AGE ESTIMATION FROM CHILDREN'S FACES

ALEXANDRA E. HARRINGTON

A THESIS SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF ARTS

GRADUATE PROGRAM IN PSYCHOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

OCTOBER 2014

Abstract

In this thesis, we addressed the question of whether or not people could estimate age from children's faces 7 to 11 years of age. We found that undergraduates were able to make accurate relative age judgments for males and females, even in faces as little as two years apart, and that their performance improved as the age differences between the faces being compared increased. They were also able to make accurate absolute age judgments that increased with increasing face age for both genders. We also looked at estimate bias and while estimates were generally low in bias, the bias was in direction of the mean age of the stimuli. Additionally, we found that there is generally an advantage for male faces presented in frontal view. Finally, we looked at one possible factor influencing age estimates– facial expression. It was unlikely that facial expression was a primary cue informing age estimates.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Background

Processing faces is an important part of everyday life. Despite the fact that most humans are quite similar in appearance most adults are remarkably adept at recognizing not only facial identities, but emotions, sex, and age. Without this expertise, our social experiences and navigation of the world around us would be very different. Without our face processing expertise it would be more difficult to determine whether an approaching person is a stranger or a familiar friend, a source of assistance or a potential threat, a teenager or a senior. Indeed, research suggests that humans have specialized brain regions that respond preferentially to human faces: the fusiform face area (FFA) and the occipital face area (OFA) (e.g., Kanwisher, McDermott, & Chun; 1997 Rossion et al., 2003; but see also Gauthier, Skudlarski, Gore, & Anderson, 2000). Damage to these areas results in loss of the ability to recognize familiar faces (Gianotti & Marra, 2011).

As one might expect, face processing has received a great deal of research attention. Although much of this research has concentrated on face recognition, more recently there has also been interest in how humans process the age of faces. Age estimation is a task that we all carry out on a daily basis. Many of our social interactions with other people are heavily influenced by our estimate of their age. For example, the way a teenage boy approaches and interacts with a teenage girl will likely be different from the way he interacts with a senior. Apart from the importance of estimating age in the context of social behaviour, age estimation may also have serious consequences in the case of eyewitness testimony. If a witness to a crime were to provide an inaccurate age estimate, innocent suspects could be prosecuted, or guilty suspects dismissed.

A great deal of interest in facial aging also stems from the effort to develop computer systems that are able to predict the way faces will change with aging. This is important for both security (recognizing individuals years after a criminal offense has taken place) and age progression (photos used to help find missing children by approximating their present appearance after aging has occurred).

The literature on facial aging and age estimation from faces is diverse, including medical anthropometry (the structural changes of the face over time; e.g., Enlow, 1982), mathematical transformations that describe human growth (e.g., Todd, Mark, Shaw, & Pittenger, 1980), approximation of facial aging using computer algorithms (e.g., Ramanathan & Chellappa, 2006) and perceptual studies of human age processing and how it relates to particular ageing cues (e.g., George & Hole, 1995). In the following sections, I will provide an overview of the findings on human age estimation from faces before introducing the focus of this thesis, which is age estimation from children's faces.

1.1 Craniofacial Growth

The face and head undergo dramatic growth and changes with age, particularly from infancy through early adulthood (Enlow, 1982). The infant face is characterized by an overall wide and short shaped head, a large forehead with respect to the rest of the face, relatively large and wide-set eyes, a short nose (both vertically and in the extent to which it protrudes), a small mouth, a seemingly underdeveloped lower jaw, and almost entirely absent chin. The facial features as a group appear small in comparison to the cranium, with the face itself appearing quite flat when viewed in profile. As the infant matures, the jaw develops and the chin becomes more pronounced. Vertical facial growth occurs more rapidly than widening, so the face begins to resemble the longer, more narrow adult face. The nose and nasal bridge grow larger and more prominent with increasing age. The relative location of the ears appears higher with respect to the rest of the head and face. As the growth of the face begins to catch up with the eyes and eye orbits (which reach near mature size early in life), the eyes begin to appear smaller and closer together with respect to the rest of the face. The forehead begins to appear proportionally smaller compared to the rapidly growing face, although it does continue to grow at a slower rate. The forehead also becomes more slanted with age.

1.2 Variables Affecting Human Age Estimates

Overall, humans appear to be generally quite accurate at estimating age from images of faces. Most studies examining age estimation of unaltered photographs find that the absolute deviation of estimates from chronological age is less than 6 years, often between 2 and 3 years (Burt & Perrett, 1995; Dehon & Bredhart, 2001; George & Hole, 1995, 1998, 2000a; Hole & George, 2011; Moyse & Bredart, 2012; Nkengne et al., 2007; Sorqvist & Eriksson, 2007; Vestlund, Langeborg, Sorqvist, & Eriksson, 2009; Willner & Rowe, 2001). However, although it has been established that humans are generally fairly accurate in estimating the age of faces, it is still important to determine the characteristics of the face, as well as of the estimator, that are guiding these estimates.

1.2.1 Cardioidal Strain. One cue long thought to inform age estimates is cardioidal strain (Pittenger & Shaw, 1975a). D'Arcy Thompson first explained that the growth process of organisms could be described as a geometric distortion of a grid pattern placed over the 3 dimensional shape of the organism (Thompson, 1917). He claimed that the growth process involved the entire organism as an interdependent whole, rather than individual changes of body parts. He suggested that the geometric distortions were the result of physical forces on the organism, such as gravity and that these effects could be roughly described by a single mathematical transformation. It was Pittenger and Shaw (1975a) who first noted that applying a cardioidal strain transformation to face profiles produced changes that resembled craniofacial growth. Low levels of strain are associated with a small chin, lower facial features, and a relatively larger skull casing. With increased strain, the skull casing is smaller with respect to the rest of the head, the features are higher on the face, and the chin protrudes more (Pittenger, Shaw, & Mark, 1979). Varying levels of cardioidal strain applied to simple drawings of face profiles reliably produce changes in age estimates in the expected direction (Pittenger & Shaw, 1975a). Even pre-school aged children are able to categorize drawings of profiles with various levels of cardioidal strain into "baby," "boy," and "man" (Montepare & McArthur, 1986). In fact, applying cardioidal strain to drawings of other animals, and even inanimate objects (a Volkswagen Beetle!) produce this same pattern of shifting age estimates (Pittenger, Shaw, & Mark, 1979; but see Mark, Shapiro, & Shaw, 1983). However, cardioidal strain information alone does not appear to be sufficient for making age estimates from photographs of faces. George and Hole (1995) investigated the effects of different face image manipulations on age estimates, including mirror reversal, removing the external features of the face (so only the eyes, nose, mouth remained), thresholding the photographs (eliminating most textural information, colour, and all but the most extreme contrast), and pseudo-cardioidal strain (moving the internal features higher or lower on the head). The only manipulation that significantly reduced the accuracy of estimates was the threshold condition, even though this condition preserved cardioidal strain. Further, estimates in the features only condition, which does not preserve cardioidal strain, were not significantly different from age estimates of un-manipulated faces. This suggests that cardioidal strain information alone is not sufficient to accurately estimate age; at least not as accurately as when other information is available (George & Hole, 1995). In another study by George and Hole (1998) it was found that adding younger features to an older face decreased age estimates, despite cardioidal strain level remaining the same between the conditions. This further supports the idea that although cardioidal strain may influence age estimates to some extent, there are probably other cues, both local and global, that have considerably greater influence on age estimates. However, it is important to note that most of the stimuli in these studies were adult faces. Since most craniofacial growth takes place between infancy and early adulthood, cardioidal strain might be a much more useful age cue in children. Indeed, George and Hole (1995) found that applying pseudo-cardioidal strain to faces significantly affected age judgments in only the youngest group of stimulus faces (5–10 year of age). Additionally, some of the age cues available in adults such as skin texture and discolouration are probably not as informative in children's faces, perhaps making it more necessary to rely on craniofacial growth cues in children.

1.2.2 Skin. Other more global features, such as skin texture and colour also appear to be important in the estimation of age. One of the most obvious signs of aging is the change in skin texture. The outer layer of skin loses fat-like substances and becomes thinner, while the inner layer loses elasticity. This causes fine lines, wrinkles, thin or transparent skin, sagging, dryness, and susceptibility to damage ("Causes of

Aging Skin," 2010). A number of studies have shown that manipulating photographs to reduce skin texture information (either through blur, thresholding, or as an incidental consequence of creating composites of two faces) disrupts age estimates, particularly in older faces (Burt and Perrett, 1995; George & Hole, 1995, 2000). In our own lab, we have found that schematic human faces, largely devoid of any textural information, are difficult to classify into even broad age categories (DiBattista, 2008).

The colour of the skin itself also changes dramatically with aging, mostly as a result of sun exposure. Age spots, veins, and blotchy complexion can all become more prevalent with age ("Mature Skin," 2013). Burt and Perrett (1995) showed that colour information does indeed inform age judgments. They established this by identifying the average changes in colour between faces aged 20–59 and faces aged 50–54 years. When the value of this colour difference was added to younger faces, age estimates increased.

1.2.3 Local Features. The appearance of the local features of a face (e.g., eyes, nose, mouth) also seem to influence age estimates. George and Hole (1998) found that substituting the facial features of an older individual with features taken from a photograph when they were younger influenced age estimates in the direction of the facial features: that is, young features on an older face had lower age estimates than the original older face. This effect was observed despite the features themselves being in the same place on the face. Jones and Smith (1984) found that the ability of four year olds to discriminate the age of faces was most affected by masking the eyes of the stimulus face relative to any other masking condition (nose and cheeks, mouth and chin, the outline of the face). Although it is clear that facial features play an important role in age estimation, it is difficult to know whether it is the qualities of the features themselves (local), or the manipulations affecting global perception (e.g., substituting features affects interpretation of the entire face and head shape; George & Hole, 1998).

1.2.4 Facial Expressions. Though the research in this area is limited to one study, Voelkle, Ebner, Lindernberger, and Riediger (2012) found that the accuracy of age estimates was impacted by the facial expression of the stimuli. The age of neutral faces were estimated most accurately, and happy faces were most likely to be underestimated in age.

1.2.5 Hair Colour and Style. Though it has not been formally investigated to my knowledge, it is likely that qualities of the hair inform age estimates. Overall thinning, male-pattern hair loss, and greyness are all an indication of increasing age.

1.2.5 Race of Face and Race of Estimator. There may be a similar effect with respect to the race of the stimulus face and the race of the estimator, reminiscent of the well-known "other race effect" described in the face recognition literature (Dehon & Bredart, 2001). The only study investigating the influence of race on age estimates was by Dehon and Bredart (2001). They found that Caucasian people were significantly better at estimating the age of other Caucasians than of Africans; however, performance for the Africans was similar regardless of the race of the stimulus face. This might be explained by the amount of time the Africans included in this study had spent living in Belgium.

1.2.6 Age of the Face and Age of the Estimator. Some evidence suggests that the age of the stimulus face itself may affect the accuracy of age judgments. There appears to be a tendency to underestimate the age of older faces, and over-estimate the age of younger faces. For example, people appear to be extraordinarily poor at estimating the age of very old faces. On average, the estimated age of faces was almost

14 years younger than the chronological age in faces over 82 years of age (Kotter-Gruhn & Hess, 2012). Several other studies have also found a tendency to underestimate the age of older adults (Vestlund et al., 2009; Henss, 1991). Likewise, the tendency to over-estimate the age of younger faces has also come up frequently in the literature (Henss, 1991; Vestlund et al., 2009; Willner & Rowe, 2001;). In our lab, we found this tendency to over-estimate the age of young faces and underestimate the age of older faces existed in schematic faces as well (computer generated faces that preserve the geometric features of the face; DiBattista, 2008). However, Willner and Rowe (2001), found that 13 and 16 year olds were over estimated but 20 and 22 year olds were under-estimated; the estimates are biased in the direction of the mean of the age of the faces included in the experiment, rather than the population.

There is also a tendency for people to be better at estimating the age of faces similar in age to their own than those of another age group, as well as a general decrease in the accuracy of estimates with age. (Moyse & Bredard, 2012; Voelkle et al., 2012; George & Hole, 1995). For example, Voelkle et al. (2012) found that young people were more accurate and less biased in their estimates of young faces, whereas older people were more accurate and less biased in their age estimates of older faces. They also found that overall age estimation ability decreased with age.

1.2.7 Gender of the Estimator and Gender of the Face. Several studies have suggested that women may be better at estimating the age of faces than men; however the advantage is not consistent (Nkengne et al., 2008; Vestlund et al., 2009). Vestlund et al. (2009) found that women made less biased age estimates than men for faces aged 56 – 65, but there was no difference in estimate bias for age estimates of male and female faces aged 15 – 30 years. Nkengne et al. (2008) found that females made more accurate age estimates for female faces aged 15 – 65 years. However, Voelkle et al. (2012) found that male participants made less biased age estimates for male faces than females did, while performance on female faces was similar between male and female participants.

With respect to gender differences for target faces, several studies have found that age is estimated more accurately for male target faces than female target faces. Dehon and Bredart (2001) have noted that age estimates for Caucasian male faces are more accurate than those for Caucasian female faces. Voelkle et al. (2012) found that male and female participants' age estimates for female faces were less accurate and more biased than their estimates for male faces. One speculation is that this effect may be due to women investing more time and effort into looking younger (e.g., wearing makeup, having procedures done to minimize signs of aging, etc.). Consistent with this hypothesis, the difference in estimate bias between male and female faces was small in younger faces (aged 19 – 31) but much larger for middle-aged (39 – 55) or older (69 – 80) age groups (Voelkl et al., 2012).

