familiar recognition tasks. For this reason, familiar recognition tasks will always remain the "gold standard" for face recognition.

Toward this end, we conducted an experiment investigating the relative utility of shape and pigmentation cues for familiar recognition. Our approach was to create sets of faces, based on photographs of real people, that either had the same shape or the same pigmentation as the original photos. The people photographed were a group of students living together in the same dormitory. They were chosen because they knew each other, and could subsequently recognize photographs of one another. In this way we conducted a familiar recognition experiment in which the subjects were shown faces that had either the shape of their friends' faces or the pigmentation of their friends' faces, and were asked to name the friend whose face had been altered to produce the image on screen. In this way, we assessed the subjects' ability to recognize their friends using only shape cues or only pigmentation cues.

Methods

Subjects

Twenty six subjects participated in the experiment, eighteen females and eight males. The subjects ranged in age from 18 to 34, with an average age of 20. All subjects were among the thirty individuals whose photographs were used as stimuli (see Stimuli, below). Eight of the subjects were Asian females, ten were Caucasian females, and eight were Caucasian males. Of the remaining four individuals who were photographed but did not take the experiment, two Asian females and one Caucasian male took pilot versions of the experiment, and one Caucasian male was unavailable to take the experiment. All of the subjects lived in the same dormitory on the MIT campus, and were selected because of their acquaintance with the people whose faces were used as stimuli.

Stimuli

Full frontal, color photographs were taken of thirty faces, ten male and twenty female. The individuals photographed ranged in age from 18 to 34, with an average age of 20. All ten males were Caucasian, ten of the females were Caucasian, and ten were East Asian. In this way, the total set was divided into three groups of faces: Asian females, Caucasian females, and Caucasian males, each comprising ten faces.

It was important that variation in the apparent darkness of the face be caused by variation in pigmentation and not variation in illumination. For this reason, very careful attention was paid to making certain that all the faces were photographed under the same illumination. Toward this end, all photos were taken in the same room, in which all light came from two studio lamps with diffusing heads. These two lamps, the camera and tripod, and the chair on which the photographic subjects sat, were kept in locations that were fixed with respect to each other and the room. Similarly, the heights of the camera and lamps were fixed. Finally, the subjects sat on a chair, the height of which was adjusted, such that the subjects' heads were all placed at the same height. The lights were centered at 0^{o} elevation, to eliminate cast shadows and minimize the effects of shading cues (Liu et al 1999).

These photographs were then manipulated using Morph Man 3.0 (Stoik Imaging). For each group of faces (e.g. Caucasian females), the ten original faces were morphed together to produce an average face. We created the stimuli for the Shape condition by distorting this average face into the shape of each of the original ten faces. This produced ten new faces that each had the same shape as one of the original faces, but the pigmentation of the average face. Similarly, we created the stimuli for the Pigmentation condition by distorting each of the ten original faces into the shape of the average face. This produced ten new faces that each had the same pigmentation as one of the original faces, but the shape of the average face. Thus there were thirty new faces created for the Shape condition and thirty new faces created for the Pigmentation condition. These stimuli were created using a method described previously (Russell et al 2004; Russell et al (In Press)), but with different faces, and photographed in a different studio.

Examples of the stimuli from the group of Caucasian females are shown in Figure 1. The example stimuli were all selected from the same group in order to illustrate how faces from the same group in the Shape condition all had the same pigmentation, and those in the Pigmentation all had the same shape. Though the Caucasian male stimuli were produced using an image of the average ten faces, only 5 of the Caucasian male faces were actually presented to the subjects. This was due to unintentional experimenter error.

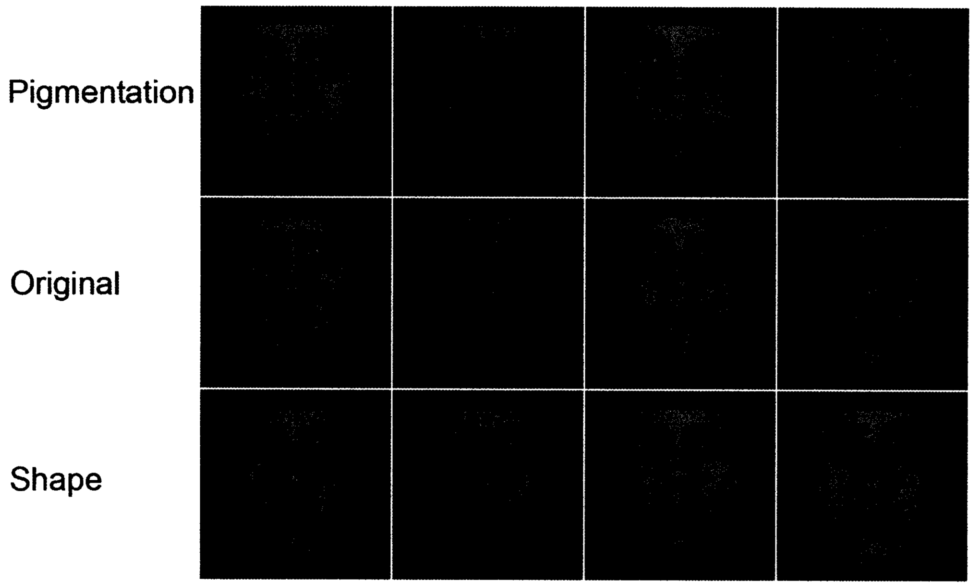

Figure 1. Some examples of Caucasian female stimuli. The Pigmentation condition faces along the top row have the same shape and are distinguishable only by pigmentation. The Shape condition faces along the bottom row have the same pigmentation as one another and are distinguishable only by shape. The faces along the middle row are photographs of actual people, and are distinguishable by both shape and pigmentation. Actual stimuli were in full color.

Procedure

Subjects first viewed a list of the names of the thirty faces that would be presented, and were asked to indicate by checking boxes whether they knew or did not know each individual. Nineteen subjects knew all the people listed. Of the remaining eleven subjects, the vast majority were unfamiliar with only a single individual from the list. Subjects were told that they would be shown images produced by manipulating photographs of those thirty individuals that had just been listed, and that their task with each presented face was to name the individual whose face had been manipulated to produce that image. The subjects were not informed as to how the images had been manipulated.

Each subject was then shown twenty-five faces—ten Asian females, ten Caucasian females, and five Caucasian males. Within each of these groups, half of the images were drawn from the Shape condition and half from the Pigmentation condition, such that every face shown was derived from a different individual. Multiple faces derived from the same individual were not shown in order to avoid priming the identities of the people depicted. There were two different versions of the experiment, with a Pigmentation version of a particular individual's face being shown in one version and a Shape version being shown in the other version. Faces were presented on the screen a laptop computer. For the first five subjects, each face appeared on screen for 10 seconds. For the subsequent twenty one subjects, the presentation time for each face was only 5 seconds. Presentation time was reduced to 5 seconds in the belief that it would provide ample time for a face to be recognized or not, and that it would reduce the possibility of strategic guessing based on knowledge of which faces might appear and which had already appeared. All subjects were given unlimited time to write the name of the individual on a sheet of paper, following presentation of each face.

Following the display of test set of faces, the subjects were shown the twenty five original (unaltered) photographs, and asked to name the individuals depicted. This last part of the experiment served as a baseline. If a subject was unable to recognize the unmanipulated photograph of the individual, their response to the manipulated image was removed from the analysis.

Results

Percent correct responses by cue type and stimulus group are shown in Figure 2. The manipulations did not obscure the apparent race and gender category of the faces, and so a subject viewing a particular face could narrow the search down to the ten individuals of each stimulus group. For this reason, chance performance for each face in this task is 10%. It is clear that performance was well above chance for all face types and regardless of whether shape or pigmentation cues were available.

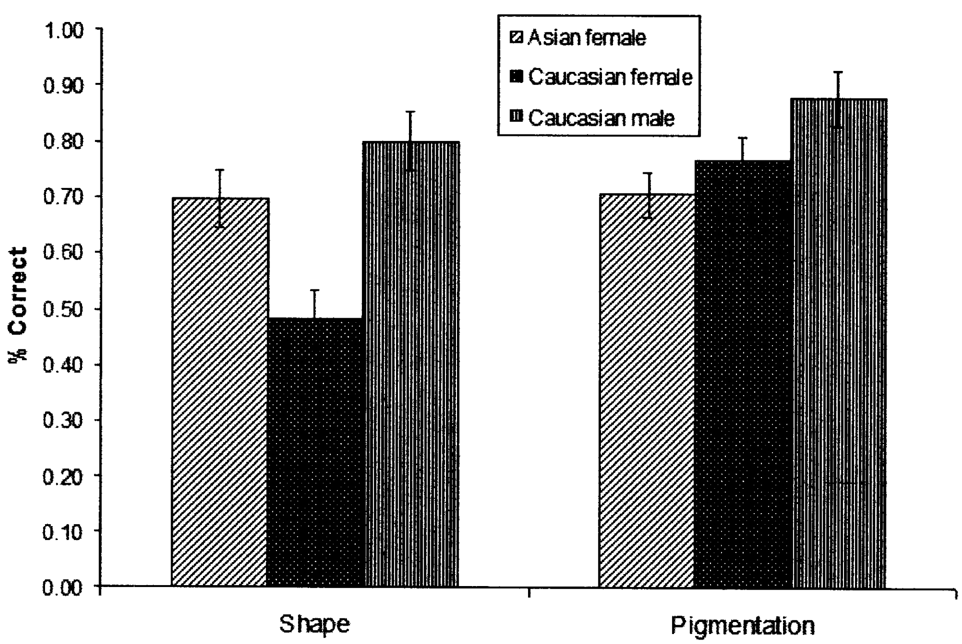

Figure 2
Results presented by cue type and stimulus group, expressed as percentage of correct responses. Error bars indicate standard errors of the mean.

There was no significant difference in performance between the two versions of the experiment or the two different presentation times, and so these factors were not considered for further analysis. We ran an ANOVA with cue type (shape, pigmentation), stimulus group (Asian female, Caucasian female, Caucasian male), and subject group (Asian female, Caucasian female, Caucasian male) as fixed factors, and percent correct responses as the dependent variable. There was a significant main effect of cue type $F(1,137) = 8.6$, $p < 0.01$, with higher performance overall in the Pigmentation condition. There was also a main effect of stimulus group $F(2,137) = 10.5$, $p < 0.001$, and an interaction between the two factors $F(2,137) = 4.5$, $p < 0.05$.

It is clear from the graph in Figure 2 that the difference between performance with shape and pigmentation cues was driven by the Caucasian female faces. This was confirmed with subsequent paired-samples t-tests, which found significant difference in performance between the Shape and Pigmentation conditions only for the Caucasian females $t(25) = 5.5$, $p < 0.001$, but not for Asian female or Caucasian males. This indicates that performance was better using pigmentation cues than shape cues only with the Caucasian female faces. There was not a significant effect of subject group $F(2,137) = 1.0$, $p = 0.39$, or any significant interactions between subject group and the other factors. Thus, there were not significant differences in how the subjects performed when viewing faces from their own or other race and gender groups.

Discussion

The primary finding of this experiment was that subjects were capable of recognizing images of their friends' faces even when only pigmentation cues were available. This provides direct evidence that pigmentation cues are useful for the recognition of familiar faces. Indeed, with all three groups of faces, performance using pigmentation cues alone was as good *or better* than performance using shape cues alone. This suggests that pigmentation cues are not only useful for familiar face recognition, but are in fact at least as important as shape cues. This is consistent with and extends previous findings that shape and pigmentation cues are about equally important for unfamiliar recognition (O'Toole et al 1999; Russell et al 2004; Russell et al (In Press)).

A drawback of the method of stimulus construction in this experiment is that shading cues were confounded with pigmentation cues rather than shape cues. Shading cues would probably fit better with shape cues, because shading can be used to infer shape (as in computer vision shape-from-shading algorithms (Horn 1986)). It is possible that relatively strong performance in the Pigmentation condition was a result of shading cues having been grouped with pigmentation rather than shape cues in this experiment.

However, we do not believe that this potential confound is a problem, for several reasons. We used diffuse lighting centered at 0^{o} elevation when photographing the faces. This resulted in images with no cast shadows and very minimized variation in shading, so that shading cues were less relevant than they are under top-lighting. In a previous study we used two kinds of stimuli: the same variety of stimuli used in the present experiment, and artificial faces with separate representations for three-dimensional shape and for surface reflectance, resulting in shading cues being grouped with shape cues (Russell et al (In Press)). In that study, which used a face matching task with unfamiliar faces, performance was almost exactly the same with the two varieties of stimuli, suggesting that there was little effect of classifying the shading cues with pigmentation or shape cues. This provides evidence that shading cues play at most a marginal role in stimuli of the variety used in the present experiment, and hence were not an important confound.

In the present experiment, there was no interaction between the subject group and the stimulus group. A well-known phenomenon, called the "own-race bias" or the "other-race effect", is that people are not as good at recognizing faces from a different racial group than their own (Meissner and Brigham 2001; Valentine et al 1995). The lack of an interaction between subject group and stimulus group indicates that we did not find an other-race effect in this experiment (or for that matter, an "other-sex effect"). However, this is likely due to the subjects all being familiar with one another, and living in a community in which both racial groups are well represented. The lack of an other-race effect here is consistent with the 'contact hypothesis'—that the other-race effect can be reduced or even reversed through increased experience with faces not of their own racial group (Chiroro and Valentine 1995; Sangrigoli et al 2005).

