Table of Contents

List of Figures

List of Tables

Abstract

Research has found a stranger's perceived age changes when they appear in a group. Typically, results indicate that a face appears younger when surrounded by younger faces. The mechanisms behind these findings are unknown, though, ensemble coding theory is a proposed explanation. Ensemble coding is an adaptive mechanism where an average across a group of stimuli is taken, allowing larger amounts of information to be processed. If ensemble coding does explain past findings of age estimation is groups, it is expected that a stranger will appear younger when flanked by a larger number of faces and a weaker effect will occur when faces are viewed for a longer period. To examine this, we presented 237 participants with a series of 30 target faces, 10 presented alone, 10 presented alongside 2 younger flanker faces and 10 presented alongside 4 younger flanker faces. Half of the participants viewed each face for 2 seconds and half viewed each face for 4 seconds. Results replicated previous findings. We found that target faces appeared younger when presented alongside younger flanker faces, and the effect increases as the number of flankers increased. The time faces were presented had no impact. Thus, it is suggested that individuals do appear younger when surrounded by younger faces and as the number of faces increases, the younger they appear. It is further suggested that ensemble coding may not explain the influence of time on age estimation, as target faces continued to appear younger when seen for four seconds.

Declaration

"This thesis contains no material which has been accepted for the award of any other degree of diploma in any University, and, to the best of my knowledge, this thesis contains no material previously published except where due reference is made. I give permission for the digital version of this thesis to be made available on the web, via the University of Adelaide's digital thesis repository, the Library Search and through web search engines, unless permission has been granted by the School to restrict access for a period of time."

Contributor Roles

ROLE	ROLE DESCRIPTION	STUDENT	SUPERVISOR 1	SUPERVISOR 2
CONCEPTUALIZATION	Ideas; formulation or evolution of overarching research goals and aims.		X	
METHODOLOGY	Development or design of methodology; creation of models.	X	X	
PROJECT ADMINISTRATION	Management and coordination responsibility for the research activity planning and execution.	X	X	
SUPERVISION	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.		X	
RESOURCES	Provision of study materials, laboratory samples, instrumentation, computing resources, or other analysis tools.		X	
SOFTWARE	Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code.	X	X	
INVESTIGATION	Conducting research - specifically performing experiments, or data/evidence collection.	X	X	
VALIDATION	Verification of the overall replication/reproducibility of results/experiments.	X	X	
DATA CURATION	Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.	X	X	
FORMAL ANALYSIS	Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data.	X	X	
VISUALIZATION	Visualization/data presentation of the results.	X	X	
WRITING – ORIGINAL DRAFT	Specifically writing the initial draft.	X		
WRITING – REVIEW & EDITING	Critical review, commentary or revision of original draft	X	X	

Do People Appear Younger When Surrounded by Younger Individuals?

The ability to accurately estimate someone's age is a fundamental part of our everyday lives. Daily, we unconsciously estimate the age of strangers, which influences our interactions (Pilz & Lou, 2022). Age estimation is crucial in many legal aspects, including assessing if an individual is old enough to purchase age-restricted items such as alcohol and tobacco (Pilz & Lou, 2022). Further, police often ask eyewitnesses of crimes how old they perceived the criminal, requiring them to estimate the age of a stranger (Thorley et al., 2022). Evidence suggests that people are often accurate to within a few years in estimating the age of strangers when they're viewed alone (Thorley, 2021). More recently, researchers have investigated how well we are able to estimate the age of a stranger when they are in a group. Awad et al. (2020) found that the perceived age of strangers changes when they're viewed within a group, with individuals appearing younger when surrounded by other younger individuals.

One proposed explanation for these findings is ensemble coding theory. Ensemble coding theory proposes that when presented with a group of stimuli, individuals automatically extract an average of the stimuli allowing them to rapidly form accurate impression of the objects (Phillips et al., 2018). This study will investigate the suitability of ensemble coding as an explanation by examining whether ensemble coding effects are accentuated when the number of younger flanker faces surrounding a target face increases, and whether they disappear when the groups of faces are seen for extended periods.