1.3 Age Estimates from Children's Faces

Although the amount of research investigating age estimation from faces has increased, the number of studies that included children's faces is surprisingly limited. Of the handful of studies that included children, in most the youngest stimulus faces were teenagers (Sorqvist & Eriksson, 2007; Vestlund et al., 2009; and Willner & Rowe, 2001). George and Hole (2000a) looked at children's faces between 1 and 10 years of age and found that participants were able to provide age estimates for children's faces quite close to the actual age of the faces. For children between 5 and 10 years of age, the mean estimate for this age group was less than 2 years away from the mean age of the group (6.3 years), even when participants were presented with inverted face. However, these faces were analyzed as an age group rather than by age in years, and the face stimuli were presented along with adult faces up to 80 years old. There were only 3 faces included in their experiment between the ages of 5 and 10 years. George and Hole (1995) also looked at adult's ability to estimate age from 5 – 10 year old children's faces. Again, as a group estimates were quite close to the mean (no numbers were provided), even when only the internal features of the faces were presented. Moyse and Bredart (2012) included children as young as 10 in their study and found that young adult subjects were able to provide estimates with errors of only .01 years for children's faces and absolute errors of only 2.07 years with children's faces. Again, these values were analyzed as a group rather than by individual age and the faces were presented along with young adult and older adult faces. These studies suggest that people are quite good at estimating age from children's faces, even from internal facial features alone or inverted faces.

In our own lab, past research has suggested that when participants are presented with schematic faces, which preserve configural information but not information about the individual facial features, they are unable to categorize faces into even very broad face categories (Di Battista, 2008). Indeed, based on a principle components analysis carried out by Di Battista on the configural data digitized from her data base of photographs, this was not surprising- there is very little information available to distinguish age groups beyond separating babies, children, and adults. In a later study, the original photographs used to create the schematic faces in the previous study were used as stimuli in a new age categorization experiment. Although performance improved considerably, undergraduates had the most difficulty categorizing faces in the 7 – 11 year age range, even though the task only required participants to place faces into broad age categories (Personal Communication, Wilkinson). In a third unpublished study (Adams, 2009) our group found that participants had the most difficulty making relative age judgments when the faces being compared were children (see Chapter 4 introduction).

In this thesis, we looked further into age estimation from children's faces. Based on past research in our own lab, we wanted to use faces in the 7 – 11 year age range since this age range seemed to have the worst performance in our previous studies (Di Battista, 2008; Adams, 2009). We used a new face database in order to increase the number of photographs we had for each age (see Chapter 2) so that performance could be analyzed on a finer scale than it has been in past studies. In addition to revisiting the question of whether or not people are able to estimate age from children's faces accurately, there are a number of more specific hypotheses we wish to explore that to our knowledge have not been previously examined.

(1) In our first experiment, we examined the ability of participants to make relative age judgments (which is older?) from male and female face stimuli in both frontal and profile view (discussed in detail in Chapter 3). Research from George and Hole (1995; 2000a) and Moyse and Bredart (2012) suggests that participants are able to make accurate estimates within 2-3 years in this age range. Additionally, a study by Pittenger and Shaw (1975b) found that participants could sort faces only 1 year apart for faces 13-18 years old more often than chance. Based on this, we expected that participants would be able to choose the older face, at least in the larger age separations

significantly more often than chance. We expected that performance would decline as the age gap between the two faces being compared decreased.

(2) In the next study (Chapter 4), we examined participants' ability to make accurate absolute age estimates for 7 – 11 year old faces. Again, since previous studies indicated that participants' mean age estimates for faces 5-10 years old fell within a few years of the actual age of the group mean, we expected that participants would be able to provide reasonably accurate estimates. We predicted that participants' estimates would increase with increasing face age.

(3) Based on the adult face age estimation literature using adult faces, we expected that there would be directional bias in the age estimates that might take either of two forms. Participants might provide over-estimates for all the face stimuli, suggesting that estimates shift towards the mean of the entire age range of faces we see in every day life (Vestlund et al., 2009; Fahsing et al., 2004; Moyse & Bredart, 2011). Alternatively, the age of younger faces might be over-estimated, but the age of the older faces under-estimated, reflecting regression of mean estimates towards the mean of the sample of faces included in the study, rather than the mean of the entire range of possible face ages (Willner & Rowe, 2009)

(4) We introduced face viewing angle (frontal view and profile view) as a variable, which hasn't been looked at before in age estimation from photographs. Most studies looking at cardioidal strain use stimuli in profile view (line drawings of face

profiles which are subjected to different levels of cardioidal strain; Pittenger and Shaw, 1975a). Additionally, many of the craniofacial growth changes described by Enlow (1982) are more visible in profile view, such as how much the jaw, nose, and brow protrude. One prediction is that because facial growth information will be more clearly visible when viewing a face from the side, estimates will be better for faces presented in profile view. Alternatively, performance may be better for faces presented in frontal view because the relationship among the facial features, skin texture, and the qualities of the local facial features are more available in frontal view. This is also the view we are more familiar with, and therefore people may be more rehearsed in estimating age from frontal views of the face.

(5) The major focus in this thesis is on male faces due to limitations in our database (discussed in detail in Chapter 2). However, because Enlow (1982) suggested that craniofacial growth in female faces is less dramatic and they maintain more "youthful" features during development, we repeated the core experiments using the limited number of female faces available to us to determine whether age could be estimated at all from female faces in this age range.

(6) The final experiment examined the role of facial expression in age estimation. This was motivated in part by feedback from our participants after the first two experiments and also by Voelkle et al.'s (2012) finding that happy adult faces look younger than other facial expressions. In this experiment, we directly compared age estimates for angry, happy, and neutral facial expression in the same children.

Chapter 2: General Methods

2.1 Participants

Participants in these studies were undergraduate students enrolled in an introductory psychology course. They were recruited using the York Psychology Department's undergraduate research participant pool. Male and female students ranging in age from 17 – 69 years (mean = 21.2) participated in the experiments. Participants were from various ethnic backgrounds, reflecting the make-up of the city of Toronto. A total of 128 students participated in the studies comprising this thesis, 42 of whom were males. All experimental procedures were approved by the York University Office of Research Ethics - Human Participants Review Sub-Committee, and written informed consent was obtained from each participant. In exchange for participation, participants received course credit in their introductory psychology course.

Table 1 summarizes all of the participants in the thesis, divided into groups based on the tasks on which they were tested and the temporal sequence of testing. No participants were involved in more than one test session. Some tasks are replications or variants of tasks used for other groups. In the subsequent chapters, data is presented by task. If several groups completed the same task and the results were similar, the data were combined.

Table 2.1. Experimental Groups

Group Number	Average Age	Age Range	Number of Participants (males)	Experiments Completed
1	22.3	19–31	19 (5)	Relative task (males only), absolute task (males only, restricted response options)
2	23.1	18–42	12 (1)	Relative task (males and females), Absolute task (males and females, restricted response options)
3* relative absolute	19.8 19.7	18–34 18-34	21 (10) 20 (9)	Relative task (males and females), Absolute task (males and females, unrestricted response options)
4	20.5	17–69	35 (16)	Expression (between subjects)
5	21.5	17–48	41 (10)	Expression (within subjects)

Note. *There was one participant in session 3 who only completed the relative judgment task.

2.2 Stimuli

2.2.1 Faces. Our laboratory was given permission to use the Dartmouth Database of Children's Faces (Dalrymple, Gomez, & Duchaine, 2013). The children in the database photographs wore black toques to cover their hair (see Figure 2.1). Each photograph was 900x900 pixels and in colour. The database consisted of male and female faces between the ages of 5 and 16 years. Each child was photographed in frontal view, left and right near profile, and left and right three-quarter view. Each child was photographed in each of these viewing angles with eight facial expressions: anger, disgust, fear, happiness, neutral, pleasure, sadness, and surprise. The number of children photographed in each age group (1 year intervals) ranged from 1 to 13 years across all age groups.

Figure 2.1: Sample faces from the Dartmouth Database of Children's Faces. Face on left is Angry 7 year old male. Face on right is neutral 9 year old female.

We selected male and female faces aged 7 to 11 years for inclusion in our experiments. This was the range for which the largest number of faces was available. Additionally, our previous research showed that age estimation was difficult for this age range (see chapter 3 introduction). Photographs of children in two viewing angles (frontal and right profile) and three different expressions (neutral, angry, and happy) were used. Individual face stimuli in this age group were excluded if they were tilting their heads or if both eyes were visible in the profile image. This left a minimum of 6 faces in each male age group, so in cases where there were more than 6 faces in a group, we randomly selected 6 representative faces to keep group numbers equal. This resulted in a total of 30 male faces. For female faces, the same procedure was followed, except the number of faces per group was three due to a limitation in the number of female faces available in the database in this age range. This resulted in a total of 15 female faces.

2.3 General Testing Procedure

After providing informed consent, participants were given a verbal description of the task followed by written instructions on the computer screen. An entire experimental session ranged from 30 – 60 minutes, depending on which conditions were tested. All experiments were run on an iMac computer (screen dimensions 27 cm x 47 cm, resolution 1680 x 1050) using programs custom written in MATLAB® version R2012b with Psychophysics Toolbox Version 3. Participants were seated approximately 100 cm from the screen and asked to maintain this distance. At this distance, the computer screen subtended 15.4 x 26.5 degrees of visual angle at the eye of the observer. Responses were entered using the mouse and the computer keyboard. The testing took place in a dimly lit room.

After completing the experiment, participants were interviewed on the amount of contact they have had with children over the past five years (contact with children questionnaire, see Appendix A), and asked to provide any insight they had as to how they made their decisions during the experiment (questions adapted from Kuefner, Macchi Cassia, Picozzi, & Bricolo, 2008).

Chapter 3: Judging Relative Age

3.1 Introduction

Most previous studies have asked participants to provide absolute estimates of age rather than rank faces by age. To our knowledge, with the exception of some ranking tasks conducted to examine the effects of Cardiodal Strain (see Pittenger & Shaw, 1975a; Mark &Todd, 1985), tests of relative age judgment with adult observers have only been done in our lab. Ranking and relative age tasks have sometimes been used with young participants (e.g., George & Hole, 2000b) but not adults. However, it is possible that on a task like age estimation, making an accurate relative age judgment or estimate is more useful, not to mention more reminiscent of what most often takes place in every day life. With the exception of alcohol sales, making accurate relative age judgments are likely sufficient for social interactions, perhaps even more informative than an absolute judgment in some cases (is a woman younger or older than the rest of the women in a group of potential mates?; which of these young siblings is the youngest and requires the most supervision/care?).

An unpublished honours thesis conducted in our lab (Adams, 2009) examined undergraduates' ability to make relative age judgments across faces from five different age groups (< 1 year of age, 1–3 years, 4–7 years, 8–11 years, 20–35 years, 36–59 years and >60 years). Each face age group was compared to the immediately adjacent groups. The study found that participants were able to choose the older face of the pair better than chance for all of the age comparison groups; however, there was an obvious dip in performance in the comparison between 4–7 year old faces and 8–11 year old faces (Only 61.2% correct for this comparison versus 86.52% in all of the other comparisons). This dip in performance was evident for both male and female faces. Pittenger and Shaw (1975b) found that school photographs from each year grades 7 through 12 could be ranked in age order with performance better than chance, even when only the internal features of the face were provided; however, participants made more mistakes in the younger and older faces within this range.

In order to further investigate the ability of undergraduates to make age estimates from faces in this age range, we conducted a relative age judgment task using new face stimuli draw from the Dartmouth Database. As discussed in chapter 2, due to the number of faces available, the faces included in our study is limited to 7–11 year olds. Stimuli used in the current study are different from those used in Adams (2009) study in that they are colour photographs, the hair is hidden under a cap, and the faces were presented in two different viewing angles.

We expected, based on our previous findings, that participants would be able to perform better than chance on this task at least on the more distant comparisons. We expected that with increasing age difference between the two face pairs, performance would improve. We also examined the effects of viewing angle and gender on relative judgment performance.

3.2 Method

This task was completed by Groups 1, 2, and 3 (see Table 1). A total of 52 (16 male) participants completed the relative judgment task with male faces. Thirty-three (11 males) also completed the same task with female faces (Groups 2 and 3). The average age of participants was 21.5 years (age range 18–42). These participants also completed absolute age judgment experiments (see chapter 4), and the order of the two tasks (relative and absolute) was randomized for each participant. Participants did not receive any feedback on their performance for any of the tasks. Therefore, at the most, participants who had done the absolute judgment task first would have known the range of possible ages.

Participants were presented with two faces with neutral expression of the same gender and viewing angle, one on the right side of the screen, and one on the left side of the screen (see Figure 3.1). The two images were presented on a grey background. The faces themselves occupied about 8 x 6.3 degrees of visual angle. Participants were instructed to use the mouse to click on the older of the two faces (two-alternative spatial forced choice). Participants were given as much time as they needed to choose a face, and clicking a face advanced the program to the next pair of stimuli. Participants were instructed to keep a consistent distance between their eyes and the screen throughout the experiment.

Trials were blocked by both gender and viewing angle and the pair of individual faces being compared were randomized each time the program was initiated. In order to increase the number of face pairs tested, each participant was tested on two unique gender and viewing angle blocks (2 genders x 2 viewing angles x 2 unique runs = 8 total experimental runs). Any repetitions of the same face pair would happen only due to chance.