There were not significant differences in performance between the Shape and Pigmentation conditions for the Asian female or Caucasian male faces. However, with

Caucasian females, performance was significantly better in the Pigmentation condition than in the Shape condition. It is possible that this effect is idiosyncratic of the faces used here, and that we shouldn't infer that it is true for the racial and gender types in larger population. However, there is reason to believe that the difference is in fact meaningful. There is evidence, primarily from studies with Caucasian faces, that male faces may differ more in terms of their shape than do female faces (Ramsey et al 2005). If (Caucasian) male faces have more shape variance than female faces, people may have developed, through perceptual learning, strategies that exploit these differences. In this case, shape would be a more robust cue for male faces, consistent with the present findings, and those of O'Toole et al. (1999). This would not explain why shape was not less important for recognition of the Asian female faces. However, less is known about variation in Asian faces, and fewer studies have been conducted using Asian faces as stimuli. It is possible that the sex difference in shape variation of Caucasian faces does not exist in Asian faces.

The present study provides strong evidence that pigmentation as well as shape cues are important for familiar face recognition. This extends previous work finding that pigmentation cues are important for unfamiliar recognition. It is particularly critical support for the notion that pigmentation is important for face recognition because familiar recognition is the 'gold standard' for face recognition.

References

Biederman I, Kalocsai P, 1997 "Neurocomputational bases of object and face recognition" *Philosophical Transactions of the Royal Society of London B* **352** 1203-1219

Braje W L, Kersten D, Tarr M J, Troje N F, 1998 "Illumination Effects in Face Recognition" *Psychobiology* **26** 371-380

Bruce V, Healey P, Burton M, Doyle T, Coombes A, Linney A, 1991 "Recognising facial surfaces" *Perception* **20** 755-769

Bruce V, Langton S, 1994 "The use of pigmentation and shading information in recognizing the sex and identities of faces" *Perception* **23** 803-22

Chiroro P, Valentine T, 1995 "An investigation of the contact hypothesis of the own-race bias in face recognition" *The Quarterly Journal of Experimental Psychology* **48A** 879-894

Enns J T, Shore D I, 1997 "Separate influences of orientation and lighting in the inverted-face effect" *Perception & Psychophysics* **59** 23-31

Galper R E, 1970 "Recognition of faces in photographic negative" *Psychonomic Science* **19** 207-208

Gibson J, J., 1979 *The ecological approach to visual perception* (Hillsdale, NJ: Lawrence Erlbaum Associates)

Hancock P J B, Bruce V, Burton A M, 2000 "Recognition of unfamiliar faces" *Trends in Cognitive Sciences* **4** 330-337

Hancock P J B, Bruce V, Burton M A, 1998 "A comparison of two computer-based face identification systems with human perceptions of faces" *Vision Research* **38** 2277-2288

Hayes T, Morrone M C, Burr D C, 1986 "Recognition of positive and negative bandpass-filtered images" *Perception* **15** 595-602

Hill H, Bruce V, 1996 "Effects of lighting on the perception of facial surfaces" *Journal of Experimental Psychology: Human Perception and Performance* **22** 986-1004

Horn B K H, 1986 *Robot Vision* (Cambridge, MA: MIT Press)

Johnston A, Hill H, Carman N, 1992 "Recognising faces: effects of lighting direction, inversion, and brightness reversal" *Perception* **21** 365-375

Lades M, Vortbruggen J C, Buhmann J, Lange J, von der Malsburg C, Wurtz R P, Konen W, 1993 "Distortion invariant object recognition in the dynamic link architecture" *IEEE Transactions on Computers* **42** 300-311

Liu C H, Collin C A, Burton A M, Chaurdhuri A, 1999 "Lighting direction affects recognition of untextured faces in photographic positive and negative" *Vision Research* **39** 4003-4009

McMullen P A, Shore D I, Henderson R B, 2000 "Testing a two-component model of face identification: Effects of inversion, contrast reversal, and direction of lighting" *Perception* **29** 609-619

Meissner C A, Brigham J C, 2001 "Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review" *Psychology, Public Policy, and Law* **7** 3-35

Moses Y, Adini Y, Ullman S, 1994 "Face recognition: the problem of compensating for illumination changes" *Proceedings of the European Conference on Computer Vision* 286-296

O'Toole A J, Vetter T, Blanz V, 1999 "Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: an application of three-dimensional morphing" *Vision Research* **39** 3145-3155

Ramsey J L, Langlois J H, Marti C N, 2005 "Infant categorization of faces: Ladies first" *Developmental Review* **25** 212-246

Russell R, Sinha P, Biederman I, Nederhouser M, 2004 "The importance of pigmentation for face recognition" *Journal of Vision* **4** 418a

Russell R, Sinha P, Biederman I, Nederhouser M, (In Press) "Is pigmentation important for face recognition? Evidence from contrast negation" *Perception*

Sangrigoli S, Pallier C, Argenti A-M, Ventureyra V A G, de Schonen S, 2005 "Reversibility of the other-race effect in face recognition during childhood" *Psychological Science* **16** 440-444

Valentine T, Chiroro P, R. D, 1995 "An account of the own-race bias and the contact hypothesis based on a "face space" model of face recognition", in *Cognitive and computational aspects of face recognition: Explorations in face space* Ed T Valentine (London: London) pp 69-94

White M, 2001 "Effect of photographic negation on matching the expressions and identities of faces" *Perception* **30** 969-981

Chapter 5: Featural contrast: A novel cue for face classification

Abstract

Male skin is darker and redder than female skin, and these differences are known to provide a cue to judgments of facial sex classification. We sought to determine whether, in addition to these differences that extend over the entire skin surface, there are also spatially patterned sex differences in face pigmentation. Comparison of carefully controlled photographs of male and female Caucasian faces provided evidence female faces have greater "featural contrast" (luminance contrast between the eyes, lips and the rest of the skin) than do male faces. Manipulation of the amount of featural contrast was found to be cue to the apparent sex of a face. Consistent with this, the amount of featural contrast in actual faces was found to correlate positively with judgments of femininity and negatively with judgments of masculinity. Together, these results implicate an important role for featural contrast in judgments of face sex.

Introduction

One of the most salient attributes of a face is its gender. We effortlessly and immediately perceive almost all faces as being clearly male or female. Even in photographs with cues such as hair, cosmetics, and facial hair minimized, subjects perform at or close to ceiling when asked to classify faces as male or female (Bruce et al 1993). This is perhaps not surprising, given the high social import of determining the sex of an individual, and that the face is one of the primary cues for making this decision.

The sex of a face, as well as the degree of masculinity or femininity of the face is also very important for judgments about how attractive a face is. A number of studies have found that there is a strong correlation between ratings of beauty and ratings of femininity of female faces, and a weaker but still highly significant correlation between ratings of beauty and ratings of masculinity of male faces (Bronstad et al 2002; Bruce and Langton 1994; O'Toole et al 1998). In turn, ratings of masculinity and femininity are strongly correlated with the ease and speed of classifying faces as male or female (O'Toole et al 1998). The importance of the task of facial sex classification is further amplified by the evidence that the perception of face gender also plays an important role in judgments about facial attractiveness.

Several recent studies have found that pigmentation plays an important role in determining the sex of a face (Bruce and Langton 1994; Hill et al 1995; O'Toole et al 1997). Here, 'pigmentation' refers to the way that the surface of the face reflects light.

There are many ways that pigmentation can vary between faces, including how much light is reflected (albedo), whether particular colors are reflected (hue), small scale variation in reflection (texture), the degree of scattering of reflected light (specularity), and the degree of subsurface scattering of light (translucence). Faces can also differ in terms of more two-dimensional attributes, such as the relative amount of light reflected by two different regions of the face.

Despite the evidence that pigmentation in general is important for deciding the sex of a face, there is little evidence about which specific attributes of pigmentation are important. For a cue to be useful for sex classification, it must differ consistently between the sexes. Given the importance of determining the sex of a face, it is likely that any cue that consistently differs between the sexes will be used to perform this task. We can take the approach, then, of first determining how male and female faces differ. This approach has been taken in the context of three-dimensional shape (Bruce et al 1993; Burton et al 1993), the results of which have been very useful for understanding sex classification. We turn now to review evidence for sex differences in face pigmentation.

Tarr and colleagues investigated have provided evidence that Caucasian male and female faces differ in color (Tarr et al 2001). Specifically, they found that male faces are redder, and female faces are greener. It was also found that, when the faces were highly blurred, there was a significant correlation between the red—green ratio of the faces and subjects' classification of the faces as male or female. Though this correlation might be largely reduced for non-blurred faces, it would still constitute a relevant finding.

Viewing blurred faces is roughly analogous to viewing faces at a distance, and it is probably useful to be able to determine the sex of a face at a distance. Together, these results provide evidence that the overall color of a (Caucasian) face is a useful cue for determining the sex of a face.

Though no studies have investigated sex difference in the overall luminance of faces, there is an extensive literature in physical anthropology demonstrating that male skin is darker than female skin. That males are about 2-3% darker than females has been corroborated by many dozens of studies, with people from all continents and across the full gamut of skin darkness (for an extensive review, see (Frost 1988)). However, the majority of this work has been performed using photometers that measure reflectance from a single patch of skin, and typically a region of the body that receives little sun exposure (such as the inside of the upper arm) to control for the possibility of differential sun exposure between the sexes causing greater tanning in one sex than the other.

In reviewing sex differences in pigmentation, there is evidence that male faces are darker and, at least for Caucasian faces, redder than female faces. However, both these differences are across the entire face—they are 'one-dimensional' differences, in that they describe the entire face in terms of a single dimension of darkness or hue. But pigmentation is spatially extended—it varies across the face surface. For this reason it would be useful to know whether there are spatially patterned sex differences in pigmentation. Do the sexes differ in terms of the relative luminance of different parts of the face? And are any such differences important for sex classification?

There were two goals of the current study. The first goal was to investigate whether there are sex differences in pigmentation that vary across the face. That is, are there two-dimensionally patterned differences in pigmentation between male and female faces? This question is addressed in two experiments. In the first experiment, averaged male and female faces were warped into the same two-dimensional shape and then subtracted from one another, to determine whether there were any face regions for which average male and average female pigmentation differs. We did indeed find such differences, and the second experiment was designed to compare samples of male and female faces, to determine whether male and female faces were *consistently* different, and for testing the hypothesis that female faces have greater "featural contrast" (defined below) than male faces. The other goal of this study was to determine whether these differences are actually useful for deciding the sex of a face, and whether they are involved in impressions of the masculinity or femininity of a face. There are several kinds of pigmentation cues that could vary across the face, but for the current study, only albedo (the amount of light reflected) was considered.

Experiment 1

The first experiment consisted of an agnostic exploration of how male and female facial pigmentation might differ. We wanted an approach where we could compare different regions of the face separately in order to find two-dimensionally patterned differences in pigmentation. Toward this end we took photographs of male and female faces under controlled lighting conditions. We then morphed together the males and females separately to create averaged faces. These images were spatially registered, such that different parts of the faces were in the same location relative to each other and the image frame. These images could then be subtracted from one another to create difference images showing where the male and female averages differ in luminance.

Such a method was described by Rowland & Perrett (1995), in the context of comparing the appearance of faces of male and female fashion models. However, our purpose differed here, in that we wanted variation in the brightness of image pixels to be caused by variation in pigmentation, but not by variation in lighting, film exposure, cosmetics, or beards. To eliminate variation in lighting, it was necessary that each face fell under exactly the same illumination as the others, and in particular there were now systematic differences between the lighting conditions for the male faces and the female faces. It would be undesirable if, say, the female faces were photographed under lower illumination than the male faces, resulting in the female faces appearing darker than the male faces because of illumination rather than pigmentation. For similar reasons, it was important to use the same photographic exposure for all the images, so that variation in brightness was not caused by variation in exposure. Females are far more likely to wear cosmetics than men, and men are far more likely to be bearded than women. To avoid introducing sex differences in pigmentation due to these factors, all faces were photographed without cosmetics and cleanly shaven.

Even with all the photographs taken under exactly the same illumination, there is another potential confound. The amount of light reflecting off of a surface is determined not only by the illumination and reflectance properties of the surface, but also angle of incidence of the lighting upon the surface, which causes *shading*. This is the phenomenon that underlies the *chiaroscuro* technique for conveying shape in a two dimensional image developed by renaissance artists, and also the shape-from-shading algorithms of computer vision (e.g. (Horn 1986)). For the present study, this means that even if the lighting is exactly the same for a set of faces, the amount of light that is reflected by different regions of a face is dependent upon the shape of the face as well as its pigmentation. To reduce the effect of shading on the appearance of the faces, illumination conditions were chosen that reduced variation in shading between the faces, and eliminated cast shadows from protruding features. However, it would be impossible to entirely remove shading variation from such a set of photographs. For this reason, we will return later to this issue.

Methods

Frontal photographs were taken of 36 male and 31 female Caucasian faces. Of these, 22 male and 22 female faces were selected randomly for use in the current experiment. A smaller group of faces were used for this experiment because of the time consuming nature of the face morphing procedure employed. Care was taken to ensure that the faces were all photographed under the same illumination. Toward this end, all photos were taken in the same room, in which all light came from two studio lamps with diffusing heads. These two lamps, the camera and tripod, and the chair on which the photographic subjects sat, were kept in locations that were fixed with respect to each other and to the room. Similarly, the heights of the camera and lamps were fixed. The height of chair on which the subjects sat was adjusted, such that the subjects' heads were all placed at the same height. In this way every head was placed at the same position relative to the camera, lighting sources, and reflective surfaces of the room, ensuring the same illumination of each head. Similarly, the aperture and shutter speed of the camera were also fixed, ensuring the same exposure of each photograph.

To reduce shading cues, very diffuse lighting was used. Additionally, the walls of the room in which the photographs were taken were painted white, which had the effect of further diffusing the lighting incident upon the faces. The lighting sources were centered at 0^{o} elevation, which has been shown to minimize the utility of shading cues for face perception (Liu et al 1999).