Studying Age Estimation

Many different methods have been used to study age estimation accuracy. The most common method used is presenting participants with passport-style photographs of an unfamiliar face and asking participants to estimate the age of the individual (Thorley et al., 2022). Less frequently used methods include sorting tasks where participants categorise and rank photographs from youngest to oldest (Rhodes, 2009). Voelkle et al., (2012), had participants estimate the age of individuals presented in photographs and found estimates of ages were inaccurate on average by 6.35 years. A similar average age estimate error was found by Amilon et al., (2007), who had participants estimate the age of an individual presented in a video recording and found an average of 5.1 years of inaccuracy for age estimation. In addition to finding out how accurate we can estimate strangers ages, researchers have also examined which facial cues are used to estimate age.

How Do We Estimate Age

Typically, people observe the facial features of an individual to make age estimates of strangers. As we age, there are a multitude of changes the face undergoes which impacts how we perceive the age of strangers (Rhodes, 2009). From childhood to adulthood, individuals undergo many facial developments which typically end at around 20 years of age (Burt & Perrett, 1995). As adults age, the eyes become smaller and sink deeper into their orbits, the ears and nose become elongated and lips becoming thinner (Thorley, 2021). These changes, including wrinkle development and the skin becoming saggier and thinner, have been found to influence an individuals perceived age (Gunn et al., (2009), Fink & Matts (2008)).

Liao et al., (2020), found that when estimating age, individuals focus on the central regions of the face, including the eyes, nose, and lips. Their study tracked participant’s eye movements when estimating the age of faces and found participants generally spend more time looking at the eyes, followed by the nose and mouth. In George & Hole’s (1998) study, they digitally swapped these facial features (eyes, nose, and mouth) between individuals of two differing ages. They found that adding older features to the younger face increased perceived ages by 40% and adding younger features to older faces decreased age perception by 33%. Evidently, age estimation is influenced by facial changes, however, there are other contributing factors to age estimation accuracy.

Age Estimation in Groups

Changes in age estimation of an individual when they appear in a group, is a contextual effect that has been studied by Awad et al., (2020). They examined how the presence of a group of flanker faces would influence an individual’s perception of a target face's age. Participants were required to estimate the age of a target face when viewed alone, flanked by younger faces and flanked by older faces. The stimuli used included passport-style photographs, with the three conditions randomly interleaved. Condition one had no flanker faces (i.e., the target was viewed alone), condition two had two younger flanker faces presented alongside the target face and condition three had two older flanker faces presented alongside the target face. The target face was always presented in the middle and in the flanked conditions the faces were presented horizontally with no gaps between the three faces. The flanker faces were always both younger or both older than the target face, with age offsets between the flanker and target of ±5, ±10, ±15 and ±20 years. Participants were not informed of the age range between stimuli. Participants were required to enter a two-digit numerical age estimation of the target face presented after viewing the target faces for two seconds. Contextual effects on age perception were found as results showed that the perceived age of the target changed when surrounded by flanker faces. The target faces were found to appear younger when flanked by younger faces and appeared older when flanked by older faces. This indicates that the presence of flankers does influence participant's estimates of strangers ages.

A similar study was conducted by Pilz & Lou (2022). The same method as Awad et al., (2020) was utilised, where participants were required to estimate the age of target faces presented in passport-style photographs for two seconds each. The three flanker conditions were randomly interleaved; however, all three conditions had a target face presented alongside two flanker faces. Condition one had a target face flanked by two identical faces as the target face, condition two had a target face flanked by two different younger faces and condition three had the target face flanked by two older target faces. The flanker faces presented in condition 2 and 3 were all within the age range of 10 years. Like Awad et al., (2020), Pilz & Lou (2022), found that target faces appear younger when flanked by younger faces and older when flanked by older faces. A more pronounced effect for the younger flanker condition was also found. The exact reason behind why these contextual effects occur are unknow, however, one potential explanation proposed by Awad et al. (2020) is ensemble coding.