The exact date of birth for the children photographed in the stimuli were not known, so to avoid comparing faces that were very close in age (e.g. a seven year old who would turn 8 shortly versus an 8 year old who just turned 8), only faces with at least

2 years between them (e.g., 7 versus 9 years) were compared. This resulted in six comparison conditions of six face pairings (36 trials in total per run). Three conditions involved faces separated in age by 2 years (7 versus 9, 8 versus, 10, 9 versus 11), two separated by 3 years (7 versus 10, 8 versus 11), and a single comparison separated by 4 years (7 versus 11). As a consequence, some individual faces occurred more frequently than others in a run (two versus three times).

Figure 3.1: Face images were 4.5 cm from the right and left side of the screen, and there was 2 cm of space between the two face images. Participants clicked the face they believed was older to advance to a new face pair.

3.3 Analysis

Mean percent correct was calculated for each age separation (2, 3 or 4 years) and viewing angle (three age distances, two viewing angles for each gender of face stimuli). A within-subjects 3x2 ANOVA was used to examine the effects of comparison distance (2, 3, or 4 years) and viewing angle (frontal or right profile) for each gender separately.

All scores were compared to chance (50%). When there were multiple separations at the same comparison distance (e.g., 7 versus 9 and 8 versus 10), data were combined. In cases where asumptions of sphericity were violated, Greenhouse Geisser corrected F and p values were used.

3.4 Results

3.4.1 Comparison Distance – Male Data. Percentage of correct responses is plotted against comparison distance in Figure 3.2. All means exceeded chance performance (single sample t-tests; $p < .001$ in all cases). Two-way repeated measures ANOVA revealed significant main effects of comparison distance, ($F_{(2, 102)} = 191.461$, $p < .001$), and viewing angle, ($F_{(1, 51)} = 18.486$, $p < .001$); the interaction between these factors was not statistically significant ($F_{(2, 102)} = .080$, $p = .907$). As is evident in Figure 3.2, performance improved as a function of the number of years separating the face pair, and was consistently better for frontal male faces.

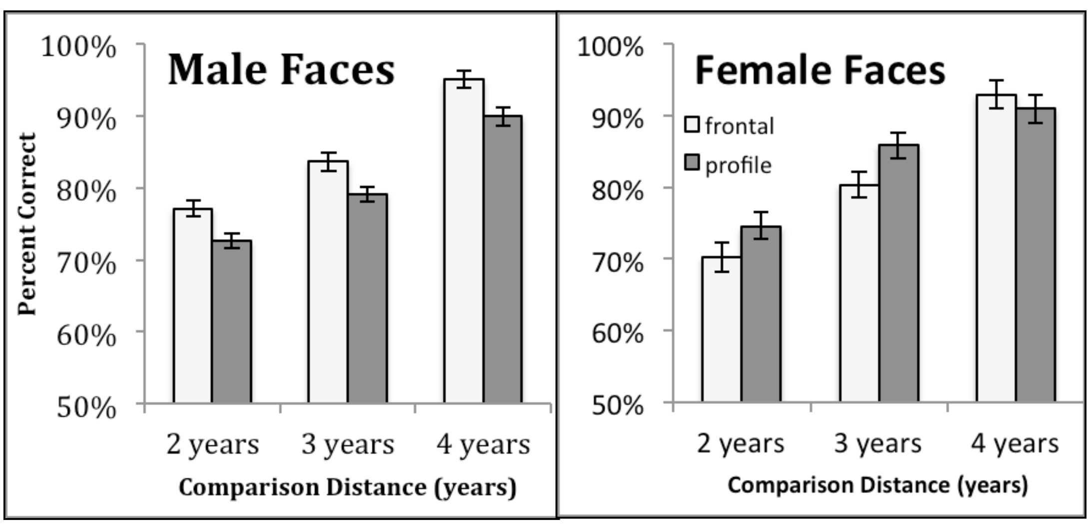


Figure 3.2: Percent correct responses for each age comparison distance category (2 years apart, 3 years apart, 4 years apart) separated by viewing angle (frontal and profile view). Left panel shows male face stimuli, right panel shows female face stimuli. Error bars are standard deviations.

3.4.2 Comparison Distance – Female Faces. We also examined the effect of comparison distance on relative age judgments for female faces (groups 2 and 3). Performance at all levels was above chance (single sample *t-p* < .001 in all cases). As with male faces, two way repeated measures ANOVA revealed a significant main effect of comparison distance, ($F_{(2, 64)} = 60.507$, $p < .001$), with performance improving as the age difference between the face pair increased (see Figure 3.2 right panel). The effect of viewing angle approached, but did not reach statistical significance ($F_{(1, 32)} = 3.610$, $p = .066$). A significant interaction was observed between viewing angle and comparison distance, ($F_{(2, 64)} = 3.551$, $p = .034$), with performance for faces in profile view in the 2 and 3 year comparisons, but performance better for frontal view for the 4 year comparison faces. Bonferroni post-hoc analysis (adjusted *p*-value .0167) revealed that there was only a significant difference between frontal and profile view for comparisons 3 years apart ($t_{(32)} = -2.722$, $p = .01$). The 2-year comparison approached significance ($t_{(32)} = -2.305$, $p = .028$). Interestingly, the effect of viewing angle was in the opposite direction of that observed for male faces – performance was better with female faces in profile view than in frontal view. There was no significant difference in performance between frontal and profile faces at the 4-year comparison distance ($t_{(32)} = .751$, $p = .458$).

3.5 Discussion

In summary, we found that undergraduates were able to accurately choose the older of two faces at rates better than chance, even when the faces were from children as little as two years apart in age over the age range of 7 to 11 years. When the age separation increased to four years apart, participants responded correctly 95% of the time with male faces in frontal view and 93% of the time with female faces in frontal view; nearly perfect performance.

Interestingly, for male faces, participants were consistently better at choosing the older face when the faces were presented in frontal view. This was contrary to our expectations. Although some features of male craniofacial development (protruding brow, jaw development) seemed as though they would be more visible in profile view, performance was actually better in frontal view. For female faces, the results for viewing angle were less consistent and less pronounced. Although there was a significant difference between frontal and profile performance at the 3-year comparison distance, it was actually in the opposite direction to that observed with the male faces – performance was better for faces presented in profile view. However, because we had only three faces per age group for females, results must be interpreted with caution.

Having established that participants were surprisingly good at deciding which of two faces was older, the next step was to determine how well they could provide an absolute age estimate for a single face, which is the subject of the next chapter.

Chapter 4: Absolute Judgments

4.1 Introduction

The number of studies that have investigated how well age can be estimated from children's faces is quite limited. George and Hole (1995) found that young (16–25) and older (51–60) adults were able to make age estimates for children's faces aged 5–10 years quite accurately when the faces were not manipulated in any way. Similarly, Moyse and Bredart (2012) tested children (aged 10–14), young adults (aged 20–30) and older adults (65–75) and found that for faces aged 10–14, all three groups provided fairly accurate age estimates. Young adults, on average had absolute estimate errors (accuracy) of only 2.07 years. Average absolute estimate errors for children and older adult participants were 2.31 and 2.56 years. The directional estimate errors (bias) for the young adult age group was only .01. Bias was .19 for children and 1.52 for older adults. Participants of all age groups over-estimated the age of faces in the 10–14 year old range, though the extent of estimate bias varied between the three participant age groups. In another study, Willner and Rowe (2001) found that alcohol servers between the ages of 18 and 50+ also over-estimated the age of 13 year olds, particularly for female faces, where on average faces were estimated to be 2.8 years older than they actually were. For male faces, bias was only .4 years.

Unpublished data from our lab found that undergraduates were able to categorize faces into one of seven age categories (0–1 years, 1–3 years, 4–6 years, 7–11 years, 20–35 years, 36–59 years and 60+ years) more accurately than chance. However, they found that performance was poorest when participants were categorizing faces in the 7–11 year age group. With male face stimuli, only 44% of faces were correctly categorized as being 7–11 years old (37% of the time, faces were incorrectly placed in the 4–6 year age category (underestimated)), and with female faces, only 36.6% of the faces were categorized correctly (35% were incorrectly estimated to fall in the 4–6 year old age group and 14.6% were incorrectly categorized as 20–35 year olds).

In all of the above experiments, the child face stimuli were included in the same experiments as adult face stimuli and presented in a random order. The faces in George and Hole (1995) study ranged from 5–70 years. Moyse and Bredart (2012) had a 10–14 year old group, a 20–30 year old age group, and a 65–75 year old age group. Willner and Rowe (2001) tested faces aged 13, 16, 20, and 22. In our lab, the faces ranged from 0–60+ years. It is possible that when participants are presented with a large range of ages, their estimates may pull towards the mean age of the faces. This may be an explanation for the tendency to over-estimate children's ages that was observed in most of the previous studies examining age estimation from children's faces. This makes it difficult to determine whether the tendency to over-estimate children's faces is an effect of regression towards the population mean or the sample mean since particularly in the case of George and Hole (1995) and Moyse and Bredart (2012), they means are quite similar. In our experiment, we included only children's faces, so the mean age of the stimuli was much lower than the population mean, allowing us to see how this affected the tendency to over-estimate the age of children's faces observed previously.

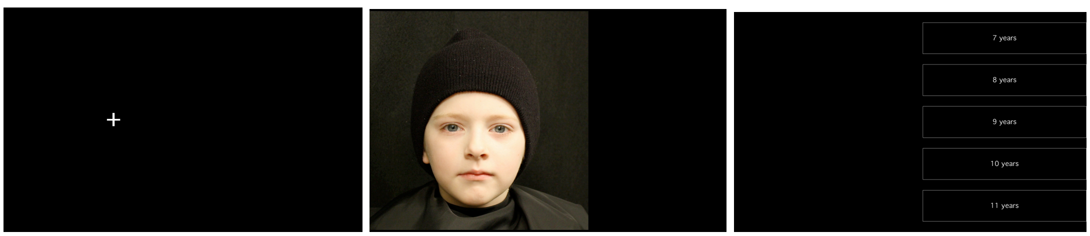
Additionally, age estimation of children's faces in all of these experiments was compared as a group (e.g., 5–10 year old age group) to other aged face age groups, rather than examined year by year. Because of this, we are unable to determine how subtle of an age discrimination people are able to make from children's faces. In our experiment, we used only a small age range, and therefore were able to make more subtle comparisons between individual ages.

Finally, for all but our own experiment, exposure times were quite long. George and Hole (1995) and Moyse and Bredart (2012) allowed participants as much time as they needed to make an age estimate and Willner and Rowe (2001) allowed participants approximately 5 seconds to decide. Since human faces can be matched with as little as 90 ms of exposure, we limited the exposure time to 1000 ms so that participants would have sufficient time to process the faces, but would not be allowed unlimited time to scrutinize faces feature by feature (Veres-Injac & Schwaninger, 2009).

Furthermore, based on the viewing angle effect found in our relative age estimation experiment, we once again tested participants on both frontal view and profile view faces. We expected to find an advantage for frontal view faces in age estimates for male faces. We planned to test participants with both restricted response options (so they could only choose an age between 7 and 11 years) and with unrestricted response options (participants were free to choose any age). We expected that participants would perform well on the restricted task because their estimates could only deviate from the actual age by 5 years due to the restriction, and because they were provided with the range of ages that the faces belonged to. With the unrestricted response task, we expected the participants to make less accurate estimates, but how the range of the estimates would be affected is unclear.

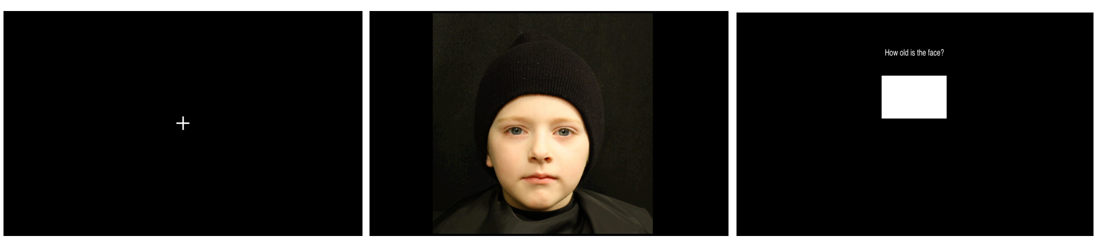
4.2 Method

In this task, participants were asked to make an age estimate for a single face stimulus with a neutral expression, presented in either the centre of the computer screen (unrestricted task) or left of centre (restricted task) (see Figure 4.1). Thirty-one participants (6 males) completed this task (see table 2.1), all of whom also did the relative task (chapter 3). The order of the two tasks was randomized across participants. Gender block order was randomized across participants.



2000 ms 1000 ms Response options

Sample absolute estimation trial with restricted response options



2000 ms 1000 ms Response box

Sample absolute estimation trial with unrestricted response options

Figure 4.1: Sample trials for restricted and unrestricted absolute estimation task. Left panel presented for 2000 ms, centre panel presented for 1000 ms, right panel presented until participant responds.

For the restricted task, participants were told at the outset that the faces ranged in age from 7–11 (restricted condition), whereas in the unrestricted task participants were blind to the age range of the faces.