Our interest was in determining the natural variation of pigmentation between the sexes, and so we ensured that all the faces were free of cosmetics when they were photographed. Similarly, we were interested in determining whether there are pigmentation differences between the sexes in the absence of beards, and so we ensured that all the faces were clean shaven when they were photographed. All faces were photographed with neutral expressions and without glasses or facial jewelry.

All 22 male faces, and separately all 22 female faces, were averaged together with Morph Man 3.0 (Stoik Imaging). This created averaged male and female faces. These are the left most images in Figure 1. These two images were then warped into the same, androgynous shape. We use the term image 'warping' to refer to movement of image pixels in the image plane (Benson and Perrett 1991), without the additional blending of pixel values of different images that is a part of image averaging (or 'morphing'). This was performed by morphing the two images together, but creating images in which the pixel values of the male and female images were not averaged together. In this way, one image was created with the average male pigmentation and androgynous shape, and another image was created with the average female pigmentation and androgynous shape. These are the images in the center column of Figure 1. Only the face outline and the internal features have the same shape—the hair, neck and shoulders differ between the two images. To eliminate these features, the images were cropped. These cropped images are on the right of Figure 1.

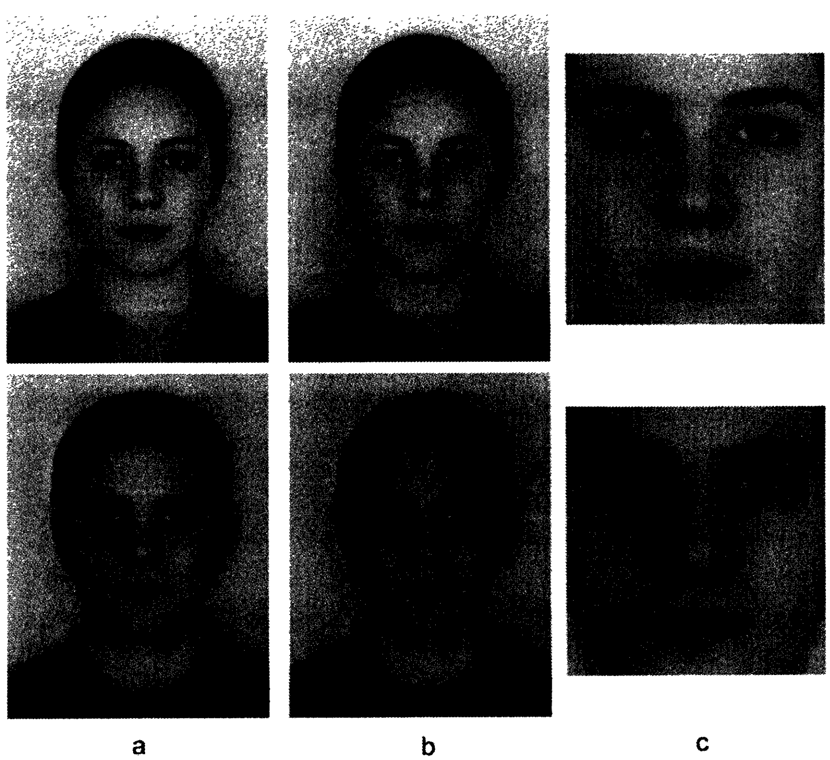

Figure 1. Images of averaged male and female faces. a) These images were produced by averaging images of 22 female faces, for the top image, and 22 male faces, for the bottom image. b) These images were produced by warping the faces on the left into the same androgynous shape. Note that the shape of the two images is the same only from the outline of the face inward—the hair, neck, and shoulders were not matched across the images. C) Cropped versions of the images in the center.

Even though the images on the right of Figure 1 do not differ in terms of shape, it is still readily apparent that the top image is female and the bottom is male. This confirms previous findings pigmentation cues are useful for performing sex classification. However, we want to go beyond this, and determine how male and female face pigmentation differs.

Results and discussion

A digital image can be represented as a matrix of pixel values. For an eight-bit, grayscale image, a single value between 0 (black) and 255 (white) represents the brightness of each pixel. To determine where two images differ, we can subtract the two matrices representing the images. We did exactly this, with grayscale versions of the images shown on the right of Figure 1. This subtraction determines which pixels are brighter, darker, or equally bright between the two images. To aid visualization, we created three images from this subtraction. These are presented in Figure 2. In the top image, white pixels indicate image regions that are brighter in the female average than in the male average. In the middle image, white pixels indicate image regions that are equally bright across the two images. In the bottom image white pixels indicate image regions that are brighter in the male average than in the female average.

Figure 2
Images produced by subtracting the images on the right side of Figure 1. In the left image, white pixels indicate areas of the face where the average female pigmentation is lighter than that of the male average. In the middle image, white pixels indicate areas of the face where the average male and female pigmentation were equal. In the right image, white pixels indicate areas of the face where male average pigmentation was lighter than the average female pigmentation.

The female average is clearly brighter than the male average in every part of the face, except for the eyes and the lips. The only regions in which some pixels are brighter in the male average are the eyes and lips. However, many of the pixels in these regions are brighter in the female average, and yet more are equally bright across the two images. It appears that while male skin is darker than female skin, the eyes and lips of the male and female average faces are about equally dark.

This is an interesting way for the pigmentation of the sexes to differ, because it is related to a proposal about how faces are detected. Under a wide variety of natural lighting conditions, the eyes and lips appear darker than the surrounding regions of the face (Sinha 2002). Because of this, the eyes and lips form an inverted triangle of dark spots surrounded by lighter areas, that has been proposed to form a unique 'signature' of faces. This notion has been implemented in an automated face detection system that searches for image regions that match this "ratio-template" of dark areas in an inverted triangle configuration surrounded by lighter areas (Sinha 2002). We adopt the convention here of referring to the ratio of the average luminance of the eyes and lips to the average luminance of the surrounding face regions as "featural contrast". Thus, another way to characterize the results of the current experiment is to say that the average female face has greater featural contrast than the average male face.

At this point we can return to the issue of shape and shading, to consider whether the sex differences between the male and female averages could be explained by shape differences. Work with laser-scan generated three-dimensional representations of faces has found that male brows are more protuberant (stick out further) than female brows, and male eyes are more deeply recessed than female eyes (Bruce et al 1993). Because recessed areas will receive less light overall, they will tend to be darker. Because male eyes are more recessed than female eyes, if shading were to have played a role, it would have been to make the contrast between the eyes and their sounding regions greater in the male faces than in the female faces. But we found the exact opposite pattern—that the female faces on average had greater contrast between the eyes and the surrounding regions than did the male faces. This suggests that shading probably did not play any significant role. There is little difference in protuberance between male and female mouths, again suggesting that shading played little to no role in causing the sex difference in featural contrast around the mouth.

These results are consistent with the physical anthropology results finding that male skin is darker than female skin. However, those studies did not investigate the face, and there are no previous data about the relative darkness of the eyes and lips. The current results suggest that, unlike the skin, there is no clear sex difference in how dark the eyes and lips are. They also suggest that female faces have greater featural contrast than male faces. But we only compared the mean values for male and female faces, not the samples from which they were drawn. In order to determine whether this difference is consistent across male and female faces, we need to compare the distributions of luminance for male and female faces. That is the goal of Experiment 2.

Experiment 2

The goal of Experiment 2 is to test the hypothesis that female faces have significantly more featural contrast than do male faces. Our approach to doing this was to hand-label on a large set of faces the locations of the eyes, lips, and the immediately surrounding regions. For each face we recorded four numbers that represented the average luminance in each of the regions. We then compared these numbers for the male and female faces, to determine whether they differed significantly in pigmentation. We also used these averages to calculate the featural contrast of each male and female face, and to determine whether there was a sex difference in this attribute.

Methods

We used all the faces from the larger database—36 males and 31 females. This constituted a superset of the faces used in the previous experiment. Grayscale versions of each of these images were hand-labeled in Matlab (The MathWorks, Inc.), to indicate the locations of the eyes, lips, and annulus shaped regions surrounding these features. A schematic depiction of way that these regions were defined appears in Figure 3. Note that the eyes included a narrow band of skin surrounding the whites of the eyes. This was included because the difference images from Experiment 1 found that it was also not different between the sexes.

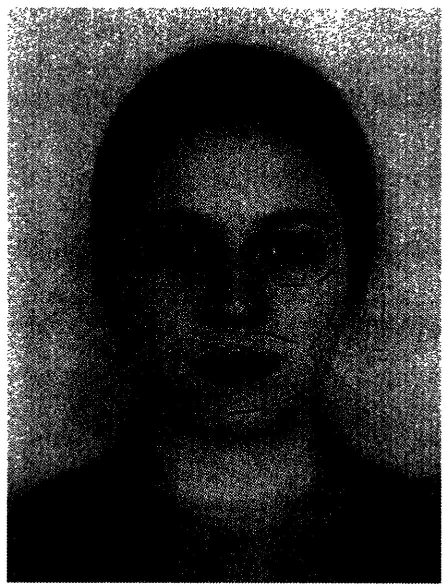

Figure 3. This figure shows the average female face, with lines drawn to indicate the approximate boundaries of the regions selected by hand on the photographs of male and female faces.

For each face, the average value of the pixels in each of these four regions (the eyes, the area surrounding the eyes, the mouth, the area surrounding the mouth) was determined, as was the Michelson contrast between the eyes and surrounding areas, and between the mouth and surrounding areas. Michelson contrast is the difference between light and dark areas, divided by the sum of the light and dark areas, which can be formally defined as $C_{Michelson} = (L_{max} - L_{min}) / (L_{max} + L_{min})$. The exact choice of contrast metric is not consequential at the relatively low contrast present in the face, as Michelson contrast and Weber contrast (the difference between light and dark areas, divided by the background (light) area) yield almost exactly the same results.

Results and discussion

The average luminance of the different regions, and the average Michelson contrasts are presented in Table 1. The average luminances for the different regions are presented as the eight-bit grayscale values. On this scale, higher numbers indicate brighter regions, with 0 representing black and 255 representing white. Michelson contrast can vary from 0 to 1, with higher numbers indicating higher contrast. To determine whether there were significant sex differences in these values, we compared them across the male and female faces in t-tests. The *t* and *p* values for these tests are also presented in Table 1.

		males (n=36)	**females** (n=32)	**T-test**
higher numbers indicate brighter regions:	eyes mean:	108	110	p = 0.26
	eyes surround mean:	141	149	p < 0.01
	mouth mean:	129	131	p = 0.24
	mouth surround mean:	156	164	p < 0.01
higher numbers indicate greater contrast:	Michelson contrast eyes:	0.137	0.149	p = 0.21
	Michelson contrast mouth:	0.095	0.110	p < 0.05

Table 1

These results followed the same pattern as was found in Experiment 1, and confirm that the sex difference in featural contrast is statistically significant for the mouth. The featural contrast of the eyes was larger in females than males, but the difference was not statistically significant at the $p<0.05$ level. The very small differences in luminance between both the eyes and the lips of the male and female faces are not significant. However, the there are significant differences between the male and female faces in terms of the luminance of the surrounding regions. The skin surrounding both the eyes and the lips was 2.5% darker for the male faces than the female faces. This is exactly in line with the findings from other studies using photometers, providing evidence that the sex difference in skin coloration extends to the face. The convergent results from photographic and photometric methods also provide evidence for the validity of the present photographic approach.

We can now firmly conclude that, for Caucasian faces, there is a spatially organized sex difference in pigmentation. The eyes and lips are equally dark across the sexes, but the rest of the face is darker in males than females. We can also express this sex difference as a sex difference in contrast, with female faces having greater featural contrast than male faces. Having now established a novel sex difference in pigmentation, it would be interesting to know whether the difference is actually used as a cue for sex classification.

Experiment 3

At this point, we would like to present a kind of experiment in which the reader can take part as a subject. Your task is to look at the faces below, and to decide which sex they are. Most observers will find that the face on the left appears female and the face on the right male. Some observers will see the sex of the faces less clearly, but will still find that the face on the left appears *more* female than the face on the right, which appears more male.

Figure 4. Two faces produced by manipulating an androgynous face. The androgynous face was created by morphing together the male and female average faces. To create these images, the eyes and lips were unchanged, but the rest of the image was lightened (to create the left image) or darkened (to create the right image). The left face has more featural contrast than the right face.

If you inspect the images closely, you will see that they are actually very similar to one another. Indeed, both were created by manipulating an androgynous face that was created by morphing together the male and female averages. In both images, the eyes and lips have not been changed, meaning that they are exactly the same across the two images. The rest of the images, outside of the eyes and lips, have been lightened (left image) or darkened (right image). That is to say that the two images are exactly the same, except most of the image on the right has been made darker, and most of the image on the right has been made lighter. Though the eyes and lips of the left face appear darker than those of the right, they are not—this appearance is the illusion of simultaneous brightness contrast.

The manipulation of the androgynous face to produce the two faces in Figure 4 was not arbitrary. It was chosen to simulate the difference between male and female faces that was discovered in Experiments 1 & 2. This shows that, given a lack of other sexually dimorphic features, this single sex difference in pigmentation is sufficient to make a face appear male or female, indicating its utility for sex classification. But is the important difference specifically that the male skin be darker than the female skin? Or is the degree of featural contrast the relevant cues for sex classification? Inspect Figure 5, again by asking whether the faces are male or female.

Figure 5. Two faces produced by manipulating an androgynous face. The androgynous face was created by morphing together the male and female average faces. To create these images, the eyes and lips were darkened (to create the left image) or lightened (to create the right image), but

the rest of the image was held constant in both images. The left face has more featural contrast than the right face.