Ensemble Coding

We frequently encounter groups of similar objects in our visual environment. Research indicates that our visual field has limited attentional and short-term memory capacity to process and store everything we view at a short glance (Haberman & Whitney, 2009). Studies suggest that rather than coding every element we view, the visual system utilises a mechanism known as ensemble coding (Haberman & Whitney, 2009). Ensemble coding allows individuals to quickly and effectively extract a summary statistic or average of the group of stimuli being viewed (Phillips et al., 2018). It allows individuals to process a rapid "gist" of the characteristics of large groups of stimuli that are being viewed for a short period of time (Elias et al., 2017). Using this mechanism reduces computational load for individuals when they do not have a long period of time to process each individual objects or characteristic in their visual field (Elias et al., 2017). Ensemble coding is also flexible and is used across a range of visual features (Elias et al., 2017). Effects of ensemble coding have been found in many low-level features such as motion and spatial orientation as well as mid-level features including perceived size of objects and depth perception (Whitney & Yamanashi, 2018). Recently, research has highlighted the significance of ensemble coding in high-level objects such as facial perception. Ensemble perception has been found to be useful in groups of face stimuli, allowing observers to quickly access the emotional tone or intent of a crowd at a short glance (Whitney & Yamanashi, 2018). de Fockert & Wolfenstein (2009), tested these ensemble coding effects in facial identity. They had participants observe four different faces for 2000ms, followed by a single test face and were asked to indicate whether the test face was previously presented in the set of four faces. The test was either a member of the preceding set of four, a member of a different set of four or a morphed average of the four previously viewed faces. Results found that if the target face was a morphed average, the probability of it being perceived as a member of the preceding four was significantly higher. These results support findings that ensemble coding is used to represent an average of facial emotions and individuals do extract averages of group faces. Similar results have also been found in studies looking at emotional expressions and facial attractiveness. (See Haberman & Whitney (2007, 2009) and Lou & Zhou 2018). However, there are many additional factors that contribute to ensemble coding representations.

The rapid process of ensemble coding allows individuals to form representations of groups of stimuli that are viewed for as little as 50ms; however, coding stimuli within a group tends to take longer, around 500ms for mid-level perception and closer to 2 seconds for higher-level perception such as facial identity and expressions in groups (Neumann, et al., 2018). It is suggested that the relationship between ensemble coding and representation depends on processing time, referred to as the time-dependent assumption (Liu et al., 2023). Li et al. (2016), examined this effect by manipulating the available processing time during a membership identification and mean discrimination task. Participants were first presented with a set of four emotionally differing faces followed by a test face. They were asked to indicate whether the test face was a member of the previously displayed set or not by answering “yes” for members and “no” for non-members. The target face presented to the participants was either a face with the mean emotion intensity of the preceding set or was a face not shown in the set at all. Exposure times were manipulated and counterbalanced between participants, with the set of four faces appearing for either 50ms, 500ms or 2000ms. Results found that when participants were exposed to the set of four faces for 50ms there was a higher rate of “yes” responses for the test face being a member of the preceding set of four. The same results were found when the stimuli were presented for 500ms, however, when stimuli were presented for 2000ms, the ratio of “yes” responses did not increase. These results are consistent with evidence of ensemble coding, where participants unconsciously represent the mean information of a set of stimuli when there is limited time available to process the information. When exposure time is increased, individuals have more time to process information. See Neumann et al., (2018) for similar findings. Findings from studies observing the influence of time on ensemble coding demonstrate that as viewing time increases, ensemble coding effects decrease, emphasising that the ability to process multiple faces is dependent on time.

Further, ensemble coding has been found to be sensitive to the number of stimuli presented. This effect was studied by Robitaille & Harris (2011), where they had participants observe the average size of circles and were then asked to indicate whether the mean size of the set was larger or smaller than the preceding target circle. Sets of 2, 4, 6, 8, 10 and 12 circles were used, and results showed that increasing the number of stimuli presented increased participant's estimation accuracy. These results lead to the conclusion that ensemble coding relies on a distribution attention model that operates across the whole display and therefore, ensemble coding benefits from a larger sample for the summary statistic to be calculated from. However, opposing results have been found in ensemble coding in facial emotion studies. Haberman & Whitney (2009), found that accuracy in determining the location of an individual face in sets of 1, 2, 3 or 5 declines from perfect accuracy in a set of 1 to 50% accuracy for a set of 4. These results indicate that as the number of stimuli increases, ensemble coding effects decrease. See Neumann et al., (2018) for similar findings.

Awad et al. (2020) propose ensemble coding as an explanation for their findings in their study on age estimation. This was proposed as an explanation as their results indicate that when a target face was surrounded by a group of younger faces, it appeared younger and when it was surrounded by older faces it appeared older, compared to when the target face was viewed alone. These results can be explained using ensemble coding as the theory proposes that individuals take an average of the group of stimuli being presented for a limited amount of time. Therefore, it is plausible that the participants in Awad et al's study estimated an average age of the faces being viewed causing the target face to appear older or younger when viewed in a group than when viewed alone. However, limited research has been done on the effects of ensemble coding on age estimation of faces.