Each face was presented 2 two viewing angles (frontal view and profile view), which were randomized across trials within a block. When the faces were displayed on the computer screen, the stimulus faces occupied 8 x 6 degrees of visual angle from the participant's viewing distance of 100 cm. At the beginning of each trial, a 2000 ms white fixation cross (~1 x 1 degrees of visual angle) was displayed on the screen, in the centre of the region where the face would appear (centre of the screen for the unrestricted task, left of centre for the restricted task). The face then appeared for 1000 ms on a black background. After the face disappeared, participants were either given five age response options (restricted task) on the right hand side of the screen and instructed to click one (See Figure 4.1, upper panel), or instructed to type in their age estimate into a text box at the centre of the computer screen (See Figure 4.1, lower panel). They were not instructed to provide estimates in whole years; however all but one participant did so. Participants were instructed to enter their responses carefully and to double check their response before pushing enter to advance to the next stimulus. The participants were allowed as much time as necessary make their estimate. Entering a response initiated the next trial.

There were a total of 60 trials for male faces (five age groups x six individual stimuli per age group x two viewing angles per individual face) and 30 trials for female faces (5 age groups x 3 individuals per age group x 2 viewing angles per individual face).

4.3 Analysis

4.3.1 Missing data. Due to an oversight with the program, participants in Group 3 were able to leave blank responses, or responses that were clearly in error (for example, entering an age of 66 instead of 6). Although this did not occur frequently, in the cases where data was missing, the average of the rest of the estimates in that category (e.g., 7 year olds in profile) was substituted for that data point. In the case of estimates that were outside of the range of 1–20 years, double digits of the same number (e.g., 66) were replaced with the single digit (6). For two digit estimates of different numbers (e.g., 76), the average of the two numbers was used to replace the missing data point (6.5). In cases where more than one data point was missing within an age/viewing angle group, the participant would be removed from analysis (this did not happen for this task).

4.3.2 Analysis. In addition to overall mean age estimates, deviation scores were calculated for each age estimate made (actual age subtracted from the age estimated) and then averaged across trials and participants to provide directional error scores for each age group and viewing angle. These signed deviations provide a measure of estimate bias. However, if there was no systematic error, over-estimates and under-estimates would cancel each other out (for example, three estimates of age 4 and three estimates of age 10 for six 7 year old face would yield a perfect mean error score for 7 year olds, even though the estimates themselves were quite far from the actual age of the faces). For this reason, we also calculated an accuracy score by taking the absolute deviation score for each estimate (absolute value of the estimated age minus the actual age) and calculating the mean absolute deviation for each viewing angle and age.

For mean age estimates, one-way repeated ANOVAs were conducted on the mean age estimates for age alone. We used frontal view estimates only in order to examine the effect of face age alone without any viewing angle effects of interactions. ANOVAs were conducted separately for each face gender and response option (restricted versus unrestricted) condition.

For mean bias, and accuracy, 2x5 (viewing angle x age) repeated measures ANOVAs were conducted to examine the effects of both age and viewing angle on age estimates. All analyses were conducted separately for each gender and response option (restricted versus unrestricted) condition. Note that for the groups in which the absolute estimation task responses were restricted to the age range of the faces presented (groups 1 and 2), except for sign (+/-), bias and accuracy values are identical for 7 year old and 11 year old face estimates. This is because there is no possibility of underestimating 7 year olds or over-estimating 11 year olds due to the restriction of age estimate responses to ages 7 through 11.

In cases where assumptions of sphericity were violated, Greenhouse Geisser corrected *F* and *p* values were used for all repeated-measures statistical tests.

4.4 Results

4.4.1 Mean age estimates. The mean age estimates for the restricted and unrestricted conditions are presented in Figure 4.2. In both the restricted and unrestricted condition it is clear that participants' mean estimates of age increase as actual age increases, and this is true regardless of gender or viewing angle. However, the range of estimates is compressed, particularly for the restricted condition. In order to explore this more fully, analyses will be on bias and accuracy.

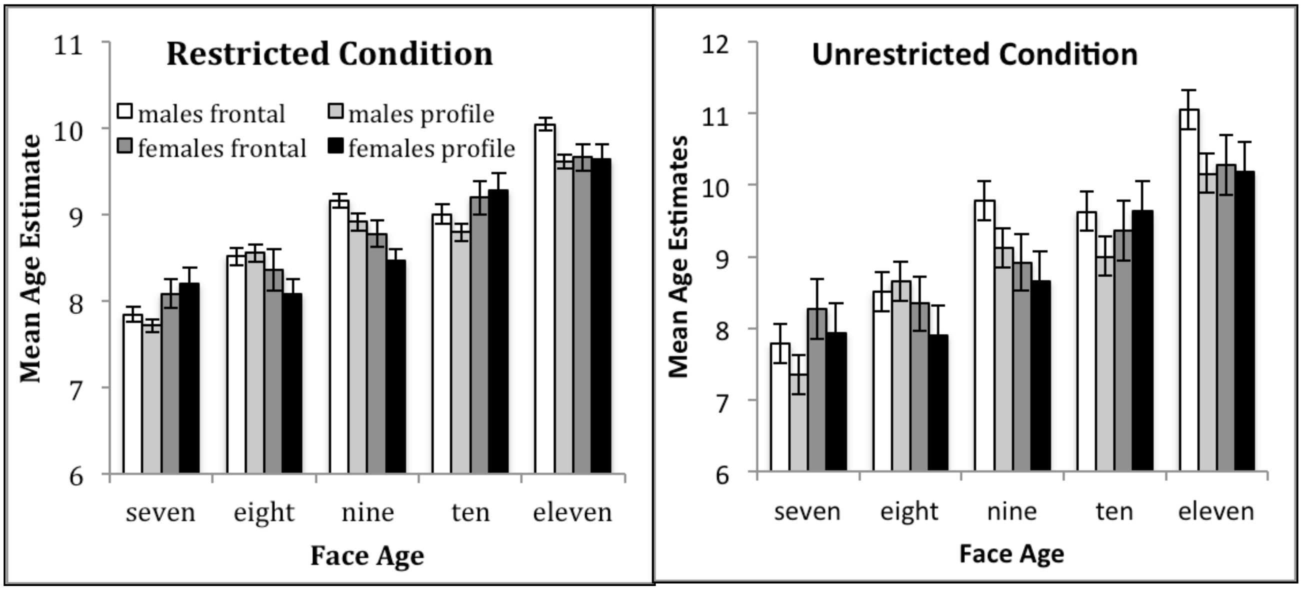


Figure 4.2: Mean age estimates for male (solid) and female (striped) faces in frontal (light grey) and profile (dark grey) view. Left panel shows restricted response condition and right panel shows unrestricted condition. Error bars show standard errors.

4.4.2 Bias.

4.4.2.1 Restricted response options (Groups 1 and 2). We examined the influence of age and viewing angle on estimate bias (directional error) .The data for male faces are presented in the left panel of Figure 4.3. A 2x5 repeated-measures ANOVA was conducted and revealed a significant main effect of age, ($F_{(4, 120)}$ = 263.18, p < .001). There is a clear tendency to over-estimate the age of the younger faces and under-estimate the age of the older faces.

The main effect of viewing angle was also significant, ($F_{(1, 30)}$ = 10.960, p = .002), as was interaction between viewing angle and age, ($F_{(4, 120)}$ = 4.597, p =.002). Faces in frontal view were estimated to look older than faces in profile view, and this effect is particularly evident in the older faces. Bonferroni post hoc analysis (adjusted p value = .01) revealed that bias values differed significantly between frontal and profile view for the oldest age group (11 year olds) ($t_{(30)}$= 4.258, p < .001). Additionally, the differences in estimate bias between frontal and profile in 9 year olds was approaching significance ($t_{(30)} = 2.716$, $p = .011$).

4.4.2.2 Unrestricted response options (Group 3). We also looked at the effects of viewing angle and age on estimate bias when participant response options were not restricted to the age range of the faces and these data are presented in the right panel of Figure 4.3. As seen in the restricted case, a 2 x 5 repeated measures ANOVA revealed significant main effects of age, ($F_{(4, 76)} = 25.929$, $p < .001$) and viewing angle ($F_{(1, 19)} = 31.547$, $p < .001$), and a significant interaction between viewing angle and age ($F_{(4, 76)} = 5.785$, $p < .001$). The results were similar to those observed when the response options were restricted, in that participants continued to over-estimate the ages of younger children, and underestimate the ages of older children. The responses were not as compressed as they were when the response options were restricted, but the pattern did persist. This is important, because it establishes that although the restriction of the response options in the previous task may have contributed to the strength of this compression effect, it does not completely explain it. Also similar to the previous experiment, the age groups in which the viewing angles differ significantly are the older age groups. Bonferroni post-hoc analysis (adjusted p value = .01) revealed that nine, ten, and eleven year old face estimates all differed significantly between viewing angles ($t_{(30)} = 3.884$, $p = .001$; $t_{(30)} = 4.184$, $p = .001$; $t_{(30)} = 6.008$, $p < .001$), with bias score values for faces in these age groups being more negative for faces in profile view than faces in frontal view. Alternatively, this can be thought of as profile view faces looking "younger" than frontal view faces.

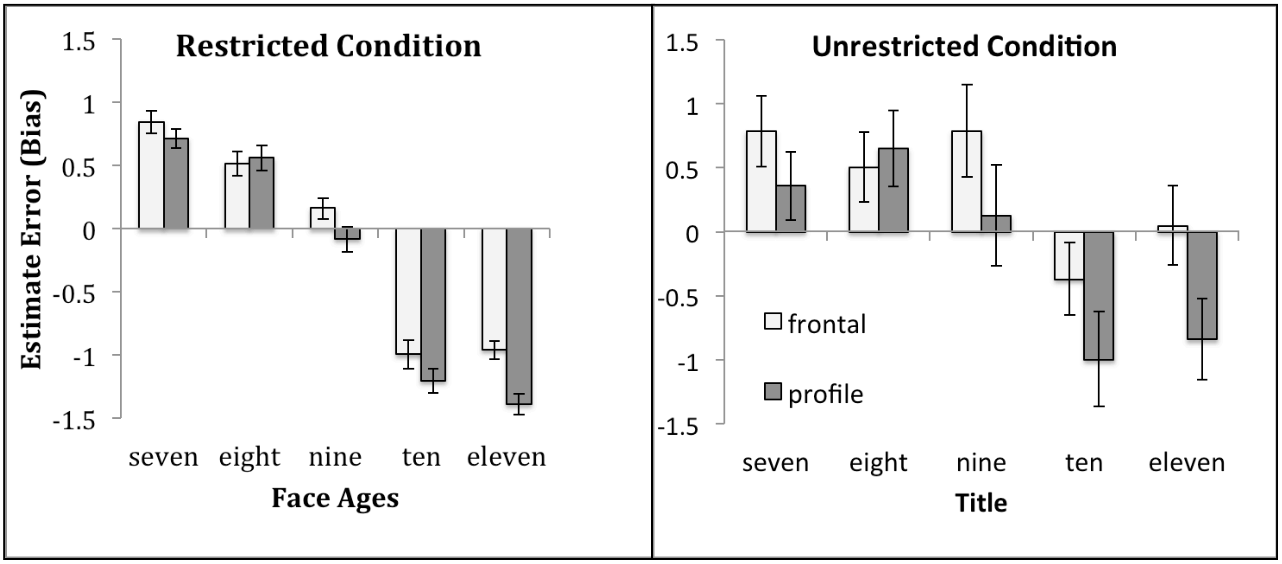


Figure 4.3: Mean deviation score (bias) for each face stimulus age group (7 years through 11 years) separated by face viewing angle (frontal view white, profile view grey). Male face stimuli. Restricted Condition presented in the left panel, Unrestricted condition presented in the right panel. Positive scores indicate over-estimates, negative scores indicate under-estimates. Error bars show standard errors.

4.4.3 Accuracy.

4.4.3.1 Restricted response group. The accuracy data are presented in left panel of Figure 4.4. Estimates were generally quite accurate, with the greatest mean absolute deviation being less than 1.5 years; however, this is not surprising given response options were restricted. The main effect of 2 x 5 repeated-measures ANOVA for age was significant, ($F_{(4, 120)} = 14.477$, $p < .001$), with estimate error generally increasing with increasing face age. The main effect of viewing angle, ($F_{(1, 30)} = 10.276$, $p = .003$), was also significant, as was the interaction between viewing angle and age ($F_{(4, 120)} = 6.999$, $p < .001$). The significant difference between frontal and profile absolute errors in 11 year old faces found with Bonferroni post-hoc analysis explains this interaction ($t_{(30)} = -4.258$ $p < .001$). The difference between viewing angles for nine year olds also approached significance ($t_{(30)} = -2.670$, $p = .012$). Overall, faces in frontal view were generally estimated more accurately than faces in profile view.

4.4.3.2 Unrestricted response options. A 2 x 5 repeated-measures ANOVA was conducted and the results are presented in Figure 4.4 in the right panel. There was not a significant main effect of age ($F_{(4,76)}$=2.045, $p = .096$). There was a significant main effect of viewing angle, ($F_{(1, 19)} = 6.255$, $p = .022$), in that participants generally tended to make more accurate estimates when faces were in frontal view than in profile view (Figure 4.4, right panel). The age by viewing angle interaction was not significant ($F_{(4, 76)}$=1.507, $p = .209$).

Estimates were markedly less accurate when response options were no longer restricted. Absolute deviation scores ranged from about 1.5 years to 2 years when response options were not restricted, and between 0.8 years and 1.4 years when response options were restricted.