In Figure 5, as in Figure 4, the face on the left appears female and the face on the right male, or at least the face on the left appears more female than the one on the right. To create these images, the same androgynous face used to create the images of Figure 4 was manipulated. This time, the eyes and lips were manipulated, and the rest of the face was held constant. The face on the left was created by darkening the eyes and the lips, and the face on the right was created by lightening the eyes and the lips.

The images of Figure 4 were manipulated to simulate the differences in pigmentation between male and female faces that were discovered in Experiments 1 & 2. The images of Figure 5 do not simulate these differences—they differ in terms of the luminance of their eyes and lips, while the skin does not differ. Yet perceptually, these manipulations have nearly the same effect. This suggests that what was relevant about the manipulations was not which part of the face became darker or lighter, but instead that the featural contrast was increased or decreased.

This provides evidence that the cue that is useful for sex classification is not the absolute darkness or lightness of the face. Instead, the useful cue is the relative luminance of the eyes and lips and the rest of the face—the amount of featural contrast.

Experiment 4

The manipulations of the androgynous face provide compelling evidence that the contrast difference is important for sex classification. But it would also be useful to know whether judgments about normal, un-manipulated faces are affected by the amount of featural contrast naturally present in the face. We chose to investigate ratings of masculinity and femininity. Ratings of femininity/masculinity are strongly correlated with reaction time to classify faces as male or female, as well as judgments of attractiveness (O'Toole et al 1998). We will look to see whether the Michelson contrast values determined in Experiment 2 predict the variance of ratings of masculinity and femininity for the same set of faces.

Methods

28 subjects provided ratings of masculinity and femininity for the same set of faces described in Experiment 2. The faces were cropped to a similar degree as the faces shown in Figures 4 & 5. One group of subjects viewed male faces first, and another group viewed female faces first. For the male faces, subjects were instructed to indicate how masculine they found each face, by pressing a number key from 1 to 7, with 1 indicating that they found the face not at all masculine, and 7 indicating that they found the face very masculine. For the female faces, the same instructions were given, but with the word 'feminine' rather than 'masculine'. The faces were presented in random order on a computer screen, one at a time, and subjects had unlimited time to view each face and respond.

Results

There were no significant effects of the sex of the subject, or the order in which the male and female faces were presented, and so these factors were removed from subsequent analyses. Ratings of masculinity or femininity were regressed with Michelson featural contrasts for the eyes and mouth. For the female faces, ratings of femininity were positively correlated with the mouth contrast, r=0.23, p<0.05, and with the eye contrast, r=0.21, p=0.06. For the male faces, ratings of masculinity were negatively correlated with the mouth contrast, r=-0.40, p<0.001 and with the eye contrast, r=-0.19, p=0.09.

The correlation between featural contrast and ratings of femininity and masculinity found in Experiment 4 closely mirrored the degree to which the featural contrast differed significantly between the male and female faces, as determined in Experiment 2. In Experiment 2 the distributions of male and female mouth featural contrast were significantly different from one another. However, the distributions of eye featural contrast were not significantly different, though there was a trend toward the female eye contrast being larger. In Experiment 4, the mouth featural contrast was found to be significantly correlated with judgments of masculinity and femininity, while there was only a trend toward the eye featural contrast being correlated with the same judgments. The consistency of the sex difference in featural contrast is closely linked to the utility of the cue for making sexually dimorphic judgments about faces.

Discussion

Comparison of photographs of male and female Caucasian faces found significant differences in the pigmentation of male and female faces. Male faces were found to be darker everywhere but the eyes and lips, which were equally dark in male and female faces. The amount of "featural contrast"—the luminance difference between the eyes and lips and the rest of the face—was found to be greater in female than male faces. Evidence from manipulated images of an averaged androgynous face and from ratings of actual faces suggested that the amount of featural contrast in a face is a useful cue for deciding the sex of the face and the degree of masculinity or femininity of the face, for at least Caucasian faces. This work joins other studies in finding a role for pigmentation cues in sex classification (Bruce et al 1994; Hill et al 1995; O'Toole et al 1997), and in finding specific sex differences in male and female pigmentation (Frost 1988; Jablonski and Chaplin 2000). These findings also join work finding a role for pigmentation in face perception more broadly, including face recognition (Bruce and Langton 1994; O'Toole et al 1999b; Russell et al 2004; Russell et al (In Press); Vuong et al 2005) and facial attractiveness (Jones et al 2004; O'Toole et al 1999a).

Though the correlations found between ratings of masculinity and femininity and featural contrast in Experiment 4 were significant, they accounted for only a small portion of the variance. Presumably this is because there are many other sexually dimorphic features that serve as cues for this judgment. These other sexually dimorphic features include shape cues, such as the protuberance of the brows and nose, and pigmentation cues, such the overall redness of the face and darkening caused by stubble. Other than cropping the images to exclude hair and necks, we did nothing to minimize these other cues, and hence the role they played in determining the judgments of masculinity/femininity. Given the large number of other available sexually dimorphic cues, it is meaningful that featural contrast played as significant a role as it did.

A clear limitation of the current study is that we only investigated faces from one racial group. It will be important to determine whether the same or other differences exist between the sexes in other racial groups whose overall pigmentation is substantially different from that of Caucasians. Another limitation of this study was its restriction to comparisons of luminance differences between male and female faces. An interesting study would be to investigate sex differences in spatial variation across the face in terms of hue, texture, and other pigmentation cues.

It was determined in Experiments 1 & 2 that male facial skin is darker than female facial skin. This adds to the large body of evidence that male skin is darker than female skin, and extends it by confirming that it is the case for facial skin as well as the skin of other body regions. However, these findings also raise a new question—if male skin is darker than female skin, why are male eyes and lips not darker than female eyes and lips? Answering this question will likely involve understanding why male and female skin is differently pigmented, which in turn may involve understanding why pigmentation varies at all.

One recent attempt to explain why skin pigmentation differs has suggested that it is a compromise between healthy and unhealthy results of exposure to ultraviolet (UV) radiation (Jablonski and Chaplin 2000). Exposure to UV light breaks down folic acid, which is necessary for cell division and producing new DNA. However, exposure to UV light also leads to the synthesis of previtamin D_3, which is necessary for preventing a variety of diseases such as rickets, osteomalacia, and osteoporosis, and also helps enhance calcium absorption. Pregnancy entails greater need for calcium, the absorption of which can be increased through higher blood concentrations of previtamin D_3. The result of this requirement of pregnancy is that men and women of childbearing age have slightly different needs for exposure to UV light, with women needing more. The primary method for shielding the body from UV light is darker pigmentation from melanin, hemoglobin, and carotene. In this account, the lighter skin of females is ultimately a result of their greater need for UV light to synthesize previtamin D_3 to support the greater calcium needs of pregnancy.

It is possible that the eyes and lips do not differ across the sexes because the balance between photo-protection and photo-absorption is different in these structures than in the rest of the skin. The eyes are particularly sensitive to light, and both the eyes and lips are particularly important for social communication. Because these facial features form a very small portion of the body area exposed to UV light, yet are sensitive and particularly important, it is likely that the need for photoprotection in these areas outweighs the small potential benefit of slightly more surface area for the synthesis of previtamin D_3. This may in fact partially explain why (for at least lighter skinned peoples) the eyes and the mouth are darker than the rest of the face in both sexes. The idea here is that while the sexes differ in their needs for sunlight absorption through the skin, they do not differ in their needs for sunlight absorption through their eyes and lips. A prediction of this proposal is that because there is no sex difference in skin darkness in people younger than menarche (Frost 1988), neither will there be a sex difference in featural contrast.

Several studies have found positive correlations between judgments of masculinity or femininity and the attractiveness of male and female faces, namely that more feminine female faces are considered more attractive, and more masculine male faces are considered more attractive (Bronstad et al 2002; Bruce et al 1994; O'Toole et al 1998). Here featural contrast was positively correlated with femininity and negatively correlated with masculinity. This sheds light on recent work finding that increasing or decreasing featural contrast has opposite effects on the attractiveness of male and female faces; female faces were rated more attractive with increased featural contrast than with decreased featural contrast, while the opposite pattern was found for male faces (Russell 2003). It follows from this that the sex difference in featural contrast has implications for judgments of facial attractiveness as well as facial sex classification.

This relationship between sex differences in featural contrast and beauty may explain part of why cosmetics are used the way they are. Cosmetics are commonly used by females, but not males, in an attempt to increase the apparent beauty of their faces. The use of cosmetics are not arbitrary; "Nosestick" and "ear shadow" do not exist, but lipstick and eye shadow are among the most commonly used cosmetics. And lipstick and eye shadow, are quite consistently used to darken rather than lighten the lips and eyes. We can re-describe this by saying that cosmetics are used to increase the featural contrast of female (but not male) faces. Recall that none of the individuals photographed for these experiments were wearing cosmetics, yet a sex difference was found in featural contrast. It would be unlikely for cosmetics to so exactly mimic exaggeration of this sex difference in pigmentation purely by coincidence. It is more likely that cosmetics are used the way they are precisely because they serve to accentuate the featural contrast, making the (female) face more feminine and, hence, attractive.

References

Benson P J, Perrett D I, 1991 "Synthesizing Continuous-Tone Caricatures" *Image and Vision Computing* **9** 123-129

Bronstad P M, Ramsey J L, Langlois J H, 2002 "Femininity = Attractiveness but Masculinity and Attractiveness Merely Share Variance", in *14th American Psychological Society Annual Convention* (New Orleans, LA: New Orleans, LA)

Bruce V, Burton A M, Dench N, 1994 "What's distinctive about a distinctive face?" *Quarterly Journal of Experimental Psychology* **47A** 119-49

Bruce V, Burton A M, Hanna E, Healey P, Mason O, Coombes A, Fright R, Linney A, 1993 "Sex discrimination: how do we tell the difference between male and female faces?" *Perception* **22** 131-52

Bruce V, Langton S, 1994 "The use of pigmentation and shading information in recognizing the sex and identities of faces" *Perception* **23** 803-22

Burton A M, Bruce V, Dench N, 1993 "What's the difference between men and women? Evidence from facial measurement" *Perception* **22** 153-76

Frost P, 1988 "Human skin color: A possible relationship between its sexual dimorphism and its social perception" *Perspectives in Biology and Medicine* **32** 38-58

Hill H, Bruce V, Akamatsu S, 1995 "Perceiving the sex and race of faces: the role of shape and colour" *Proceedings of the Royal Society of London B* **261** 367-73

Horn B K H, 1986 *Robot Vision* (Cambridge, MA: MIT Press)

Jablonski N G, Chaplin G, 2000 "The evolution of human skin coloration" *Journal of Human Evolution* **39** 57-106

Jones B C, Little A C, Burt D M, Perret D I, 2004 "When facial attractiveness is only skin deep" *Perception* **33** 569-576

Liu C H, Collin C A, Burton A M, Chaurdhuri A, 1999 "Lighting direction affects recognition of untextured faces in photographic positive and negative" *Vision Research* **39** 4003-4009

O'Toole A J, Deffenbacher K A, Valentin D, McKee K, Huff D, Abdi H, 1998 "The perception of face gender: The role of stimulus structure in recognition and classification" *Memory and Cognition* **26** 146-160

O'Toole A J, Price T, Vetter T, Bartlett J C, Blanz V, 1999a "3D shape and 2D surface textures of human faces: the role of "averages" in attractiveness and age" *Image and Vision Computing* **18** 9-19

O'Toole A J, Vetter T, Blanz V, 1999b "Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: an application of three-dimensional morphing" *Vision Research* **39** 3145-3155

O'Toole A J, Vetter T, Troje N F, Bulthoff H H, 1997 "Sex classification is better with three dimensional head structure than with image intensity information"

Rowland D A, Perrett D I, 1995 "Manipulating Facial Appearance through Shape and Color" *IEEE Computer Graphics and Applications* **15** 70-76

Russell R, 2003 "Sex, beauty, and the relative luminance of facial features" *Perception* **32** 1093-1107

Russell R, Sinha P, Biederman I, Nederhouser M, 2004 "The importance of pigmentation for face recognition" *Journal of Vision* **4** 418a

Russell R, Sinha P, Biederman I, Nederhouser M, (In Press) "Is pigmentation important for face recognition? Evidence from contrast negation" *Perception*

Sinha P, 2002 "Qualitative representations for recognition" *Lecture Notes in Computer Science* **LNCS 2525** 249-262

Tarr M J, Kersten D, Cheng Y, Rossion B, 2001 "It's Pat! Sexing Faces Using Only Red and Green" *Journal of Vision* **1** 337

Vuong Q C, Peissig J J, Harrison M C, Tarr M J, 2005 "The role of surface pigmentation for recognition revealed by contrast reversal in faces and Greebles" *Vision Research* **45** 1213-1223

Chapter 6: Sex, beauty, and the relative luminance of facial features

Published as: Russell, R. (2003). Sex, beauty, and the relative luminance of facial features. *Perception,* 32, 1093-1107

Abstract

It has been suggested that the consistent luminance difference between the darker regions of the eyes and mouth and the lighter regions that surround them forms a pattern unique to faces. One of the more consistent uses of cosmetics to make the female face more attractive is to darken the eyes and mouth relative to the surrounding skin. This study investigates the hypothesis that the size of the luminance difference between the eyes and mouth and the rest of the face affects the attractiveness of male and female faces differently. This possibility was tested in four experiments in which attractiveness ratings were obtained for images of faces in which the luminance difference between the eyes and mouth and the rest of the face had been manipulated. Female faces were found to be more attractive when this luminance difference was increased than when it was decreased, and the opposite was found for male faces. An interpretation consistent with these results is that the luminance difference between the eyes and mouth and the rest of the face is naturally greater in women than men. In this case increasing or decreasing the luminance difference will make a face more feminine or masculine, respectively, and hence, more or less attractive depending on the sex of the face. Implications for the causes of cosmetics usage are discussed.