Aims and Hypotheses

Age estimation studies have found that faces appear younger when flanked by younger faces and older when flanked by older faces. Ensemble coding has been proposed as an explanation for these findings.

Studies on ensemble coding theory have found ensemble coding effects appear to be sensitive to two factors: time and the number of stimuli. It has been found that as viewing time of stimuli increases, ensemble coding effects decrease. Further, research on the number of stimuli presented has found conflicting results. Studies observing the effects of the number of stimuli presented using shapes has found that as the number of stimuli increases, ensemble coding effects also increase, however, facial emotion studies have found that as the number of stimuli increases, ensemble coding effects decrease.

Manipulating these factors in the present study will enable us to determine whether ensemble coding does influence age estimation of strangers. This has two aims. First, to examine whether contextual effects increase as the number of younger flankers surrounding the target faces increases from two to four faces. Second, to examine whether contextual effects persist as the amount of time spent viewing the stimuli increases from two seconds to four seconds.

Participants were presented with 30 different target faces and were asked to provide a two-digit numerical estimate of the target faces perceived age. Target faces were either presented alone, flanked by two younger faces, or flanked by four younger faces. Half the participants viewed stimuli for two seconds and half viewed them for four seconds. All flanker faces were younger than the target face. Younger flanker faces were selected as both Awad et al. (2020) and Pilz & Lou's (2022) studies have previously found a more pronounced effect of target faces appearing younger when presented alongside younger flanker faces.

In the past, age estimation studies have only examined contextual effects when the target face is present alone versus presented with two flanker faces and stimuli has only been presented for two seconds. The present study will build on past findings by observing if two flankers versus four flankers influences age estimation. The time the stimuli is presented for will also be examined by presenting faces for 2 seconds as well as 4 seconds.

Past research indicates that ensemble encoding effects disappear when groups of objects are viewed for more than 1.6 seconds due to more information processing time. Awad et al. found that ensemble coding effects still occur after displaying stimuli for 2 seconds. To examine if Awad et al's results are due to ensemble coding, we will replicate their study by displaying stimuli for 2 seconds to half of our participants and for 4 seconds to half of the participants. Weaker effects should be found when faces are presented for 4 seconds if ensemble coding does have an impact.

Conflicting results have been found for the effects of the number of stimuli presented and ensemble coding as studies of shapes have found stronger contextual effects as number of stimuli increases, however, facial expression studies have found contextual effects decrease as number of stimuli increases. Past studies have not examined the effects of more than two flankers presented alongside a target face, therefore, our study will examine the influence of increasing the number of flankers on age estimation.

Given these conflicting findings around ensemble coding effects in shapes and faces, we are unsure of the impact the number of stimuli presented will have on the participants perceived age of the target faces. However, we expect that contextual effects will occur when faces are seen for two seconds but not for four seconds due to ensemble coding.

Our study will test if Awad et al. and Pilz & Lou's findings can potentially be explained by ensemble coding, or whether an alternative explanation may be more appropriate.

Method

Participants

An a-priori power analysis was calculated in G Power to ensure our study would detect a small-sized effect if the data were analysed using the anticipated 2 x 3 mixed measures ANOVA. The power analysis determined that a sample size of 164 participants would be required to detect a small effect (*Cohen's* $f = 0.10$, $1-\beta=0.80$, $\alpha=0.05$).

The final sample for the study included 237 participants (male = 49 , female = 186, other = 2), aged 17 – 40 years (*M* = 19 , SD = 3.04), 68% from European backgrounds, 23% from Asian background and the remaining from other diverse backgrounds. All participants were students at the University of Adelaide currently enrolled in Psych 1A (PSYCHOL1000). Students were recruited through the school of Psychology's online Research Participation System and the study was conducted online using Qualtrics. All participants received course credit for partaking in the study. The study was approved by the Human Research Ethics Committee of the University of Adelaide and all participants gave informed consent prior to participation.

Design

The study had a 2x3 mixed-subjects design, with two independent variables. The first independent variable was the number of flankers presented along a target face. This variable had a within-subject's design and three levels: (1) 0 flankers, (2) 2 younger flankers, and (3) 4 younger flankers. The second independent variable was coding duration, which is the length of time the faces are seen before participants estimate their age. This variable had a between subject's design and two levels: (1) 2 seconds and (2) 4 seconds. The study had one dependent variable, which was the participants' estimates of the target faces age in years.