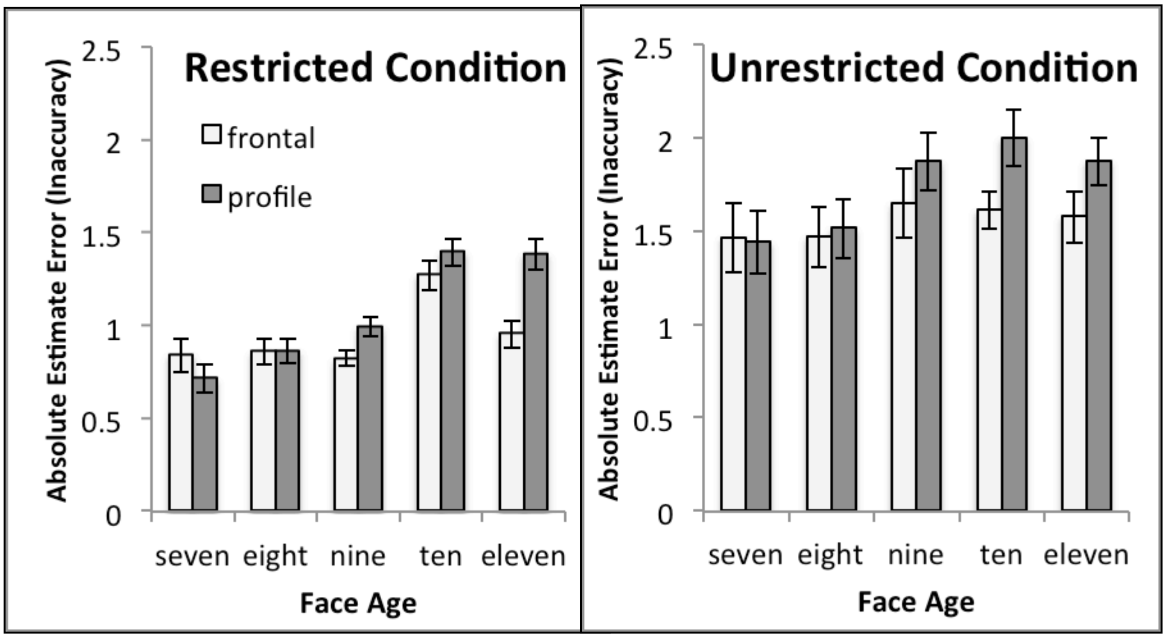


Figure 4.4: Mean absolute deviation score (inaccuracy) for each face stimulus age group, separated by viewing angle (frontal view white, profile view grey). Male face stimuli. Restricted Condition shown in left panel, Unrestricted condition shown in right panel. Higher estimate error indicates less accurate estimates. Error bars show standard errors.

4.4.4 Female Faces.

4.4.4.1 Bias. For female faces, we tested the effects of age and viewing angle on bias in both restricted and unrestricted response conditions. The main effect of age was significant in both the restricted ($F_{(4, 44)}$=65.195, p<.001) and unrestricted conditions ($F_{(4,76)}$=20.970, p<.001). The main effect of viewing angle was not significant for either restricted ($F_{(1,11)}$=.491, p=.498) or unrestricted condition ($F_{(1,19)}$=2.238, p=.151), nor was the interaction between age and viewing angle for either the restricted ($F_{(4,44)}$=1.155, p=.344) or the unrestricted condition ($F_{(4,76)}$=1.487, p=.215). Data are presented in Figure 4.5, collapsed across viewing angle.

4.4.4.2 Accuracy. The accuracy data are presented in figure 4.5 in the left panel. The main effect of age was significant for both the restricted ($F_{(4, 44)}$=3.724, p=.011) and the unrestricted condition ($F_{(4, 76)}$=3.477, p=.012). The main effect of viewing angle was not significant for either the restricted ($F_{(1,11)}$=.061, p=.809) or the unrestricted condition ($F_{(1, 19)}$=.814, p=.378. The interaction between age and viewing angle was not significant for either the restricted ($F_{(4, 44)}$=1.156, p=.343) or the unrestricted condition ($F_{(4, 76)}$=.428, p=.322).

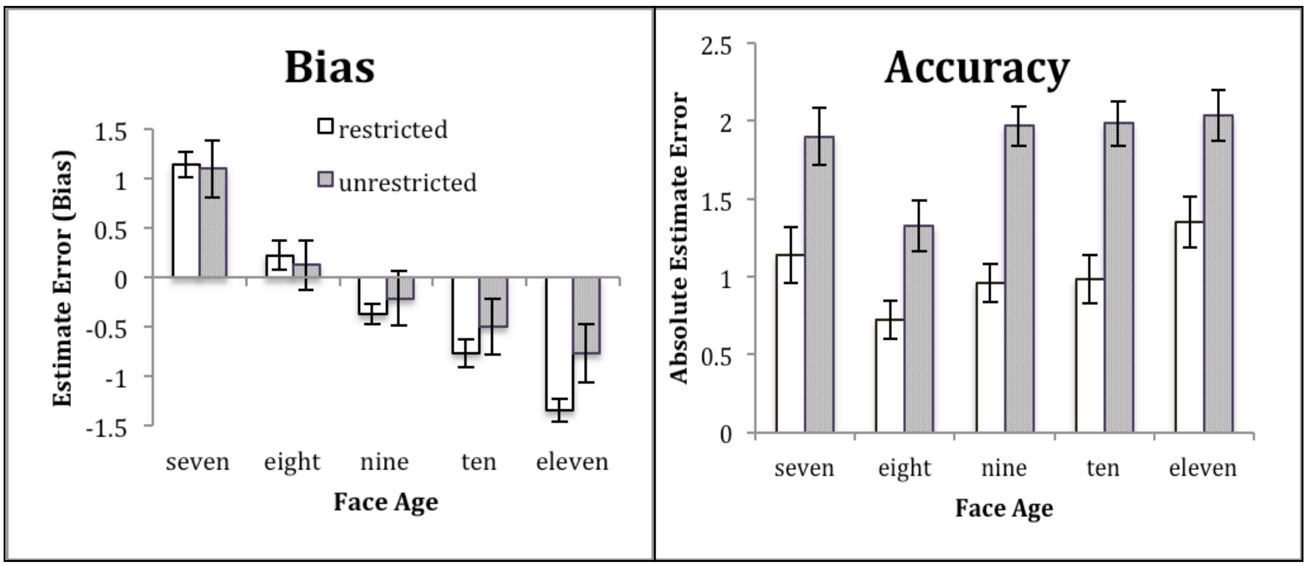


Figure 4.5: Left panel shows estimate bias for female faces separated by response restriction condition, collapsed across viewing angle. Right panel shows estimate accuracy for female faces separated by response restriction condition, collapsed across viewing angle. Error bars show standard errors.

4.4.5 Gender of Face Stimuli. Although the number of female faces in the database was limited, we did conduct preliminary analyses comparing male and female face stimuli on mean age estimates in both restricted and unrestricted response conditions. Because of the limited number of female faces, these results should be interpreted with caution. Face age was significant in all conditions ($p < .001$ in all cases, see Table 4.1). The main effect of face gender was not significant in any case. The interaction between face age and face gender was significant for a number of cases, but Bonferroni post-hoc comparisons showed no consistent pattern and age comparisons that were significant differed across test conditions (Table 4.1). Because there was no main effect of face gender observed, and the interactions observed were inconsistent, it is likely that the interaction effects observed are attributable to the small number of female faces available in the sample and individual differences in the face stimuli.

Table 4.1 Gender Comparison Results

Frontal				Post-Hoc Comparisons				
Response condition	Face Gender	Face Age	Gender * Age	7 year olds	8 year olds	9 year olds	10 year olds	11 year olds
Restricted	$F_{(1,11)}$=2.021, p=.183	$F_{(4,44)}$=64.618, p<.001	$F_{(4,44)}$=1.914, p=.125	NA	NA	NA	NA	NA
Unrestricted	$F_{(1,19)}$=2.063, p=.167	$F_{(4,76)}$=76.250, p<.001	$F_{(4,76)}$=6.537, p<.001	ns	ns	ns	ns	**M>F***
Profile								
Restricted	$F_{(1,11)}$=.535 p=.480	$F_{(4,44)}$=52.776 p<.001	$F_{(4,44)}$=6.917 p<.001	ns	ns	ns	ns	ns
Unrestricted	$F_{(1,19)}$=.000, p=.993	$F_{(4,76)}$=54.035, p<.001	$F_{(4,76)}$=8.162, p<.001	ns	**M>F***	ns	**F>M***	ns

* In cases where post-hoc comparison was significant, direction of difference (judged significantly older) is reported in the right panel. Bonferroni post hoc critical p value when corrected for multiple comparisons was p<.01.

4.4.6 Contact with Children Interview Data. It has previously been suggested that there is relationship between the amount of contact a person has had with children in a particular age group and their age estimation performance for that group (Kohichi (2000). We wished to examine this effect for estimation of the age of children aged 7–11. After completing the experiment, all participants were interviewed to assess the amount of contact they had had with children over the past 5 years (questions adapted from Kuefner et al., 2008). Because we were working with undergraduate participants, most of them did not have a lot of regular contact with children, so we were not able to examine the effect of contact on age estimation. Only 5 of the 128 participants interviewed had children of their own and only 20 reported having job or volunteer experience within the past 5 years that put them in contact with children in this age range at least once a week for longer than 6 months. It would be necessary to directly recruit participants with extensive contact with 7–11 year olds (e.g., primary school teachers or pediatric nurses) in order to investigate this contact hypothesis.

4.4.7 Subjective Reports on the Estimation Process. Additionally, at the conclusion of each experiment, participants were asked what sort of cues, strategies, or techniques they used to help them to decide the age of the faces, or which of two faces was older. Response data from participants in Group 1 was coded into different "cue" types and sub-categories: non-eye features (sub category examples: "softer features", "nose", "small features" "cheek bones"), overall facial structure ("round face", "longer face", "adult like"), expression/emotions (e.g., "shy", "angry/aggressive", "smirking/smiling/happy"), eyes ("big/baby eyes", "bags/circles", "eyebrows"), vague references to size (“taller”, “small head”), skin (“rough/exposed/wrinkled”, and “other” (“ears/hat”, “posture”). Surprisingly, the type of cue reported most often was the subtle variations in facial expression between the faces, although all stimuli were taken from the neutral expression category in the database. Most often, participants reported that faces that were “smiling” or “smirking” looked younger, while faces that looked “annoyed”, “angry” or “sad” were reported to look older. Participants also often reported that the eyes were helpful in deciding the age of a face, with “big” or “baby eyes” suggesting a younger child, darkness or under-eye circles suggesting an older child. Surprisingly, participants also reported using the skin to inform their estimate, with “rough”, “exposed”, “experienced”, “laugh lines” or “wrinkles” being a clue that the child was older.

4.4.8 Size of the Head in the Stimulus Image. Because there was some variability in camera distance when photographing children for inclusion in the database, we tested to make sure there were no differences in the size of the heads between age groups that might be influencing estimates, for example “bigger” heads (closer photograph) looking older than “smaller” heads (more distant photograph). We measured both the inter-pupillary distance (IPD), and the distance between the mid-point between the pupils and the chin (Length) and compared these measurements between the 5 age groups. There was no significant difference between the age groups on either IPD ($F_{(4, 25)}$ = .487, p=.745) or Length ($F_{(4, 25)}$ = .788, p=.544). Additionally, there was no significant correlation between mean estimated age for the face stimuli and either the IPD ($r_{(29)}$ =

.005, p = .981), or Length ($r_{(29)}$ = .074, p = .699). It is unlikely that the size of the face in the image frame is influencing age estimates.

4.5 Discussion

Overall, we have shown that young adults are able to make fairly accurate and unbiased age estimates for all five age groups and in both viewing angles even when they were given no age range for the faces. While previous studies found consistent over-estimates for children's faces even older than those included in our study, we found that the age of the older faces in our experiment were under-estimated, suggesting that the over-estimation effect observed in other experiments may be an effect of the entire range of faces included in the study, and estimates shifting towards the mean age of the all of the faces included in the experiment.

The viewing angle effect observed in the relative judgment data with male faces was also seen in the absolute estimation data, in that there was often an advantage for frontal views over profile views. This effect was not observed when observiers viewed the female faces.

While we have shown that people are consistently able to provide accurate and unbiased age estimates, we still do not know how they are doing it. In the next experiment, we investigate one possible influence- facial expression.

Chapter 5 – Effect of Facial Expression on Age Estimation

5.1 Introduction

As discussed in the previous chapter, in order to get a better idea of what factors people may have been using to inform their age estimates, at the end of each experimental session the researcher asked the participant to describe what sorts of "cues, strategies or techniques" they thought they used to help them choose the age for each face, or to decide which of two faces was older. Surprisingly, the most frequently reported strategy reported by participants was using subtle differences in the facial expressions the children were making in the photographs. Most frequently, participants reported that children who were smirking or smiling looked younger, while children who looked more angry or annoyed looked older.

To our knowledge, only one publication (Voekle et al., 2012) has investigated the effects of facial expression on human age estimates and it did not include children's faces as stimuli. In Voekle's study, participants fell into one of three age groups: 20–31, 44–55, and 70–80. Face stimuli also fell into three age groups: 19–31, 39–55, and 69–80. Both bias and accuracy were assessed for happy, angry, disgusted, fearful, sad, and neutral faces. They found that accuracy was best for neutral facial expressions; however neutral and happy faces were also the most biased (under-estimated).