Introduction

Recent research has firmly replaced the belief that 'beauty is in the eye of the beholder' with the notion that the perception of the attractiveness of a given face is largely consistent between observers, regardless of their age, sex, or cultural background (Etcoff 1999; Langlois et al 2000; Thornhill and Gangestad 1999; Zebrowitz 1997). A number of specific visual attributes that contribute to the attractiveness of a face have been proposed, including averageness (Langlois and Roggman 1990; Rhodes and Tremewan 1996), symmetry (Perrett et al 1999; Rhodes et al 1998), and sexual dimorphisms (Cunningham et al 1990; O'Toole et al 1998; Perrett et al 1998; Rhodes et al 2000).

Notably absent is explicit mention of photometric, or luminance attributes of facial attractiveness. The literature on facial attractiveness is almost entirely concerned with the location and shape of features. Much less attention has been paid to the relative luminance and coloration of those features. While a number of studies, such as those using photographic averages, have implicitly investigated attractiveness in terms of the relative luminance of different features, none have identified specific determinants of attractiveness in these terms.

A body of literature has shown that relative luminance and coloration patterns play a role in other aspects of face processing, such as face detection (Thoresz and Sinha 2001; Watt 1994), identification (Yip and Sinha 2002), and age estimation (Burt and Perrett 1995). Coloration has been found to play a particularly important role in sex discrimination (Bruce et al 1993; Hill et al 1995; Tarr et al 2001). Because past studies have suggested that attractiveness is related to other facial attributes, such as identity (Rhodes and Tremewan 1996), age estimation (Deffenbacher et al 1998; Zebrowitz et al 1993), and especially sex discrimination, it stands to reason that luminance patterns also play an important role in attractiveness.

Recent findings with low pass filtered images of faces suggest that large-scale luminance patterns are important in determining attractiveness (Sadr et al 2002). In this study, attractiveness ratings of faces at varying levels of blur were obtained. The attractiveness ratings given to blurred faces were consistent with the attractiveness ratings given to the same faces without any blurring. This suggests that much of what determines the attractiveness of faces is not the high frequency information in the fine features, but rather the low frequency information that was preserved in the blurred images. The relevance of this work to the present study is that we are likely to find luminance effects on attractiveness not in the fine details of the face, but rather in the placement and relative luminance of large regions of dark and light.

It has been suggested that the consistent luminance difference between the darker regions of the eyes and mouth and the lighter regions that surround them forms a pattern unique to faces. Roger Watt's work in image segmentation demonstrated that the horizontal regions centered over the eyes and over the mouth are darker than the horizontal areas above and below (Watt 1994). More recently Thoresz and Sinha have implemented a face detection algorithm based on the notion that the eyes and mouth are darker than the surrounding face regions (Thoresz and Sinha 2001). The algorithm uses a coarse template to look for image regions with this pattern of light and dark relationships, considering only the spatial and relative luminance relations, but not the absolute luminances. The algorithm is successful at detecting novel faces, including examples of several races. This is additional evidence that the face has a distinctive luminance pattern, with the eyes and mouth darker than the surrounding regions of the face. This distinctive luminance pattern of faces can be thought of as the "face pattern". Because attributes that are important in one domain of face processing are typically important in other domains, the face pattern likely plays a role not only in face detection, but also in aspects of face perception, such as attractiveness.

Though there are a number of uses of cosmetics, among the most common are darkening the eyes and mouth. Most cosmetics applied directly around the eyes or lips to change their coloration darken those features. This darkening of the eyes and mouth, without darkening other regions of the face, accentuates the face pattern by increasing its amplitude. It is likely not accidental that cosmetics are typically used in a way that accentuates the face pattern. Cosmetic use that changes the luminance or coloration of the face is far less common among men than women, historically as well as in the present (Corson 1972; Gunn 1973). It is a reasonable supposition that primarily women use cosmetics to accentuate the face pattern because only they are made more attractive by this transformation. This suggests that the relationship between attractiveness and the luminance difference between the eyes and mouth and the rest of the face (the amplitude of the face pattern) differs by sex, rather than being common to all faces.

The present study is designed to determine whether the size of the luminance difference between the eyes and mouth and the rest of the face affects the attractiveness of male and female faces differently. Toward this end, four experiments were conducted in which subjects rated the attractiveness of images of faces that had been manipulated by changing the luminance difference between the eyes and mouth and the rest of the face. The criterion for evaluating the hypothesis was whether the results showed an interaction between the sex of the face and the manipulation.

The first two experiments were the most critical. In both experiments, versions of the stimuli faces were made such that the luminance difference was increased, decreased, or left unchanged, in order to investigate the effect of this manipulation on the rated attractiveness of those faces. In the first experiment, versions of each face were made in which the eyes and mouths of the faces were darkened, lightened, or left unchanged, while the rest of the face was untouched. In the second experiment, versions of each face were made in which the eyes and mouths of the faces were left untouched, while the rest of the face was darkened, lightened, or left unchanged. Thus, in the first two experiments, there were versions of each face being rated in which the luminance difference between the eyes and mouth and the rest of the face was increased, decreased, or left unchanged, though the change was effected through different means. The reason manipulating the luminance difference in two different ways was to make more certain that it was the relative luminance difference between the eyes and mouth and the rest of the face and not the absolute luminance that caused any changes to the rated attractiveness of the faces. The attractiveness ratings of the faces were compared by sex and by condition (whether the luminance difference was decreased, unchanged, or increased). The presence of an interaction between sex and condition was to be taken as grounds for accepting the hypothesis, while the lack of an interaction was to be taken as grounds for rejecting it.

The third and fourth experiments were performed as controls for the first two. The third experiment, unlike the other three, used images of the faces that included the entire head and neck. The images were the same as those used in experiment 1, except that more of the head was visible. This experiment was performed to assess whether the results of the first two experiments would be the same when the external features of the face were visible. Again, presence of a sex by condition interaction was to be taken as evidence in support of the hypothesis. In the fourth experiment, there were also three versions of each face, though the entire image was darkened or lightened, with no portion being left untouched. This experiment was performed to ensure that the results of the other experiments, in which the images in some conditions were darker than those in other conditions, were not due to a low level effect of overall luminance. Because the manipulation did not change the relative luminance difference between the eyes and mouth and the rest of the face, presence of a sex by condition interaction was to be taken as evidence against the hypothesis in the fourth experiment.

General Methods

Stimuli

90 images, 45 male and 45 female, were selected from the Aberdeen set in the University of Stirling PICS database (http://pics.psych.stir.ac.uk). The set consists of mostly Caucasian faces with a wide range of adult ages. Image manipulations were performed using the image processing software Adobe Photoshop. The images were converted to grayscale to avoid confounding effects of hue and saturation. Except in experiment 3, the images were cropped such that only the inner features of the face were visible. The cropping was performed in order to reduce noise caused by extraneous variables such as hairstyle.

Subjects

In experiments 1 and 3, twenty paid subjects participated, ten male and ten female. In experiments 2 and 4, eighteen paid subjects participated, nine male and nine female.

Procedure

The procedure was the same for each of the four experiments. A within-subject design was used, with presentation order of the faces pseudo-randomized and counter balanced across subjects for sex and condition. Each subject saw all of the stimuli images. Images were displayed on the CRT monitor until the subject made a rating by pressing a key, at which time the next image appeared. Faces were rated on a 7-point Likert scale of attractiveness, with a response of 1 indicating that the face was very unattractive and a response of 7 indicating that the face was very attractive. For each experiment a three-way analysis of variance (ANOVA) was carried out on the attractiveness ratings, with sex of subject as a between-subjects factor, and condition and sex of face as repeated measures. Post-hoc pairwise comparisons (Tukey's HSD) were conducted when the analysis of variance found significant differences between conditions.

Experiment 1

Stimuli

To manipulate the size of the luminance difference between the eyes and mouth and the rest of the face, the stimuli in experiment 1 had the eyes and mouth darkened, lightened, or left unchanged, while the rest of the face always remained unchanged. This resulted in an increased, decreased, or unchanged luminance difference between the eyes and mouth and the rest of the face.

The burn tool in Adobe Photoshop was used to selectively darken the eyes and mouth, and the dodge tool was used to selectively lighten the eyes and mouth. The eye and mouth regions were hand defined for each face. The eye regions included the iris, sclera, and a narrow band of skin around the lashes. The mouth region consisted of the lips. Custom brush sizes were used for each feature of each face. The same brush was used to darken and lighten a given feature of a particular face, to ensure that the same areas of the face were being affected in each version. The "shadows" option was selected for both dodging and burning, causing only the darker pixels to be affected by the tool. This resulted in the sclera retaining its normal whiteness, and a more natural look for the manipulated features. Overall brightness manipulation would have caused the sclera to look dark—an unnatural and unhealthy feature. The luminance of the irises was increased or decreased by approximately 35 on a 256 level scale. When the features were darkened, the luminance difference with the rest of the face was increased , and when the features were lightened, the difference was decreased. Difference increased and decreased versions of each of the 90 original images were produced. Each stimulus face appeared in three conditions: difference increased, unchanged, and difference decreased. Examples of stimuli from experiment 1 are shown in Figure 1.

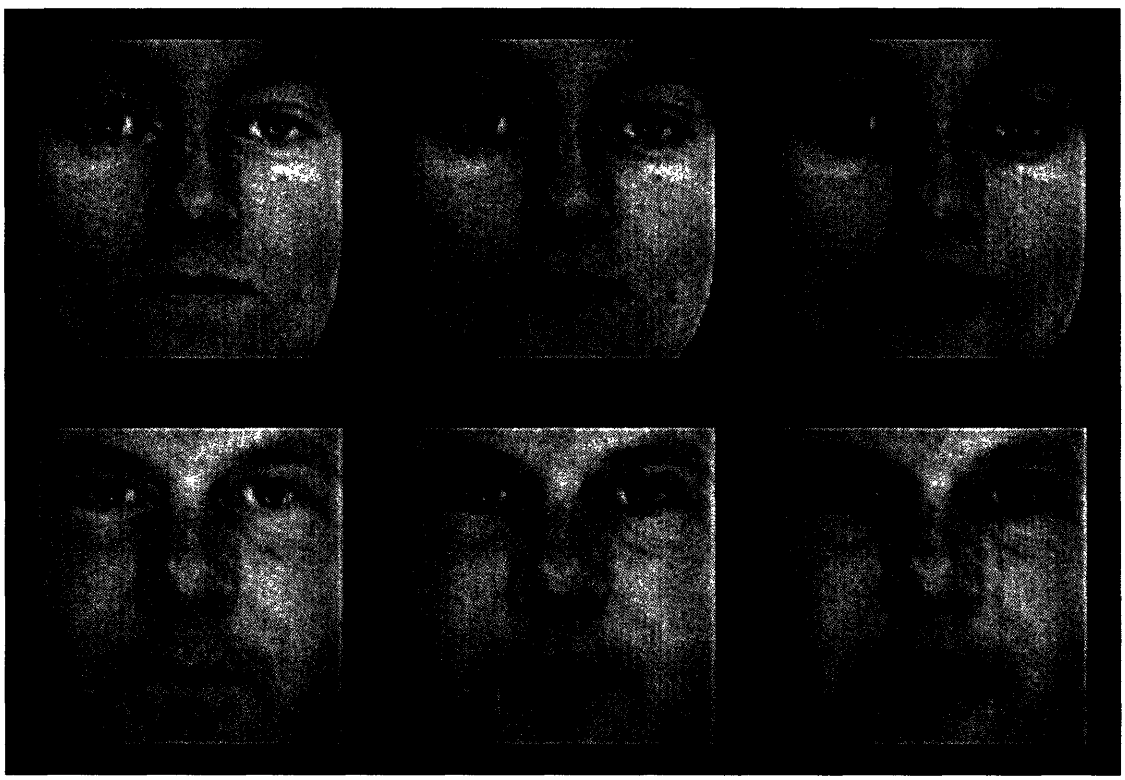

Figure 1. From left to right are examples of luminance difference decreased, unchanged, and luminance difference increased versions of female (top) and male (bottom) faces from experiment 1.

Results

The results of experiment 1 are shown in Figure 2. There was a significant interaction between condition and the sex of the face, ($F_{2, 36} = 22.49$, $p < 0.001$). Pairwise comparisons of the female faces found that the faces in the difference increased and unchanged conditions elicited significantly more attractive ratings than the faces in the difference decreased condition. The faces in the difference increased condition were rated more attractive than those in the unchanged condition, but pairwise comparison showed the difference to be not significant. The male faces showed the exact opposite configuration of results, with the faces in the difference decreased and unchanged conditions being rated significantly more attractive than the faces in the difference increased condition. Though the male faces were rated more attractive in the difference decreased condition than the unchanged condition, the difference was not significant. There were no other significant effects.