Stimuli

180 different passport-style photographs of stranger's faces were used from the Minear & Park (2004) and Ebner et al. (2010) databases. The photographs were all

presented in colour and only showed the individuals head and shoulders. All faces were forward facing and had neutral facial expressions. The faces presented were all Caucasian and varied in hair and eye colour as well as hair length varying between photographs. None of the faces presented had any form of accessories obstructing the view of the face (e.g., glasses). The age range of the 180 faces presented was between 17 and 75 years. 90 of the faces presented were female aged between 18 and 64 years, and 90 were male between 18 and 75 years. The size of the image presented varied depending on the device used to complete this online study.

From the 180 photographs used within the study, participants were required to estimate the age of only 30 different faces. Each participant viewed a total of 90 different faces, 30 being the target face and 60 being younger flanker faces presented alongside the target face. The age of the 30 target faces presented to participants ranges between 27 and 75 years (M = 40.47, SD = 14.05). 15 of the faces being estimated were females aged between 27 and 64 (M= 36.4, SD = 10.4) and 15 were males aged between 28 and 75 (M = 44.53, SD = 16.3). The target faces presented were always 10 +/- 1 years older than the flanker faces presented alongside it (e.g. if the target face was 29 years old, the two or four flanker faces presented alongside it were all 18-20 years old).

Each individual photograph was digitally edited to ensure all faces appeared against a white background and appeared either alone, in a group of three or in a group of five. Thus, there were 3 versions of each photograph. Participants either viewed the target face alone in the middle of the screen, in a group of three, presented horizontally with a flanker face to the left and flanker face to the right of the target face or in a group of five with a flanker to left, right, above and below the target face, all with no gaps in-between them. The target face was always presented in the centre of the screen. See *Figure 1* for stimuli presentation.

Figure 1 - Stimuli Layout

Layout of stimuli in study. Target face presented alone, target face presented with 2 flanker faces and target face presented with 4 flanker faces.

Procedure

The study was conducted online using Qualtrics. Participants were encouraged to complete the study on a laptop or computer in full screen in a quiet, distraction free location of their choice. Instructions on how to complete the study in full screen were provided. An information page was initially shown to inform participants that the study would be examining their ability to estimate the age of a stranger’s face and if the number of faces in the group presented influences their age estimation. Participants were informed on how many faces they were estimating the age of and how many faces could appear on the screen at one time. There was no mention of the duration they would be viewing faces for. Any potential harms, benefits of the project and participants right to withdraw at any time was made clear prior to participants consenting to participate. After consenting, participants were required to complete a demographic's questionnaire. Instructions were then provided to inform participants on how the study would be run. These instructions stated that they would now be presented with a series of passport-style photographs that will appear in the middle of the screen and they are required to estimate the age of the target face by entering a 2-digit number into a textbox. Participants were informed that if faces appeared alone, they were required to estimate the age of that individual face. If faces appeared in a group of three or five, they were informed that they were required to estimate the age of only the face presented in the middle of the screen.