We decided to examine the effect of facial expression on age estimates for children's faces. Based on our anecdotal evidence from participants and from Voekle et al's (2012) findings with adult faces, we predicted that children with happy faces would look younger than angry faces, which was also consistent with Voekle et al's findings.

Angry faces, on the other hand, were expected to look older than happy faces based on participant reports. Neutral faces were expected to fall in between angry and happy estimates.

We repeated this experiment twice. In the first experiment, faces were presented in blocks by expression and, in the other, expressions were interleaved.

5.2 Method

The methods used were the same as in the absolute task (chapter 4) except as noted below. Participant estimates were not restricted and they were not provided with the age range of the faces. Happy, angry and neutral faces versions of the same 7–11 year old faces were included in the experiment. Participants were presented with an image of a face and asked to estimate the age. A total of 76 participants (Groups 4 and 5 in Table 1; males = 26) completed this absolute age estimation task, divided into 2 designs. For the first group (Blocked design) (N=35, males = 16, mean age = 20.5) the task was blocked by expression and block order was randomized over participants. The second group (Interleaved) (N = 41, males = 10, mean age = 21.5) completed the task with all three facial expressions interleaved within the same run.

Subjects were tested in 180 trials: six individual faces by two viewing angles by three expressions for each of five age conditions. Female faces were not included in this experiment due to the low number of stimulus faces available in the Dartmouth database.

5.2.1 Analysis

5.2.1.1 Missing data. In the Blocked condition, missing and unusual (e.g., estimate of 66 years) data were replaced as described in chapter 4. Two participant were removed from the analysis because more than one data point was missing for the same face/viewing angle/expression group. For the Interleaved condition, the program was re-written so that if estimates fell outside of the 1–20 age range, contained letters, or were left empty, participants were prompted to re-enter their estimate, so we no longer had missing data.

5.2.1.2 Statistical analysis. Mean age estimate, mean estimate error (bias), and mean absolute estimate error (accuracy) were calculated for each age, expression, and viewing angle combination. Data for the Blocked group were analyzed using a mixed-ANOVA (viewing angle, age, and expression within group factors and task order as a between-groups factor). For the Interleaved group, data were analyzed using 2 x 5 x 3 (viewing angle x face age x expression) repeated-measures ANOVA. In cases where assumptions of sphericity were violated, Greenhouse Geisser corrected F and p values were used for all repeated-measures statistical tests.

5.3 Results

The main effects of age and viewing angle observed for this task were the same as chapter 4 for the most part, and therefore will not be reported unless they are different from the effects observed in chapter 4.

5.3.1 Blocked Design

5.3.1.1 Bias. The main effect of expression was not significant ($F_{(2,60)} = 1.758$, p =.181). The interaction between expression and age was significant ($F_{(8, 240)} = 5.680$, $p <$ .001). As can be seen in Figure 5.1 in the left panel, for happy and neutral faces the tendency to over-estimate the age of younger faces seen previously with neutral faces was apparent again, but for angry faces, age was under-estimated across all age groups. Bonferroni post-hoc (critical p-value when adjusted for multiple comparisons = .005) analysis revealed differences that reached or approached significance between age estimates for angry and happy faces and angry versus neutral faces at 7 ($t_{(32)} = -3.137$, $p =$ .004; $t_{(32)} = -2.870$, p=.007) and 9 ($t_{(32)} = -2.976$, p=.006; $t_{(32)} = -2.857$, p=.007) years. No comparisons between happy and neutral faces were significant. Contrary to the predictions made based on participant self-reports, where significant differences were observed, angry faces were judged to be younger than neutral faces, not older; there was no evidence that happy faces were judged to be younger than neutral faces.

In the right panel of Figure 5.1, the same data are presented collapsed across age and separated by task order. People who saw angry faces first estimated ages to be younger than people who saw happy or neutral faces first. This was the case for all expressions, but particularly for angry faces. This is reflected in the borderline significant interaction between expression and task order (($F_{(4, 120)} = 2.362$, $p = .063$). This suggests that the expression by age interaction observed on this task is either due to being exposed to angry faces first, somehow biasing responses for the remained of the experiment or alternatively, due to some idiosyncratic characteristics of the group of participants who saw angry faces first.

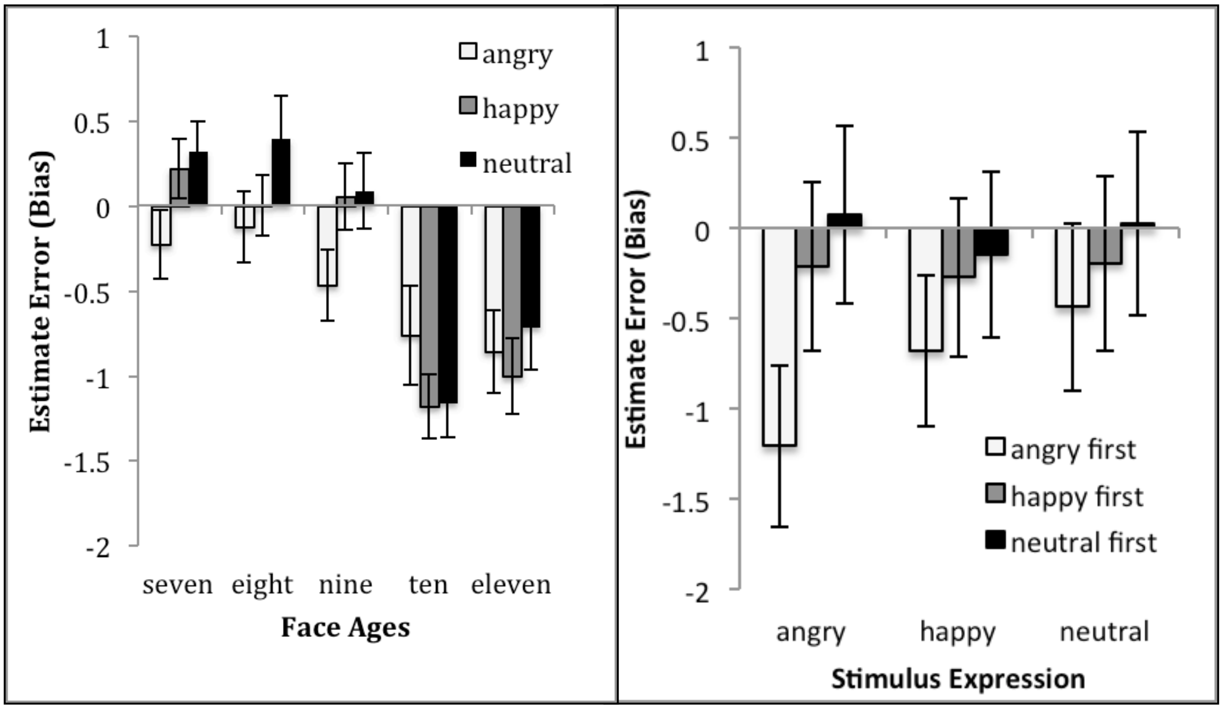


Figure 5.1: Left panel: Mean deviation score (bias) for each face stimulus age group (7 years through 11 years) separated by expression, but collapsed across viewing angle. Right panel: Estimate error for angry, happy and neutral face stimuli separated by block order. White bars are estimates from participants who saw angry faces first, grey are estimates from participants who saw happy faces first, and black are estimates from participants who saw neutral faces first. Male face stimuli. Positive scores indicate over-estimates, negative scores indicate under-estimates. Participant response options were unrestricted. Error bars are standard errors.

5.3.1.2 Accuracy. A 2 x 5 x 3 x 3 (viewing angle x face age x expression x task order) mixed-methods ANOVA was used to analyze the within-subjects effects of age, viewing angle, and expression and between subjects effect of task order. There was not a significant main effect of expression ($F_{(2, 64)}$ = .345, p = .710). Although the interaction between expression and age reached statistical significance, ($F_{(8, 256)}$ = 2.106, p = .036) (See Figure 5.2 left panel), there was no consistent pattern in the relative ordering of stimulus expressions across age. No individual post-hoc comparisons were statistically significant.

The interaction between expression and task order was also significant ($F_{(4, 120)}$ = 2.826, p = .032). Just as the participants who saw angry faces first showed greater bias (see above), they also showed less accurate responding, at least for angry and neutral faces, again suggesting that this may reflect either idiosyncratic characteristics of these participants (Figure 5.2 right panel).

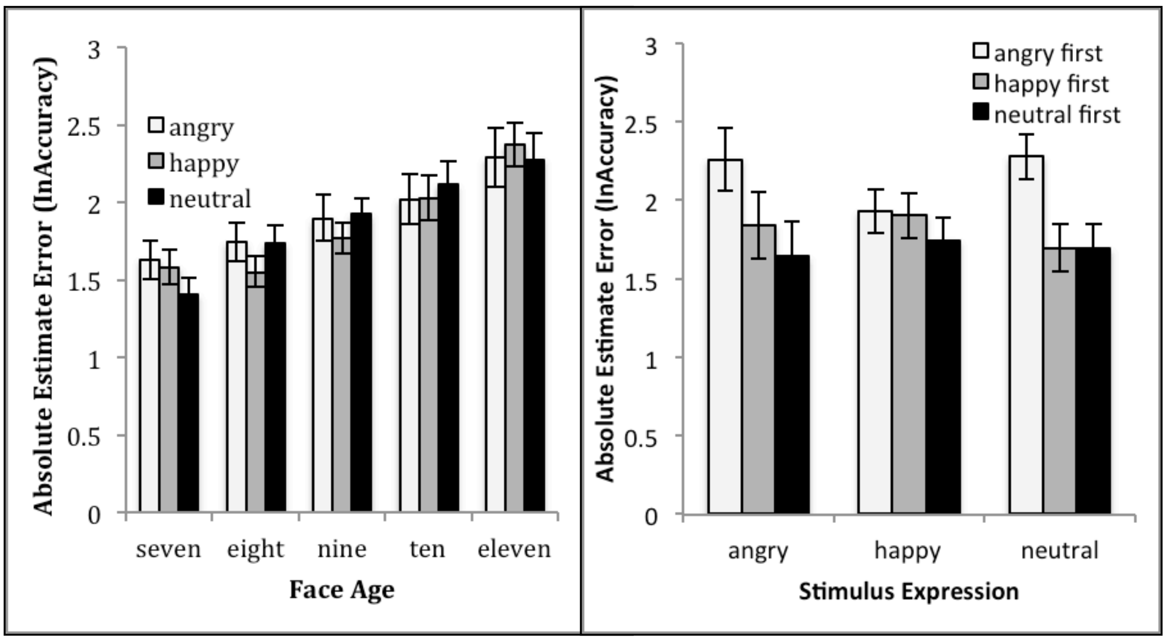
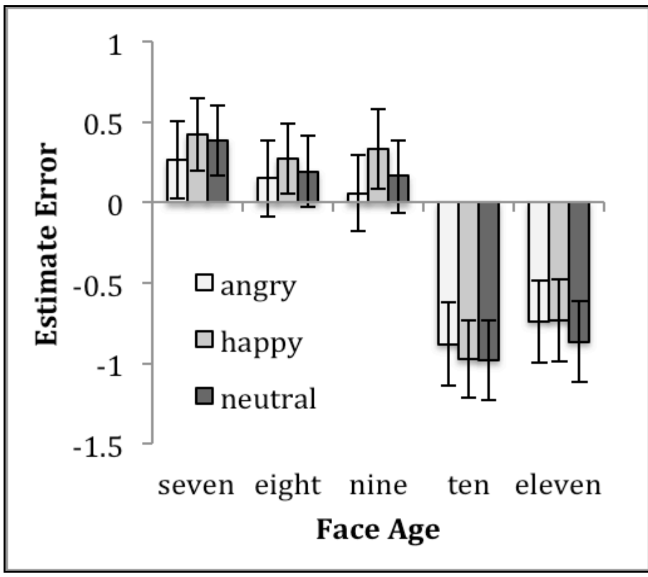


Figure 5.2. Left panel: Mean absolute deviation score (inaccuracy) for each face stimulus age group, separated by expression and collapsed across viewing angle. Right panel: Absolute estimate error for angry, happy and neutral face stimuli separated by block order. Light grey are estimates from participants who saw angry faces first, dark grey are estimates from participants who saw happy faces first, and black are estimates from participants who saw neutral faces first. Male face stimuli. (Higher estimate error indicates less accurate estimate). Participant response options were unrestricted. Error bars are standard errors.

5.3.2 Interleaved group - Effects of facial expression – Within-subjects

5.3.2.1 Bias. There was not a significant main effect of expression, ($F_{(2, 80)}$ = 1.463, p = .239) (see Figure 5.3 left panel). The interaction between viewing angle and expression was significant, ($F_{(8, 320)}$ = 9.134, p < .001) (see Figure 5.3 right panel). No other interactions involving expression were significant.