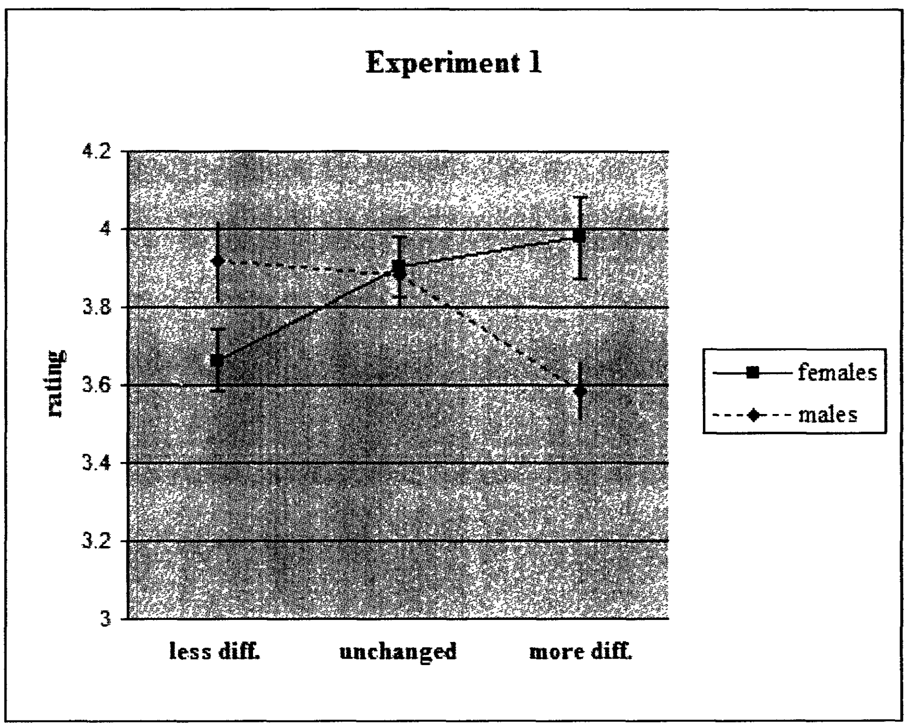

Figure 2. Results from experiment 1. Ratings for male and female faces are plotted for the three conditions. Higher numbers indicate higher ratings. Error bars are one standard error.

The results show a clear interaction between the size of the luminance difference and the sex of the face being rated. These results support the claim that male and female faces are affected differently by changes to the luminance difference between the eyes and mouth and the rest of the face. Specifically, female faces were rated more attractive when the difference between the eyes and mouth was increased than when it was decreased, while for male faces the opposite was the case.

Experiment 2

Stimuli

To manipulate the size of the luminance difference, the stimuli for experiment 2 had the eyes and mouth always remaining unchanged, while the rest of the face was darkened, lightened, or left unchanged. Darkening the rest of the face resulted in decreased luminance difference between the eyes and mouth and the rest of the face, while lightening the rest of the face resulted in increased difference.

To create the stimuli for experiment 2, a mask was drawn around the eyes and mouths of each of the original images. The eyes and mouth were defined using the same criteria as used in experiment 1. The eye regions included the iris, sclera, and a narrow band of skin around the lashes. The mouth region consisted of the lips. The mask was then inverted and feathered by five pixels to create a graded boundary. The brightness of the remainder of the image was then increased or decreased by 20 points on scale of 0 to 255, to produce two new images. Unlike in experiment 1, the method of manipulating the luminance difference between the eyes and mouth and the rest of the face did not affect the local contrast within the eyes and mouth. In one image the entire face, except the eyes and mouth, was lightened, and in the other image it was darkened. Because the eyes and mouth were held constant, luminance difference was increased when the rest of the face was lightened, and luminance difference was decreased when the rest of the face was darkened. As in experiment 1, each stimulus face appears in three conditions: difference increased, unchanged, and difference decreased. The stimuli were cropped to the same dimensions as in experiment 1. Examples of stimuli from experiment 2 are shown in Figure 3.

Figure 3. From left to right are examples of luminance difference decreased, unchanged, and luminance difference increased versions of female (top) and male (bottom) faces from experiment 2.

Results

The results of experiment 2 are shown in Figure 4. There was a significant interaction between condition and the sex of the face, ($F_{2, 32} = 8.55$, $p < 0.001$). Pairwise comparisons of the female faces found no significant differences between any of the conditions. However, for the male faces, significant differences were found between all three of the conditions, with the difference decreased faces more attractive than the unchanged faces, which were in turn more attractive than the difference increased faces. There was also a significant main effect of condition, ($F_{2, 32} = 9.82$, $p < 0.001$), that was driven by the effect of the manipulations on the male faces. Pairwise comparisons of the conditions, for the male and female faces combined, showed the faces in the difference decreased and unchanged conditions to be significantly more attractive than the faces in the difference increased condition. Though the faces were rated more attractive in the difference decreased condition than in the unchanged condition, the difference was not significant.

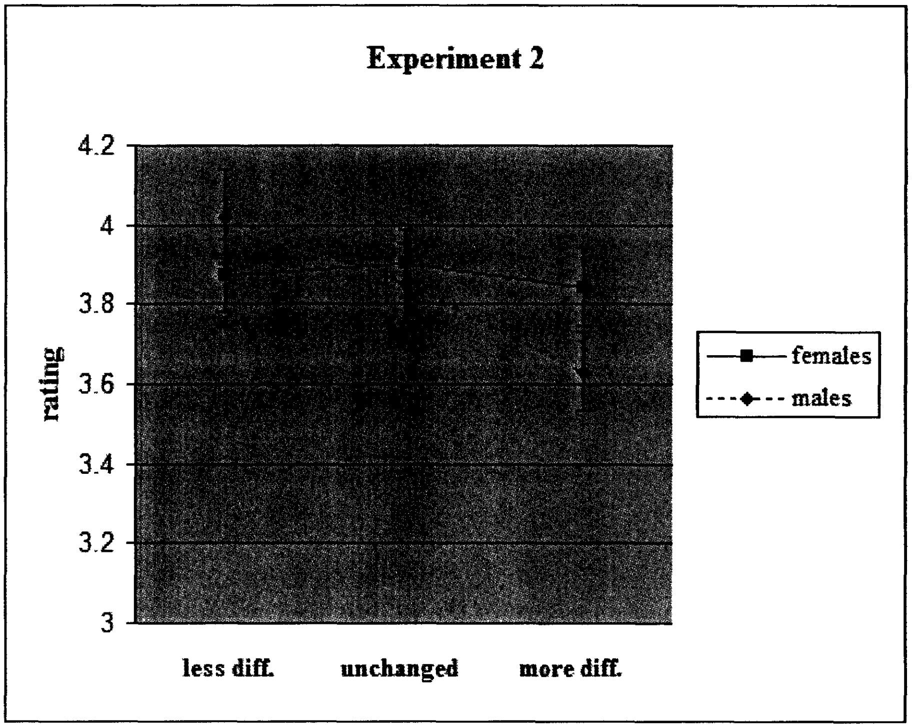

Figure 4. Results from experiment 2. Ratings for male and female faces are plotted for the three conditions. Higher numbers indicate higher ratings. Error bars are one standard error.

The results of experiment 2 show a clear interaction between the luminance difference and the sex of the face being rated, again supporting the claim that male and female faces are affected differently by to the size of the luminance difference between the eyes and mouth and the rest of the face. The results for male faces in experiment 2 were consistent with the notion that male faces are more attractive when the luminance difference between the eyes and mouth and the rest of the face is decreased, and less attractive when it is increased. However, the same manipulations had no effect at all on the female faces, casting some uncertainty on the assertion that this luminance difference plays a role in female attractiveness. Both the male and female faces in the difference decreased condition looked tanned, and in the difference increased condition they looked pasty. It is possible that this interacted with the effect of manipulating the luminance difference of the face pattern, resulting in the overall main effect of condition and the lack of difference between conditions for the female faces.

Experiment 3

Stimuli

The same faces, with the same manipulations, were used in experiment 3 as in experiment 1, except that they were left uncropped, so that the hair and bounding contours of the face were visible. This experiment was performed to ensure that the results of the first two experiments were not due to artifacts of the cropping procedure. Examples of stimuli from experiment 3 are shown in Figure 5.

Figure 5. From left to right are examples of luminance difference decreased, unchanged, and luminance difference increased versions of female (top) and male (bottom) faces from experiment 3.

Results

The results of experiment 3 are shown in Figure 6. There was a significant interaction between condition and sex of face, ($F_{2, 36} = 40.06$, $p < 0.001$). Pairwise comparisons of the female faces found the unchanged faces to be significantly more attractive than the difference decreased faces. Though the unchanged faces were rated more attractive than the difference increased faces, which were in turn rated more attractive than the difference decreased faces, neither difference was significant. Pairwise comparisons of the male faces showed the faces in the difference decreased and unchanged conditions to be rated significantly more attractive than the faces in the difference increased condition. Though the male faces were rated more attractive in the difference decreased condition than the unchanged condition, this difference was not significant. There were also significant main effects of condition, ($F_{2, 36} = 13.13$, $p < 0.001$), and sex of the face, ($F_{1, 18} = 15.24$, $p < 0.001$), with female faces rated more attractive than male faces. Pairwise comparisons between the conditions, for the male and female faces combined, revealed that the faces in the difference decreased and unchanged conditions were rated significantly more attractive than the faces in the difference increased condition. Though the faces were rated more attractive in the unchanged condition than in the difference decreased condition, the difference was not significant.

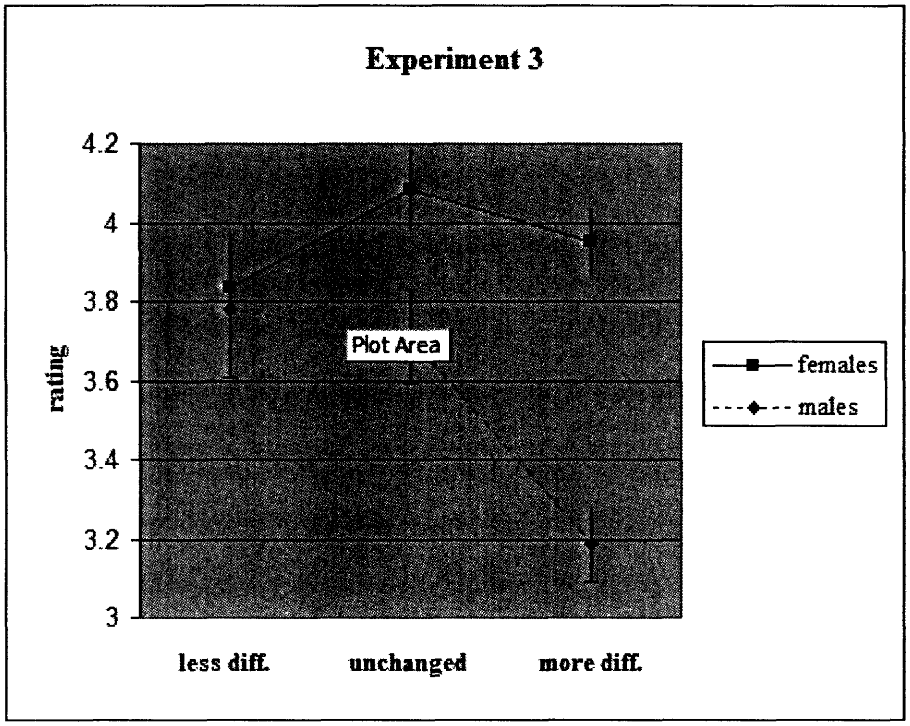

Figure 6. Results from experiment 3. Ratings for male and female faces are plotted for the three conditions. Higher numbers indicate higher ratings. Error bars are one standard error.

In experiment 3, as in the first two experiments with cropped faces, there was a significant interaction between the luminance difference and the sex of the face being rated, again supporting the claim that male and female faces are affected differently by the size of the luminance difference between the eyes and mouth and the rest of the face. However, the results of experiment 3 also differed from those of the two experiments with cropped faces in ways that reduce the comparability of the experiments. There were main effects of both sex and condition, with the females being rated more attractive than the males, and the difference increased faces being rated less attractive than the other conditions. The set of faces was well matched between the sexes for attractiveness without hair and facial outlines, but inclusion of those features made the male faces less attractive, and the female faces more attractive. This difference unfortunately makes the results of experiment 3 less suitable for comparison with the results of the experiments in which only part of the face is visible. The main effect of condition, with the difference increased faces less attractive than the other faces, was likely caused by an artifact of the manipulation rather than of the cropping. Normally the eyes and mouth of a face are not significantly darker than the hair. However, for most of the faces in the difference increased condition of experiment 3, the eyes and mouth were the darkest regions of the entire head. This unusual darkness of the features relative to the hair likely caused the lowered ratings for the faces of both sexes in the difference increased condition. Though the intensity of the manipulations was appropriate for the faces with internal features only, they were probably too large for entire heads. Yet despite these critical differences between the whole-head stimuli set and the cropped stimuli set, both yielded significant interactions between sex and condition.

Experiment 4

Stimuli

Experiment 4 was designed to test the possibility that the main effect of condition in experiment 2, and to a lesser extent experiment 1, was actually caused by an effect of overall luminance on attractiveness. The stimuli in experiment 4 were created by increasing or decreasing the luminance of the entire image by 20 points on a scale of 0 to 255. This was the same manipulation that was performed in experiment 2, except that it was performed over the entire image, including the mouth and eyes as well as the rest of the face. The relative difference between the eyes and mouth and the rest of the face was unaffected by this manipulation. Each stimulus face appeared in three conditions: darkened, unchanged, and lightened. The stimuli were cropped to the same dimensions as in experiments 1 and 2. Examples of stimuli from the experiment 4 are shown in Figure 7.

Figure 7. From left to right are examples of darkened, unchanged, and brightened versions of female (top) and male (bottom) faces from experiment 4.

Results

Figure 8 shows the attractiveness ratings of male and female faces under the three experimental conditions: lightened, unchanged, and darkened. There was no interaction between condition and sex of face, ($F_{2, 32} = 0.56, p = 0.6$). Though the female faces were rated more attractive than the male faces in all three conditions, the effect was not significant, ($F_{1, 16} = 1.79, p = 0.2$). The only significant effect in experiment 4 was a three-way interaction between sex of the face, sex of the rater, and condition ($F_{2, 32}$ = 8.15, $p < 0.005$). Figure 9 shows the attractiveness ratings of male and female faces, as rated by male and female subjects, under the three experimental conditions. This interaction is likely caused by the different ratings given by male and female subjects to the unchanged male faces. Post-hoc testing revealed no significant pairwise effects between conditions, as modified by sex of face and sex of rater.