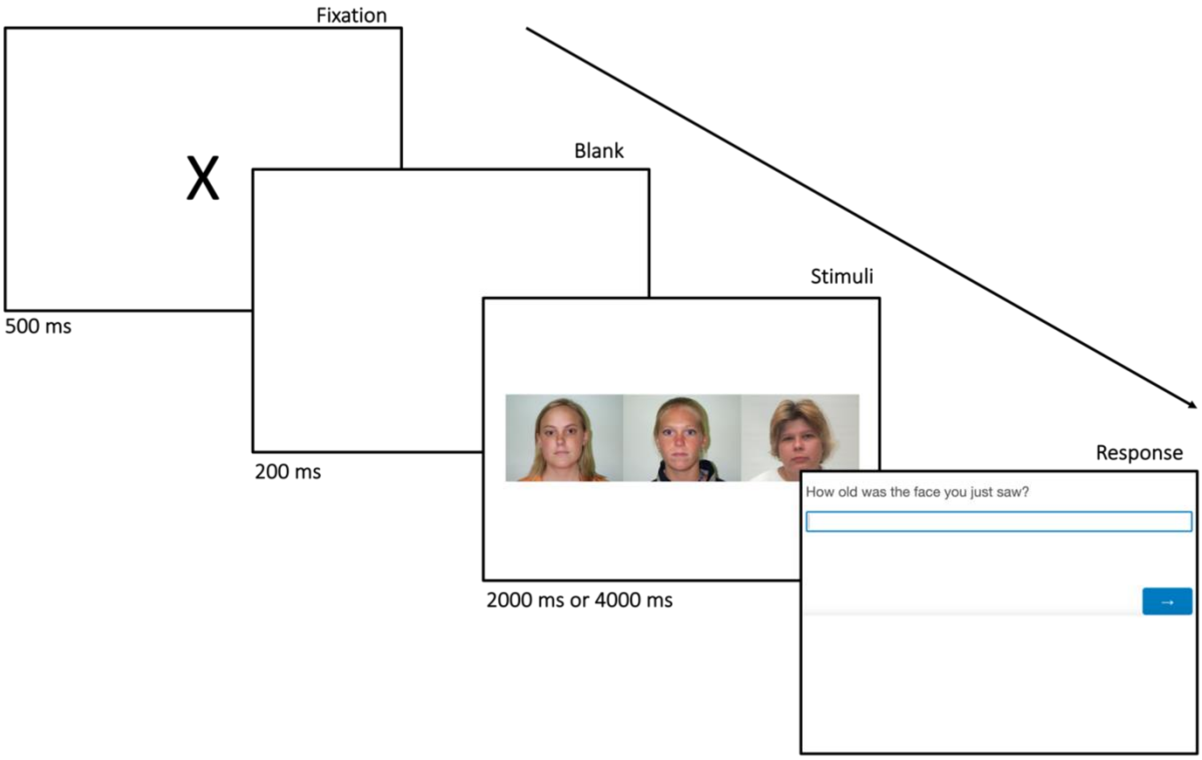
Participants then commenced a pilot practice study, consisting of three practice age estimation trials. The first image displayed was an individually presented face, the second was presented in a group of three with a flanker to the left and to the right of the target face and the final image was in a group of five with a flanker to the left, right, above and below the target face. The target face in each trial round was outlined in red to highlight to the participant which face they were required to estimate the age of. The stimuli presented in the practice trials were presented in the same format as the main study; however, the faces viewed in the practice study did not appear again in the present study and faces were not outlined in red in the real study. Following completion of the pilot study, participants were informed that the main study was about to commence.

Participants were then presented with the 30 target faces. Each target face was only presented once to the participant in a fully randomised order. The stimuli were presented in the middle of the screen against a white background. Participants always saw ten faces alone, ten faces within a group of three with a younger flanker either side and ten faces in a group of five with four younger flankers faces surrounding the target face. Five of the ten target faces were always male and five were female. The 30 target faces used within the study were counterbalanced across participants (e.g. participant 1 saw a 27 year old female alone, participants 2 saw her in a group of 3 with 2 younger flankers and participant 3 saw her in a group of 5 with four younger flankers). Participants were randomly allocated to either view the faces for 2 seconds or for 4 seconds. Further counterbalancing was done across participants to see the stimuli for either 2 seconds or 4 seconds creating 6 possible conditions participants could be randomly allocated to (e.g. participants 1 saw a 27 year old female alone for 2 seconds. Participant 2 saw her in a group of 3 with 2 younger flankers for 2 seconds. Participant 3 saw her in a group of 5 with four younger flankers for 2 seconds. Participant 4 saw her alone for 4 seconds. Participant 5 saw her in a group of 3 with 2 younger flankers for 4 seconds. Participant 6 saw her in a group of 5 with four younger flankers for 4 seconds). Participants were randomly allocated to one of the six study conditions and faces were presented in a random order for each participant.

Each trial began with a 500ms fixation point (X displayed in the middle of the screen), followed by a blank white screen for 200ms and then the stimulus was displayed for either 2 seconds or 4 seconds. Participants were then instructed to estimate the age of the target face. See figure 2 for illustration of trial sequence. A 2-digit numerical value was required to be entered in the textbox provided for participants to proceed to the next face. The study was completed after the participant had estimated the age of 30 target faces. The duration of the study on average was 8 minutes for the 2 second condition and 13 minutes for the 4 second condition.

Figure 2 - Trial Sequence Set Up

Trial sequence set up example. Fixation point X displayed in the middle first, followed by a blank screen. Face stimuli then presented where the participants are required to estimate the age of the target face presented in the middle, followed by the text box participants are required to enter a 2-digit age estimate into.