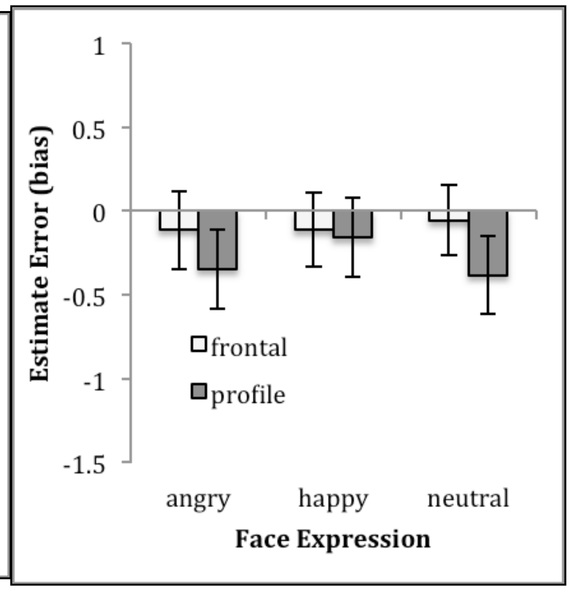


Figure 5.3: Left pane: Mean deviation score (bias) for each face stimulus age group (7 years through 11 years) separated by expression, but collapsed across viewing angle. Right panel: Mean deviation score (bias) for each face stimulus expression, separated by viewing angle, but collapsed across age. Male face stimuli. Positive scores indicate over-estimates, negative scores indicate under-estimates. Participant response options were unrestricted. Error bars show standard errors.

5.3.2.2 Accuracy. There was no effect of expression ($F_{(2,80)}$ = .269, p = .765), nor were any of the interactions involving expression significant (see Figure 5.4).

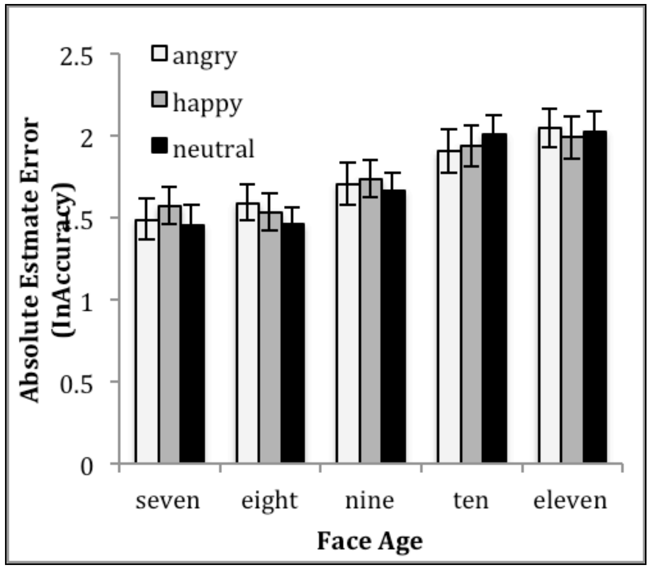


Figure 5.4: Mean absolute deviation score (inaccuracy) for each age group (7, 8, 9, 10, and 11 years) separated by expression, but collapsed across viewing angle. Male face stimuli. Higher estimate errors indicate less-accurate estimates. Participant response options were unrestricted. Error bars show standard errors.

5.4 Discussion

Our predictions regarding expression based on past research and anecdotal reports were not supported. Although we did find some interactions between face age and expression, they were not what we expected. Angry faces, which participants reported tended to look older, were actually estimated as younger than the happy and neutral faces, particularly in the youngest faces on the Blocked task. This effect was observed in the blocked design, however was not replicated in the interleaved experiment. It appears that only participants in the group who saw angry faces first had this tendency to underestimate the age of angry faces. We cannot resolve from the current data whether this is due to first exposure to angry faces biasing age judgments for other expressions, or due to particular characteristics of these participants. This will only be resolved in replication of this experiment.

Overall, the effects of age were consistent between this experiment and our findings in chapter 4; participants continued to over-estimate younger faces and underestimate older faces most of the time, although there were some discrepancies in the Blocked task (angry faces were underestimated across all age groups).

Although the main effect of viewing angle was not observed in all analyses, the interaction between age and viewing angle was significant in all but one case, and the interaction was the same as that observed in chapter 4, with faces in profile view generally appearing younger than those in frontal view.

Chapter 6 – General Discussion

Estimating age is an important ability we all use on a daily basis in order to interact in an age-appropriate way with the people around us. In this thesis, we examined the ability of university undergraduates to estimate age from children's faces in the 7 to 11 year age range. We assessed this ability using both a relative and an absolute age judgment task with both male and female face stimuli. We looked at both the bias and accuracy of absolute age estimates. The influence of viewing angle on age estimates was also examined. Finally, we looked at one possible factor influencing age estimates– facial expression.

6.1 Relative Estimates

Participants were able to correctly choose the older of two children's faces as little as 2 years apart for both male and female faces at rates significantly better than chance. As expected, as the age distance between the faces being compared increased, performance improved. For faces 4 years apart, performance was almost perfect (95% of the time for male faces in frontal view). Pittenger and Shaw (1975b) collected annual school photographs of the same individuals in grades 7 – 12 and had participants sort faces of both the same individual and different individuals by age. They found that the ranking was most accurate for faces around 14–15 years old, even when the faces came from different individuals. They attributed this to the fact that physical growth changes occur most rapidly in this age group due to the onset of puberty. Our results were consistent with those of Pittenger and Shaw (1975b) in that we found that participants were able to rank faces by age quite well, even when the faces were younger, suggesting that even more gradual growth changes are sufficient for making accurate relative age judgments.

6.2 Absolute Age Estimates

We also found that undergraduates were able to provide age estimates that were quite accurate for children's faces between the ages of 7 and 11 years, regardless of gender, expression, and viewing angle. Estimated age differed significantly as a function of actual age in all of our experiments, increasing as actual age increased. Although the estimates were generally quite close to the actual face age, there was a clear tendency for the estimate errors to be biased in the direction of the mean age of the stimulus faces; that is, participants tended to over-estimate the age of youngest faces and under-estimate the age of oldest faces on all estimation tasks. Even when participants were not provided with an age range, this effect was still present.

Previous studies have reported that the age of children and young adults is generally over-estimated and the age of older adults is generally under-estimated (Vestlund et al., 2009; Fahsing, et al., 2004; Moyse & Bredart, 2012). This effect is often attributed to regression towards the mean of the entire range of possible face ages. For example, Fahsing et al., (2004) examined archival data from eye-witness testimonies and compared reports to the actual age of perpetrators. They found that faces 16–21 years old were over-estimated and faces 40–50 years old under-estimated in age. Vestlund et al. (2009) found that target faces age 15–19, 20–24, and 25–30 were all over-estimated in one experiment, and in a second experiment

15–24 year olds and 34–46 year olds were over-estimated, while 56–64 year olds were underestimated. Our findings were not consistent with this interpretation that estimates shift towards the mean of the entire range of possible face ages. Our findings are also not consistent with an own-age anchor effect. An own-age anchor effect is when estimates shift towards the participant's own age. If this were the case, all face ages would be over-estimated since all of our participants were older than all of the face stimuli evaluated. Our results are more in line with the findings of Willner and Rowe (2001), who found that 13 and 16 year olds were over estimated and 20 and 22 year olds were under-estimated; the estimates are biased in the direction of the mean of the age of the faces included in the experiment, rather than the population.

6.3 Viewing Angle Effects

A viewing angle effect was observed across all relative estimation experiments for male faces. Participants were better able to rank faces by age when they were in frontal view than when they were in profile view. An effect of viewing angle was also observed in the absolute estimation tasks, such that estimates were less accurate and more biased in profile view than frontal view, at least in the older three age groups. There are no prior studies to our knowledge that have examined the effect of viewing angle on age estimation. However, a number of studies have investigated the effect of viewing angle on face recognition. A frontal view advantage was found in male faces that was consistent with the findings of Bruce, Valentine, and Baddeley (1987), Hill and Bruce (1996), and McKone (2009) in the face recognition literature. Bruce et al. (1987) found that participants were slower to respond and made more mistakes when asked to decide if two sequentially presented faces were the same person when either of the faces was in profile view. Performance was particularly bad if both faces were presented in profile view. Hill and Bruce (1996) found that even in the recognition of visually degraded (but very familiar) faces, performance was markedly poorer in profile view than ¾ or frontal view. McKone (2009) also reported slower performance and less accurate identification of faces in profile view than in frontal or ¾ view. One obvious explanation is that we simply get a lot more exposure to and experience with frontal view faces, particularly in one-on-one interaction. Even in photographs encountered in the media, only 5% are profile views (McKone, 2009). So we would be better at estimating age from frontal view faces simply because we've had more practice doing it. Alternatively, the effect of viewing angle may be related to the differences in information available in frontal view versus profile view. In frontal view, the internal facial features, and their relationship to each other are much more clear, while in profile view, information about the angle of the forehead, and protrusion of the nose and chin becomes available.

6.4 Facial Expression

Our predictions regarding expression were not supported. Participants often reported after completing the experiments in Chapter 3 and 4 that they used subtle differences in nominally neutral facial expression as a cue to inform their age estimates. Children who looked like they were smiling or smirking were subjectively reported looked younger, while those who looked angry looked older. Similar anecdotal reports were also described by Pittenger and Shaw (1975a). When we explicitly tested the effects of distinct facial expressions (angry, happy and neutral) we found no such effect. Angry faces were actually estimated as *younger* than the happy and neutral faces in the youngest faces on the Blocked task (however as stated in Chapter 5, this was likely more an effect of individual differences in that particular group of participants than a meaningful effect of expression). In our interleaved experiment we found no main effect of expression. While it is possible that facial expression may be a cue used by some individuals to inform age estimates, we found no consistent effect of facial expressions in our experiments.

6.5 Other Factors Contributing to Age Estimates

Although participants reported cues related to cardioidal strain only occasionally (small chin, large forehead), these cues still may have informed their estimates. Pseudo-cardioidal strain manipulations have been used in the past (see George & Hole, 1995) with adult and children's faces to examine the influence of this manipulation on age estimates and found that shifting the internal features shifted age estimates in the expected direction. It is possible that even on the fine scale used in our study, cardioidal strain was at least one of the factors informing participant's estimates. However, George and Hole (1995) also tested young adult subjects age estimation from children aged 5-10 with only the internal facial features and found that for young participants, their estimates were no different for the features only than they were for the original faces. If cardioidal strain information alone was informing age estimates, performance should have been poorer with estimates made from only the internal facial features, unless participants are able to somehow extrapolate all of the cardioidal strain information from the relationship between the features alone. So while cardioidal strain may be one factor influencing age estimates for faces in the 7–11 year age range, it seems unlikely that it is the only factor.

Craniofacial growth changes occur rapidly during development (Enlow, 1982). It is possible that changes in craniofacial growth are informing age estimates in this age group. Many of the changes associated with craniofacial development (long nose, jaw and chin development, protruding forehead) are much more visible in profile view, and yet, age discrimination performance was poorer in profile view. However, craniofacial changes visible from the front of the face (relative size of the eyes, relative height of the ears, relationship between feature size and rest of the head) may have informed estimates. Since in the George and Hole (1995) study participants were able to make accurate estimates from the internal features alone, it is likely that some of these changes were informing their estimates. Even children as young as 6 years old are able to make relative age estimates from only internal features (George & Hole, 2000b). Given these findings, there are a number of possible cues that are informing age estimates: First, the relationship between the internal features may be informing estimates. This is known as configural processing. Another second possibility is that the characteristics of some or all of the features themselves, rather than their relationship with one another, are informing estimates (eg., nose changing from a "button nose" to more adult nose. A third, albeit unlikely possibility is that more global qualities such as skin texture and colour are informing estimates. While this is likely highly informative in adult faces, it seems unlikely that this would be a particularly informative cue in children, though it was reported by some participants in post-experiment interviews.

Interestingly, in another study, George and Hole (2000a) found that inverting children's faces had absolutely no effect on age estimation performance. Inverting a face is thought to disrupt configural processing (relationship between the facial features) considerably. In fact, estimates were actually closer to the actual age of the face for inverted faces than they were for unmanipulated faces (George & Hole, 2000a). Based on this finding, it makes it unlikely that the relationship between the features (configural information) alone is informing estimates. Additionally, presenting the faces as negatives (brightness levels reversed) also did not disrupt age estimation accuracy, even though this manipulation would probably make the characteristics and shape of the individual facial features themselves difficult to resolve. Additionally, in our study, the relatively short exposure time would make it difficult for participants to spend a great amount of time scrutinizing the qualities of individual features. So again, it is unlikely that the characteristics of the individual features themselves are the only cue informing estimates.

The most likely explanation is that we are able to make fairly accurate age estimates using any one of these age cues, and in the absence of one, another is sufficient. Indeed, George and Hole (2000a) found that when they both reversed the brightness *and* inverted the face stimuli, performance deteriorated considerably. However, since George and Hole (1995; 2000a) analyzed performance of the age group as a whole rather than examining estimates individual by year, we may find that when we look at performance on this finer scale, we will be better able to see the relative influences of each of these different age cues.

6.6 Limitations

One limitation was that our participants were all university undergraduates. However, given the limited amount of contact most of our participants had with children, they were surprisingly good at estimating age in this age group, contrary to the contact hypothesis. The contact hypothesis posits that people who have a lot of contact with a particular age group (for example, children), will make more accurate age estimates. To our knowledge, only one study has investigated the contact hypothesis in the context of age estimation (Kohichi, 2000). The details of this research are available only in Japanese; however, according to the abstract, the authors found that nursery school teachers were better at estimating the age of young children (1–5 years) while nursing home employees were better at estimating the age of seniors (66–89 years) and accuracy increased as the amount of contact with the respective age group increased. In a future study, recruiting unique populations with very high or very low levels of contact such as mothers or pediatric nurses would be a valuable extension to the studies reported here.