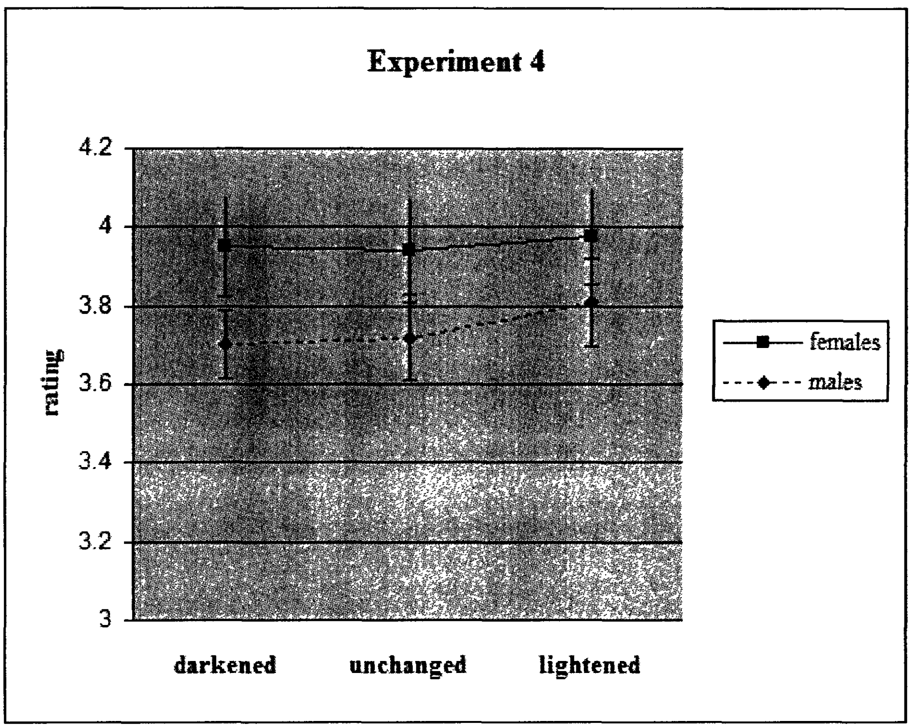

Figure 8. Results from experiment 4. Ratings for male and female faces are plotted for the three overall luminance conditions. Higher numbers indicate higher ratings. Error bars are one standard error.

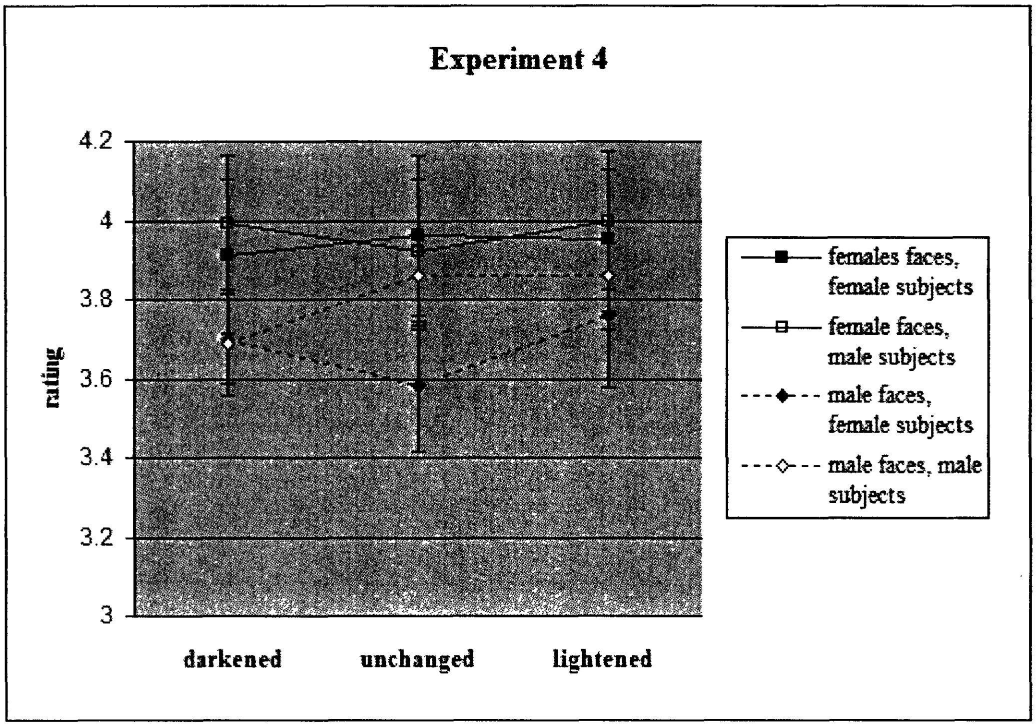

Figure 9. Results from experiment 4 showing the three-way interaction between the sex of the face, the sex of the rater, and condition. Ratings for male and female faces and male and female raters are plotted for the three overall luminance conditions. Higher numbers indicate higher ratings. Error bars are one standard error.

The results of experiment 4 discount the possibility that there is a low level effect of luminance on the judgments of attractiveness made in experiment 2. Though there was a significant three-way interaction between sex of the face, sex of the rater, and condition, it is not clear that there is any meaningful interpretation of this effect. More importantly, given the lack of main effects or two-way interactions, it is not clear what such a result might mean, or that it is relevant to the study. Critically, unlike the other three experiments in which the luminance difference between the eyes and mouth and the rest of the face was manipulated, there was no interaction between condition and sex of face in experiment 4.

Discussion

The affect on attractiveness of manipulating the luminance difference between the eyes and mouth and the rest of the face was investigated in four experiments. In all three experiments in which faces were manipulated such that the luminance difference between the eyes and mouth and the rest of the face could be larger or smaller, there was an interaction between the manipulation and the sex of the face being rated. In the one experiment in which overall luminance but not relative luminance was changed, there was no interaction between the manipulation and the sex of the face being rated. More specifically, changing the luminance of only the eyes and mouth (experiments 1 & 3) or only the rest of the face (experiment 2) affected the attractiveness of male and female faces differently. However, changing the luminance value of the entire image (experiment 4) did not affect the attractiveness of male and female faces differently. The results obtained support the hypothesis that the luminance difference between the eyes and mouth and the rest of the face does affect the attractiveness of male and female faces differently.

For the most part, female faces were more attractive with the luminance difference increased than with it decreased, while male faces were more attractive with the luminance difference decreased than with it increased. The female faces in experiment 2 were an exception to this. With these faces there was no effect of the manipulation, which consisted of changing the luminance of the rest of the face, while holding the eyes and mouth constant. In experiment 2, both the male and the female faces that were darkened in this way appeared to be somewhat tanned. Hence, it is possible that the appearance of tanning was a separate influence, balancing the possible reduction in attractiveness caused by reducing the luminance difference between the eyes and mouth and the rest of the face.

The results of experiment 3 did not neatly match those of experiment 1, even though the stimuli differed only by how much of the face was visible. However, much of the difference was due to the main effect between male and female faces when the hair and face boundary was visible. Though well matched across sex for facial attractiveness, the set was not well matched for the attractiveness of the entire heads. Aside from this main effect of the sex of the face, the other major difference between the results of experiments 1 and 3 was that the difference increased faces (eyes and mouth darkened) were rated less attractive for both male and female faces when the context of the external features was visible. This effect of condition was likely caused by the eyes and mouths being darker than the external features in the difference increased condition. It is likely the case that the luminance manipulations that were appropriate for the internal features-only faces of experiment 1 were simply too great for the entire heads, in which there was more context in which to read the relative luminance variations of the features. But despite these two differences between the results of experiments 1 and 3, a critical similarity remained. In both experiments there was an interaction between the sex of the face and the manipulation.

Regardless, further work will need to be performed in order to understand the affects on attractiveness of sex and the luminance difference between the eyes and mouth and the rest of the face, that were found in this study. Also, images of non-Caucasian faces will need to be tested to determine whether the results obtained in this study are common to attractiveness ratings of the faces of all races, and color images will need to be tested to make certain that the results are not an artifact of having used grayscale images.

The similarity between the appearance of the experimental manipulations on the eyes and mouth, and those caused by common cosmetics usage raises the issue of the causal relationship between cosmetics and the results of these experiments. One possibility is that because people are used to seeing the eyes and mouths of women but not men darkened by cosmetics, they find that darkened eyes and mouths are attractive on a woman, but strange on a man. In this case, cosmetics use is a cause rather than a result of the attractiveness effects suggested by the present findings. However, the results are also consistent with another explanation, that both cosmetic usage and the present results are caused by the affects on attractiveness of emphasizing a sexually dimorphic facial attribute. The sex dimorphism that could cause both cosmetics usage and the present results is a greater luminance difference between the eyes and mouth and the rest of the face in females than in males. The evidence for the existence of this dimorphism, and how its existence could cause the present results and cosmetics usage are described below.

Within all ethnic groups measured, male skin is darker than female skin (see (Frost 1988) for an extensive review of the literature). Recent evidence has suggested that skin pigmentation is a compromise between healthy and unhealthy results of exposure to ultraviolet (UV) radiation (Jablonski and Chaplin 2000). Exposure to UV light breaks down folic acid, which is necessary for cell division and producing new DNA. However, exposure to UV light also leads to the synthesis of previtamin D_3, which is necessary for preventing a variety of diseases such as rickets, osteomalacia, and osteoporosis, and also helps enhance calcium absorption. Pregnancy entails greater need for calcium, the absorption of which can be increased through higher blood concentrations of previtamin D_3. The result of this requirement of pregnancy is that men and women have slightly different needs for exposure to UV light, with women needing more. The primary method for shielding the body from UV light is darker pigmentation from melanin, hemoglobin, and carotene. In this account, the lighter skin of females is ultimately a result of their greater need for UV light to synthesize previtamin D_3 to support the greater calcium needs of pregnancy.

Systematic studies of the pigmentation of the eyes and lips in particular have not been performed. The eyes are particularly sensitive to light, and under separate genetic control for pigmentation. Both the eyes and lips are particularly important for social communication. Thus, there is reason to suspect that the pigmentation of the eyes and lips are subject to different demands than is the skin on the rest of the face. Because the facial features form a very small portion of the body area exposed to UV light, yet are sensitive and particularly important, it is likely that the need for photoprotection in these areas outweighs the small potential benefit of slightly more surface area for the synthesis of previtamin D_3. This could partially explain why the eyes and the mouth are darker than the rest of the face. Critically for the present study, it is likely that the compromise in the eyes and mouth between photoprotection and synthesis of previtamin D_3, unlike in the skin of the rest of the face, does not differ significantly between the sexes.

If it is the case that female eyes and mouths are as dark as those of males, the luminance difference between the eyes and the mouth and the rest of the face is greater in women than men, because the rest of the face is lighter in women than men. This provides reason to believe that the luminance difference between the eyes and mouth and the rest of the face could be sexually dimorphic, with a larger luminance difference in female than male faces. But how could such a sexual dimorphism explain the present results or cosmetics usage?

The effects of masculinity and femininity (the emphasis of sexual dimorphism) on the attractiveness of faces have been measured primarily in two ways. The first is by using facial morphing software to emphasize or de-emphasize the differences between faces, typically averaged male and female faces. The assumption is that the difference between these two (averaged) faces captures the features that make a face more masculine or feminine. The faces produced by morphing between and beyond these two faces are then rated for attractiveness to determine the degree of masculinity or femininity is considered to be most attractive. The other common methodology is to look for correlations between the ratings of masculinity or femininity and the ratings of attractiveness of particular faces. Increasing the femininity of a female face through morphing has consistently been shown to make it more attractive (Perrett et al 1998; Rhodes et al 2000). Also, ratings of femininity correlate strongly with attractiveness for female faces (Bruce et al 1994; O'Toole et al 1998). Evidence for the role of masculinity in determining male attractiveness is more mixed. Increasing the masculinity of a male face through morphing has both been found to make the face less attractive (Perrett et al 1998; Rhodes et al 2000) and to make it more attractive (Johnston et al 2001). The Johnston study used a slightly different methodology, in that the two faces used to create the morph were not an averaged male face and an averaged female face, but an averaged male face and a male face judged as highly masculine. However, ratings of masculinity consistently correlate strongly with attractiveness for male faces (Cunningham et al 1990; O'Toole et al 1998; Scheib et al 1999). The studies using morphing to investigate the relationship between masculinity and attractiveness are consistent in their methodology, using only the stimuli faces produced by morphing between two faces. The correlational studies investigated several (40 or more) images of actual (not morphed) faces. Thus, for the sort of stimuli used in the experiments reported here (actual faces rather than morphs), more masculine faces have been consistently found to be more attractive than less masculine (feminine) faces.

Masculinity and femininity are determined by sexually dimorphic qualities. Previous research, described above, has suggested that making a face more masculine or feminine will have opposite effects on the attractiveness of male and female faces. For example, the distance between the eye and eyebrow is greater in female than male faces. If this distance is increased, it will make a female face look more feminine, and hence *more* attractive. However, the same manipulation would also make a male face look more feminine, and hence *less* attractive. Accentuating a sexually dimorphic quality in a face will usually cause an increase in attractiveness if the face is of one sex, but a decrease in attractiveness if the face is of the other sex. That is to say that the same manipulation, to a sexually dimorphic feature, can cause opposite changes to the attractiveness of male and female faces.