Results

This experiment required participants to provide age estimations of 30 passport style target faces. Faces were presented either alone, surrounded by two different flanker faces or by four different flanker faces. The faces were presented for either 2 seconds or four seconds. To examine the influence of the number of flankers and presentation time has on our dependant variable, age estimation in years, we converted each participant's age estimations in years to bias. These bias scores were calculated by determining how many years participants underestimated or overestimated the target faces veridical age. For example, if the target face was 30 years old and the participant estimated the age to be 35, their bias score was 5 years overestimated. Likewise, if the age estimation was 25, the bias score would be 5 years underestimated. Age estimation bias was calculated for all 6 conditions and each participant had a total of three scores reflecting their overall bias score for the alone condition, 2 flanker condition and 4 flanker condition.

Prior to data analysis, four outliers were identified during initial screening of the data. The identified outliers were removed from the data set due to random and unrealistic age estimations, with figures being consistently repeated throughout the same participant's estimations.

Following initial data screening, we conducted a bias analysis to determine if there was a significant difference for each of the 6 conditions. Results indicate that participants generally overestimated target faces ages. See table 1 for mean bias scores in each condition. Mean bias scores indicate that participants on average overestimated the target face's age by 2.29 regardless of the conditions. To determine if there was significant bias in the results a one-sample t-test was run. Results were statistically significant with all values t = > 4.7, p < .001 and Cohen's d = above .43. These results show a statistically significant overestimation of target age in each of the 6 conditions.

To examine if age estimations differed when faces were seen alongside two or four flankers as well as if faces were viewed for two or four seconds, a 2x3 mix measures ANOVA was conducted. We first checked that assumptions of a 2x3 ANOVA were met.

There was homogeneity of variance, with all Levene's test p-values being above .41. Data was normally distributed in each condition, with a Shaprio-Wilk test showing all p-values being above .16.

As assumptions of a 2x3 mixed measures ANOVA were met, we proceeded conducting the ANOVA. We found a significant main effect of number of flankers $F(2, 470) = 5.67, p < 0.005, \omega^2 = .006$. As a significant effect was found, post-hoc tests were conducted to identify which groups differed from each other. Results found there was no significant difference ($p = 1.00$) between the no flanker condition ($M = 2.60$) and two flanker conditions ($M = 2.44$). A significant, small effect ($p = 0.005$, $d = 0.2$) was found between the no flanker condition and four flanker condition ($M = 1.82$), indicating that faces look younger when surrounded by 4 flankers than when presented alone. A significant effect ($p = 0.03$, $d = 0.16$) was also found between the 2-flanker condition and 4 flanker condition, demonstrating that target faces appear younger when presented with 4 flanker faces than when presented with 2 flanker faces. We did not find a significant main effect of condition; 2 seconds vs 4 seconds $F(1,235) = 1.4, p > 0.05, \omega^2 < .001$. Additionally, there was no significant interaction between the number of flankers and time condition, $F(2, 470) = 0.41, p > 0.05$.

These reported results indicate that time does not impact age estimation bias. The number of flankers presented with the target does however influence age estimation bias as participants estimated the target faces to be younger when they were presented alongside 4 flanker faces than when they were presented with only two flanker faces or when presented alone.

Table 1 - Age Estimation Bias

Mean and standard deviation age estimation bias, in years, in three flanker conditions for 2 second and 4 second condition

Number of Flankers	**Condition**	**N**	**Mean**	**SD**	**95% CI Upper**	**95% CI Lower**
None	2 Seconds	119	2.22	3.80	2.91	1.54
	4 Seconds	118	2.99	4.00	3.71	2.26
Two	2 Seconds	119	2.27	3.85	2.94	1.58
	4 Seconds	118	2.62	4.16	3.37	1.87
Four	2 Seconds	119	1.62	3.74	2.29	0.95
	4 Seconds	118	2.03	4.13	2.77	1.28

Discussion

This study had two aims. The first was to examine whether contextual effects would increase as the number of younger flanker faces surrounding a target face increased. Here, we found participants perceived the target face to be younger as the number of flanker faces presented increased. The second aim was to examine whether contextual effects persist as the amount of time spent viewing stimuli increased. Here, it was concluded that time does not impact age estimation bias. Age estimation results will now be discussed in more detail. In accordance with past research, our participants overestimated the age of the target faces presented (see Voelkle et al., 2012; Amilon et al., 2007). Past studies have found age estimation of strangers to be overestimated by an average close to 5 years (Voelkle et al., 2012; Amilon et al., 2007; Thorley et al., 2022), however, participants in our study, on average, overestimated target faces age by a mean of 2.29 years. While an overestimation was still found, the lower average of overestimation in our study can be explained by other past studies that have found that the perceived age of a stranger changes when viewed in a group (Awad et al., 2020; Pilz & Lou 2022).