Another limitation of our study was that we did not know the exact ages of the photographed children, only their age in whole years. We were not able to make comparisons between adjacent age groups in our relative judgment task because we could potentially be comparing a 7 year old who almost 8, to an 8 year old who had a birthday very recently. We would be able to test participants' ability to make even finer discriminations (only one year apart) if we had exact ages of the faces.

While we were fortunate to be able to access this database with so many children in different viewing angles and with different facial expressions, a greater number of faces would be an improvement. This database was particularly limited in the number of female faces in the age range of interest to us so we were not able to carry out directly comparable analyses of male and female faces.

6.7 Future Directions

While we have answered the question of whether or not people can make age estimates from children's faces, exactly what aspects of individual features versus their configuration influence these judgments is not yet clear. The next step would be manipulating these factors separately and investigating the relative contribution of each to age estimation from children's faces.

Although we found that participant's estimates for female faces increased with increasing face age, we weren't able to directly compare performance on male and female faces, or their interactions with participant gender due to the differences in the number of male and female faces available. In a future study, it would be interesting to directly compare gender.

Another question that came up while working on this thesis was whether or not gender could be discriminated at all in this age group when cues like hair, jewelry, makeup, and clothing were not available. We have already started investigating undergraduates' ability to discriminate between genders with this set of faces.

6.8 Summary

In this thesis, we addressed the question of whether or not people could estimate age from children's faces in the 7 to 11 year age range. We found that young adults were able to make accurate relative age judgments for both males and females, even in faces as little as two years apart, and that their performance improved as the age differences between the faces being compared increased. They were also able to make accurate absolute age judgments that increased with increasing face age for both male and female faces. We also looked at estimate bias and while estimates were generally low in bias, the bias was in direction of the mean age of the stimuli. Additionally, we found for the first time that there is an effect of viewing angle on age estimates for male faces, in that there is generally an advantage for faces presented in frontal view. Finally, we looked at one possible factor influencing age estimates– facial expression, and found that it was unlikely that facial expression was a primary cue informing age estimates.

References

Adams, L. (2009) *Who is older? Cues to relative age in faces*. Unpublished Honours Thesis, York University, Toronto.

Bruce, V., Valentine, T., & Baddeley, A. (1987). The basis of the 3/4 view advantage in face recognition. *Applied Cognitive Psychology, 1*(2), 109–120.

Burt, D. M., & Perrett, D. (1995). Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information. *Proceedings: Biological Sciences, 259*(1355), 137–143.

“Causes of Aging Skin”. (2010). *American Academy of Dermatology*. Retrieved from http://www.skincarephysicians.com/agingskinnet/basicfacts.html.

Dalrymple, K., Gomez, J., & Duchaine, B. (2013). The Dartmouth Database of Children’s Faces: Acquisition and Validation of a New Faces Stimulus Set. *PLoS ONE,* 8(11).

Dehon, H. & Bredart, S. (2001). An ‘other-race’ effect in age estimation from faces. *Perception, 30,* 1107–1113. doi: 10.1068/p3122.

DiBattista, A. (2008*). Age-dependent changes in face geometry: a psychophysical assessment.* Unpublished Master’s Thesis. York University, Toronto, On.

Enlow, D. H. (1982). *Handbook of Facial Growth*. 2nd Ed. Philadelphia: W. B. Saunders Co.

Fahsing, I. A., Ask, K., & Granhag, P. A. (2004). The man behind the mask: Accuracy and predictors of eyewitness offender descriptions. J*ournal of Applied Psychology*, 89(4), 722–729

Gainotti G., & Marra, C. (2011). Differential contributions of right and left temporo-occipital and anterior temporal lesions to face recognition disorders. *Frontiers in Human Neuroscience*, 5(55), 1–11.

Gauthier, I., Skudlarski, P., Gore, J.C., & Anderson A.W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191–197. doi: 10.1038/72140.

George, P., & Hole, G. (1995). Factors influencing the accuracy of age estimates of unfamiliar faces. *Perception, 24*(9), 1059–1073.

George, P. A., & Hole, G. J. (1998). The influence of feature-based information in the age processing of unfamiliar faces. *Perception, 27*, 295–312.

George, P., & Hole, G. (2000a). The role of spatial and surface cues in the age-processing of unfamiliar faces. *Visual Cognition, 7*(4), 485–509.

George, P., & Hole, G. (2000b). Factors influencing young children's ability to discriminate unfamiliar faces by age. *International Journal of Behavioural Development, 24*(4), 480–491. doi: 10.1080/016502500750038035.

Henss, R. (1991). Perceiving age and attractiveness in facial photographs. *Journal of Applied Social Psychology, 21*(11), 933–946.

Hole, G., & George, P. (2011). Evidence for holistic processing of facial age. *Visual Cognition, 19*(5), 585–615. doi: 10.1080/13506285.2011.562076.

Hill, H., & Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performancec, 22,* 986–1004.

Jones, G., & Smith, P.K. (1984). The eyes have it: Young children's discrimination of age in masked and unmasked facial photographs. *Journal of Experimental Child Psychology, 38,* 328–337.

Kanwisher, N., McDermott, J., & Chun, J. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11), 4302–4311.

Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of Experimental Psychology: Human Perception and Performance, 34*(4), 811–817. doi: http://dx.doi.org/10.1037/0096-1523.34.4.811

Kohichi, S. (2000). Estimating ages from faces: How does experience affect accuracy? *The Japanese Journal of Psychonomic Science, 19*(1), 8–13.

Kotter-Gruhn, D., & Hess, T. M. (2012). So you think you look young? Matching older adults' subjective ages with age estimations provided by younger, middle-aged, and older adults. *International Journal of Behavioural Development,* doi: 10.1177/0165025412454029

Mark, L. S., Shapiro, B. A., & Shaw, R. E. (1983). Structural support for the perception of growth. *Journal of Experimental Psychology: Human Perception and Performance, 12*(2), 149–159.

Mark, L. S., & Todd, J. T. (1985). Describing perceptual information about human growth in terms of geometric invariants. *Perception & Psychophysics 37*(3), 249–256.

"Mature Skin". (2013). *American Academy of Dermatology*. Retrieved from http://www.aad.org/media-resources/stats-and-facts/prevention-and-care/mature-skin#.UbiUKjnm5iA

McKone, E. (2009). Configural Processing and face viewpoint. *Journal of Experimental Psychology: Human Perception and Performance, 34*(2), 310–327. doi: 10.1037/0096-1523.34.310.

Montepare, J. M. & McArthur, L. Z. (1986). The influence of facial characteristics on children's age perceptions. *Journal of experimental child psychology, 42*, 303–314.

Moyse, E., & Bredart, S. (2012). An own-age bias in age estimation of faces. *Revue européenne de psychologie appliquée, 62*, 3–7.

Nkengne, A., Bertin, C., Stamatas, G., Giron, A., Rossi, A., Issachar, N., & Fertil, B. (2007). Influence of facial skin attributes on the perceived age of Caucasian women. *Journal of the European Academy of Dermatology and Venerology, 22*, 982-991. doi: 10.1111/j.1468-3083.2008.02698.x.

Pittenger, J. B., Shaw, R. E., & Mark, L. S. (1979). Perceptional information for the age level of faces as a higher order invariant of growth. *Journal of Experimental Psychology: Human Perception and Performance, 5*(3), 478–493.

Pittenger, J. B., & Shaw, R. E. (1975a). Aging faces as vical-elastic events: Implications for a theory of nonrigid shape perception. *Journal of Experimental Psychology: Human perception and performance, 1*(4), 374–382.

Pittenger, J. B., & Shaw, R. E. (1975b). Perception of relative and absolute age in facial photographs. *Perception and Psychophysics, 18*(2), 137-143.

Ramanathan, N., & Chellappa, R. (2006). Modeling age progression in young faces. *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

Rossion, B., Caldara, R., Seghier, M., Schuller, A-M., Lazeyras, F., & Mayer, E., (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *A Journal of Neurology, 126*(11), 2381–2395

Sorqvist, P., & Eriksson, M. (2007). Effects of training on age estimation. *Applied Cognitive Psychology, 21,* 131–135. doi: 10.1002/acp.1271.

Thompson D. W. (1917). *On growth and form*. Cambridge: Cambridge University Press.

Todd, J., Mark, L., Shaw, R., & Pittenger, J. (1980). The perception of human growth. *Scientific American, 242*(2), 132–144.

Voelkle, M. C., Ebner, N. C., Lindernberger, U., & Riediger, M. (2011). Let me guess how old you are: Effects of age, gender, and facial expression on perceptions of age. *Psychology of Aging, 27*(2), 265–277. doi: 10.1037/a0025065.

Veres-Injac, B., & Schwaninger, A. (2009). The time course of processing external and internal features of unfamiliar faces. *Psychological Research, 73*, 43–53. doi: 10.1007/s00426-008-0147-5.

Vestlund, J., Langeborg, L., Sorqvist, P., & Eriksson, M. (2009). Exerts in age estimation. *Scandinavian Journal of Psychology, 50*, 301–307. doi: 10.1111/j.1467-9450.2009.00726.x.

Willner, P., & Rowe, G. (2001). Alcohol servers' estimates of young perople's ages. *Drugs: education, prevention and policy, 8*(4), 375–383. doi: 10.1080/09687/63001001 9299.

Appendix A

Contact with children Interview

1. Do you have children?

Current Ages:

a. A few hours a year
b. A few hours a month
c. A few hours a week
d. Daily

2. Do you have any siblings tgat you've had regular contact with within the past five years?

Current ages:

a. A few hours a year
b. A few hours a month
c. A few hours a week
d. Daily

3. Do you have any nieces or nephews or other younger relatives you've been in contact with within the past five years?

Current ages?

a. A few hours a year
b. A few hours a month
c. A few hours a week
d. Daily

4. Have you had any regular contact with the children of friends or acquaintances in the past five years?

Current ages?

a. A few hours a year
b. A few hours a month
c. A few hours a week
d. Daily

5. Have you had any job or volunteer experience in the past five years that put you in contact with children regularly?

Age group?

a. A few hours a year
b. A few hours a month
c. A few hours a week
d. Daily

6. Do you watch television shows or movies with children as major characters?

Examples?

a. Very often
b. Occasionally
c. Rarely

7. Do you spend a lot of your free time engaging in activities that involve children?

Current ages?

a. A few hours a year
b. A few hours a month
c. A few hours a week
d. Daily

8. Do you have any other regular contact with children which we have not discussed?

a. A few hours a year
b. A few hours a month
c. A few hours a week
d. Daily

Appendix B

Department of Psychology, Faculty of Health & Centre for Vision Research, York University

Informed Consent – form A

Participant's name and code: ____________________________

Study Title: *Age estimation from children's faces*

Head of the Research Team: Dr. Frances Wilkinson phone #: 736-2100 ext 33184

Person directly in charge of conducting the study: Alex Markham (Graduate student)

Project Funding: NSERC

Description of the study: This study examines the way in which the human visual system analyzes information for age estimation of faces. Participating in this study involves making age judgments about children's faces presented on a computer monitor. An experimental session will be approximately ___ min in length, and only one session will be required to obtain meaningful data for this study. During testing you will be asked to sit in front of a computer monitor and to maintain a fixed distance from the monitor. You can work at your own pace, and the session will be broken into several blocks with breaks between them.

Risks and Benefits*:* This study has no known risks associated with it as it involves brief periods of viewing a computer screen. The study will not benefit you directly; however, information obtained from this study may advance our understanding of the human visual system.

Data Security and Confidentiality: All information derived from the study will be kept confidential. Data, identified by an alphanumeric code, will be stored for a period of 5 years in a locked filing cabinet in the laboratory of Dr. Frances Wilkinson and in a computer files protected by a password. Only members of the laboratory involved in the study will have access to these files. The file containing information linking your name to your data will be stored separately in secured files. In publications resulting from this study, your data will be identified by your initials, unless you specifically request otherwise, in which case a numeric code will be used. Confidentiality will be provided to the fullest extent possible by law.

***Right to withdraw*:** Your participation in this study is voluntary and you have the right to withdraw from this study at any time for any reason. Your decision to stop participating, or to refuse to answer particular questions will not affect your relationship with the researchers, or with York University. Should you decide to

withdraw from the study all hard copy and electronic records of data collected from you will be destroyed/securely erased immediately.

***Compensation**:* Compensation for participation in the study will be 1 URPP credit.

Questions about the Research? If you have any questions about the study, please contact Dr. Frances Wilkinson at (416) 736-2100 ext 33184, franw@yorku.ca.

Ethics Approval: This project has been approved by the Human Participants Review Sub-Committee, York University's Ethics Board and conforms to the standards of the
Canadian Tri-Council Research Ethics guidelines. If you have any questions about this process or about your rights as a research subject, please contact Ms. Alison Collins-Mrakas, Senior Manager & Policy Advisor for the Office of Research Ethics, 5th Floor, York Research Tower, York University (telephone 416-736-5914 or e-mail ore@yorku.ca

****

I have been informed about the nature and procedures of this study, and understand it in full. My participation in this study is entirely voluntary and I know that I may withdraw from the study at any time for any reason, and that I am under no obligation to justify my withdrawal to the experimenters. I am not waiving any of my legal rights by signing this form. My signature below indicates my consent.

__________________________ ____________________________
Signature of participant Date

__________________________ _____________________________
Signature of Investigator