In the present study, the same manipulations, increasing or decreasing the luminance difference between the eyes and mouth and the rest of the face, led to roughly opposite changes in the attractiveness ratings of male and female faces. Thus, the results of this study are consistent with the notion that the manipulations affected a sexually dimorphic attribute, which caused opposite changes to the attractiveness of male and female faces. This putative sexual dimorphism is the luminance difference between the eyes and mouth and the rest of the face. The presence of this same dimorphism may also explain the use of cosmetics by women to darken the eyes and mouth—that it accentuates a pre-existing sexual dimorphism, thereby making the face more feminine, and hence more attractive.

An obvious problem with this account is that it suggests that men also could use cosmetics to accentuate their masculinity, by lightening their eyes and lips or darkening the rest of their face. But it is far less common for males to use cosmetics than for females. Previous work has shown a stronger connection between femininity and attractiveness in women than for masculinity and attractiveness in men (Bronstad et al

2002; Bruce et al 1994; O'Toole et al 1998). Thus, it may be that men do not use cosmetics because they stand to gain less by accentuating their masculinity than do women by accentuating their femininity. Also, it could be said that males innately possess a kind of cosmetics for decreasing the luminance difference—their beards. As hair is almost always darker than skin, beards reduce the luminance difference between the mouth and the rest of the face.

Though the present results seem to support the assertion that the luminance difference between the eyes and mouth and the rest of the face is larger in females than males, we cannot yet conclude this. To conclude that males and females differ in the luminance difference between the eyes and mouth and the rest of the face, we will need to investigate the luminance patterns of male and female faces that are not wearing cosmetics, and are photographed under consistent lighting conditions from one sitter to the next. Unfortunately, the present stimuli set is not appropriate for addressing this question. The lighting and exposure of each of the images was not kept rigidly consistent, so comparing luminance relationships across images would likely tell us more about differences between the lighting and exposure conditions of the photographs than about differences between the faces. Also, some of the female faces are wearing cosmetics, but there is no documentation of exactly which faces or what cosmetics were applied to those faces. It would not be surprising to find that male and female faces differ in relative luminance when the females are wearing cosmetics and the males are not. Conclusions about the sexual dimorphism of relative luminance will require systematic examination of a set of faces photographed without cosmetics under rigidly consistent lighting. The assertion that the present study does support is that the size of the luminance difference between the eyes and mouth and the rest of the face affects the attractiveness of male and female faces differently.

References

Bronstad P M, Ramsey J L, Langlois J H, 2002 "Femininity = Attractiveness but Masculinity and Attractiveness Merely Share Variance", in *14th American Psychological Society Annual Convention* (New Orleans, LA: New Orleans, LA)

Bruce V, Burton A M, Dench N, 1994 "What's distinctive about a distinctive face?" *Quarterly Journal of Experimental Psychology* 47A 119-49

Bruce V, Burton A M, Hanna E, Healey P, Mason O, Coombes A, Fright R, Linney A, 1993 "Sex discrimination: how do we tell the difference between male and female faces?" *Perception* 22 131-52

Burt D M, Perrett D I, 1995 "Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information." *Proceedings of the Royal Society of London* 259 137-43

Corson R, 1972 *Fashions in Makeup: From Ancient to Modern Times* (London: Peter Owen)

Cunningham M R, Barbee A P, Pike C L, 1990 "What do women want? Facialmetric assessment of multiple motives in the perception of male facial physical attractiveness" *Journal of Personality and Social Psychology* 59 61-72

Deffenbacher K A, Johanson J, O'Toole A J, 1998 "Facial aging, attractiveness, and distinctiveness" *Perception* 27 1233-43

Etcoff N, 1999 *Survival of the prettiest: the science of beauty* (New York: Doubleday)

Frost P, 1988 "Human skin color: A possible relationship between its sexual dimorphism and its social perception" *Perspectives in Biology and Medicine* 32 38-58

Gunn F, 1973 *The Artful Face: A History of Cosmetics* (New York: Hippocrene)

Hill H, Bruce V, Akamatsu S, 1995 "Perceiving the sex and race of faces: the role of shape and colour" *Proceedings of the Royal Society of London B* 261 367-73

Jablonski N G, Chaplin G, 2000 "The evolution of human skin coloration" *Journal of Human Evolution* 39 57-106

Johnston V S, Hagel R, Franklin M, Fink B, Grammer K, 2001 "Male facial attractiveness: Evidence for hormone-mediated adaptive design" *Evolution and Human Behavior* 22 251-67

Langlois J H, Kalakanis L, Rubenstein A J, Larson A, Hallam M, Smoot M, 2000 "Maxims or Myths of Beauty? A Meta-Analytic and Theoretical Review" *Psychological Bulletin* 126 390-423

Langlois J H, Roggman L A, 1990 “Attractive faces are only average” *Psychological Science* 1 115-121

O'Toole A J, Deffenbacher K A, Valentin D, McKee K, Huff D, Abdi H, 1998 “The perception of face gender: The role of stimulus structure in recognition and classification” *Memory and Cognition* 26 146-160

Perrett D I, Burt D M, Penton-Voak I S, Lee K J, Rowland D A, Edwards R, 1999 “Symmetry and Human Facial Attractiveness” *Evolution and Human Behavior* 20 295-307

Perrett D I, Lee K J, Penton-Voak I, Rowland D, Yoshikawa S, Burt D M, Henzi S P, Castles D L, Akamatsu S, 1998 “Effects of sexual dimorphism on facial attractiveness” *Nature* 394 884-87

Rhodes G, Hickford C, Jeffery L, 2000 “Sex-typicality and attractiveness: are supermale and superfemale faces super-attractive?” *British Journal of Psychology* 91 125-40

Rhodes G, Proffitt F, Grady J M, Sumich A, 1998 “Facial symmetry and the perception of beauty” *Psychonomic Bulletin and Review* 5 659-69

Rhodes G, Tremewan T, 1996 “Averageness, exaggeration, and facial attractiveness” *Psychological Science* 7 105-10

Sadr J, Fatke B, Massay C, Sinha P, 2002 “Aesthetic judgments of faces in degraded images [Abstract].” *Journal of Vision* 2 743

Scheib J E, Gangestad S W, Thornhill R, 1999 “Facial attractiveness, symmetry and cues of good genes” *Proceedings of the Royal Society of London B* 266 1913-17

Tarr M J, Kersten D, Cheng Y, Rossion B, 2001 “It's Pat! Sexing Faces Using Only Red and Green [Abstract]” *Journal of Vision* 1 337

Thoresz K, Sinha P, 2001 “Qualitative representations for recognition [Abstract]” *Journal of Vision* 1 298a

Thornhill R, Gangestad S W, 1999 “Facial Attractiveness” *Trends in Cognitive Sciences* 3 452-60

Watt R J, 1994 “A computational examination of image segmentation and the initial stages of human vision” *Perception* 23 383-98

Yip A, Sinha P, 2002 “Contribution of color to face recognition” *Perception* 31 995-1005

Zebrowitz L A, 1997 *Reading Faces* (Boulder, CO: Westview Press)

Zebrowitz L A, Olson K, Hoffman K, 1993 “Stability of babyfaceness and attractiveness across the life span” *Journal of Personality and Social Psychology* 64 453-66

Chapter 7: Conclusions

The goal of the current body of work was to evaluate whether pigmentation plays an important role in face perception. Five studies involving three common perceptual judgments about faces were conducted to perform the evaluation. In each study pigmentation was found to play an important role in face perception. This forces a re-evaluation of the widely held implicit assumption that pigmentation plays an inconsequential role in face perception.

The first three studies (Chapters 2, 3, & 4) were performed to investigate the utility of shape and pigmentation cues for face recognition. Two of the studies used a face matching task with unfamiliar faces (Chapters 2, & 4). The other experiment (Chapter 3) used a naming task with familiar faces. In all three studies, pigmentation cues were found to be about as useful as shape cues. Indeed, in both studies in which color stimuli were used, pigmentation cues were found to be slightly more useful than shape cues. Along with the results of previous studies (Bruce et al 1991; Bruce and Langton 1994; O'Toole et al 1999), these data provide strong evidence that pigmentation cues are consequential for face recognition, and in fact are about as important as shape cues.

The fourth study (Chapter 5) was an investigation of whether there are consistent pigmentation differences between male and female faces, and whether any such differences are used to decide the sex of a face. This investigation found that there is a consistent sex difference in the amount of luminance contrast between the region of the eyes and the lips, and the rest of the face. Female faces have more contrast than male faces between these features and the surrounding face regions. The amount of this featural contrast was found to be positively correlated with judgments of femininity, and negatively correlated with judgments of masculinity. It was also found that this contrast difference is sufficiently important for our perception of face gender that simply changing the featural contrast of an androgynous face can bias viewers to perceive the face as male or female.

The fifth study (Chapter 6) investigated the effect of accentuating the featural contrast of male and female faces on ratings of beauty. It was found that changing featural contrast had opposite effects on male and female faces. Female faces were rated as more attractive with increased featural contrast than with decreased featural contrast, but the opposite pattern was found for male faces. This finding is consistent with the notion that pigmentation plays a role in judgments of facial beauty.

A straightforward conclusion to be drawn from these findings is that the human visual system represents the pigmentation as well as the shape of faces. Pigmentation could not play a critical role in the perception of faces if it were not represented by the visual system. In several of the studies here, there has been a close linking between the variation of a cue in the stimuli, and the utility of the cue for face perception, bringing to mind the dictum "ask not what's inside your head, but what your head's inside of". This linking between the utility of cues and the variation between faces in terms of those cues points toward an even more fundamental conclusion to be drawn from these results: from a visual standpoint, pigmentation variation is an important part of what faces actually *are*.

Faces have been assumed to be best described as collections of shapes, with pigmentation variation being a relatively unimportant afterthought. Perhaps this is because we have more language to describe shape than pigmentation. Certainly, we have better metrics for quantifying differences in shape than we do pigmentation. The current findings suggest that faces should really be conceived as collections of shapes *and* pigmentations—that what gives a face its individual and collective identity is its pigmentation as well as its shape.

Because this finding affects our understanding of the very nature of what faces *are*, it also affects virtually every aspect of our study of face perception. The finding also has implications for applied research relating to faces, including the construction of automated systems for face recognition, computer graphics, forensics and criminal identification, surgical planning, and cosmetic manipulation. More immediately, this finding argues strongly against the general use of stimuli that lack pigmentation cues in studies of face perception. It also argues against models of face perception that throw away pigmentation cues to rely exclusively on features derived only from shape cues.

There are several limitations of the work presented here, primarily relating to the kinds of faces and stimuli used, and the way that shape and pigmentation were defined. These studies relied exclusively on Caucasian and East Asian faces, and primarily on Caucasian faces. It is possible that the faces of other racial groups contain different relative amounts of shape and pigmentation variation. There is no a priori reason to believe that this is the case, but it is a possibility that the current results cannot address. For this reason, a more cautious summary of the results would be that pigmentation is important for the perception of Caucasian and East Asian faces.

The studies presented here all rest on the premise that it is meaningful and possible to separate shape and pigmentation. I argued earlier that while the division is not entirely complete, it is sufficiently complete as to be useful. A different approach toward evaluating the utility of the distinction between shape and pigmentation is to ask whether it is productive—are new and interesting questions produced by making this division. Several of the current studies have presented new and interesting findings that were directly the result of looking at faces specifically in terms of their pigmentation. For example, there have been a number of studies investigating facial sex classification; considering the issue with this conceptual distinction in mind was required for the discovery of consistent sex differences in face pigmentation. That the concept of a division between shape and pigmentation is productive gives it further validity.

The current findings do not constitute evidence that the visual system contains separate representation for the shape and pigmentation of faces. Though a distinction has been made here between cues about pigmentation and cues about shape, no evidence has been presented that addresses the question of whether the visual system makes such a distinction. We can only say that the visual system must represent face pigmentation somehow. We cannot address, on the basis of these studies, whether pigmentation is represented separately from other face cues, or together with shape and illumination cues as part of a unified representation of facial appearance.

As discussed above, an important limitation of these studies is their dependence upon faces of only two racial groups. For this reason, it would be worthwhile to investigate whether these findings about the utility of face pigmentation apply to faces of other racial groups. This would likely be particularly fruitful with respect to the sex differences in pigmentation. There were some discrepancies between the sex differences of Caucasian and East Asian faces, and it seems likely that sex differences in face pigmentation of other racial groups may contain even larger differences. This would also be relevant for evaluating the notion that cosmetics usage is related to sex differences in pigmentation. If other racial groups have other sex differences in pigmentation, we would expect that their optimal usage of cosmetics would differ accordingly.

A likely reason why researchers have been less aware of pigmentation as an important cue for face perception is that there are few useful metrics for describing pigmentation. We have many appropriate metrics for describing shape, such as distances and angles. Most metrics for describing pigmentation have no spatial extent—luminance and chrominance, for example. The current studies have hopefully made it clear that these dimensions are inadequate for describing the variation in pigmentation between faces. A critical task for future research in face perception will consist of determining the critical dimensions of face pigmentation. Progress in this area may come from studies of texture and material perception. It is also likely that many of the relevant dimensions of facial pigmentation involve large-scale spatial organization or patterning, as does featural contrast (described in chapter 5).

References

Bruce V, Healey P, Burton M, Doyle T, Coombes A, Linney A, 1991 "Recognising facial surfaces" *Perception* **20** 755-769

Bruce V, Langton S, 1994 "The use of pigmentation and shading information in recognizing the sex and identities of faces" *Perception* **23** 803-22

O'Toole A J, Vetter T, Blanz V, 1999 "Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: an application of three-dimensional morphing" *Vision Research* **39** 3145-3155