Age Estimation in Groups

Our study replicated Awad et al's (2020) and Pilz & Lou's (2022) findings on age estimation in groups. Our results found that when target faces were presented alongside two younger flanker faces, participants perceived the target face's age to be younger by an average of .2 years compared to when viewed alone. Further, when the target face was viewed alongside four younger flankers, the perceived age of the target face was younger by an average of .8 years, than when viewed alone and younger by .6 years on average when viewed with four younger flankers compared to two younger flanker faces. These results suggest that when an individual appears within a group of younger individuals, we perceived them to be younger and as the size of the group increases from a group of three to five, age estimates decrease. Similar results were found by Awad et al. (2020) and Pilz & Lou (2022), who also found that faces appeared younger when presented alongside two younger flanker faces compared to alone.

Moreover, we observed the effect timing would have on participants age estimates of target faces. Here, we did not find a significant result, indicating that time does not influence the age estimation of strangers. While past age estimation studies have not looked at the impact of time on age estimation, these results are surprising as it was anticipated that contextual effects would persist when participants viewed the stimuli for 2 seconds and decrease when participants viewed the stimuli for 4 seconds.

Age Estimation and Ensemble Coding

Our findings offer insight into the mechanisms used in age estimation. Limited research on age estimation in groups has been conducted, therefore, our study aimed to build on Awad et al. and Pilz & Lou's age estimation in groups findings, and similar results were found. An explanation proposed by Awad et al. (2020), for why individuals perceive faces to be younger when surrounded by younger flanker faces is ensemble coding theory. Ensemble coding theory proposes that when we have limited processing time available, we extract an average of large groups of stimuli we are presented with (Phillips et al., 2018). This theory could explain why participants in Awad et al., (2020) and Pilz & Lou's (2022) study perceived the target face to be younger when viewed in a group. Participants were only able to view the stimuli for 2 seconds, and therefore, ensemble coding was likely used to take an average of the all the faces and make an estimate of the target face's age. By taking an average age estimate of all three faces, where the two flanker faces were younger than the target, the target face appeared younger than it truly is.

Our study looked further into whether ensemble coding is an appropriate explanation for these findings. Studies have found that ensemble coding is influenced by two factors: the number of stimuli and available processing time. We observed the effects of these factors by combining and manipulating them both.

Past research has found conflicting results of the impact that the number of stimuli has on ensemble coding. Ensemble coding tasks looking at perception of shape size, found that as the set size increases, ensemble coding increases. Proposing that ensemble coding is used when large numbers of stimuli are presented. In accordance with ensemble coding theory and studies looking at the influence of the number of stimuli (see Robitaille & Harris, 2011), we found similar results in support of the utilisation of ensemble coding in age estimation. Participants in our study perceived the target face to be younger when presented in a group with other younger faces, compared to when viewed alone. As the group size increased from three to five, contextual effects increased, with target faces continuing to be perceived as younger than their actual age. These results indicate that participants may be using ensemble coding to make their age estimates. As the surrounding flankers are younger than the target face, the average age taken during the ensemble coding process, causes the target face to appear younger. Conflicting results have been found in previous research focusing on the influence of the number of stimuli presented on facial identity. It has been found that as the number of stimuli increases, ensemble coding effects decrease (Haberman & Whitney, 2009 and Neumann et al., 2018). A potential explanation for these conflicting results is that these studies focus on facial identity and the participants accuracy and ability to determine whether they had previously viewed the target. Therefore, perhaps ensemble coding is not used in memory tasks of facial identity, but, as our results suggest, are used in perception-based tasks such as age estimation. Our findings here support Awad et al's proposed explanation that ensemble coding is used in the age estimation of strangers.

To further examine the effects of ensemble coding on age estimation in groups, we manipulated viewing time. Past studies have found that ensemble coding effects decrease after 1.6 seconds, presumably due to having more time to process information (Neumann et al., 2018). We, therefore, expected to find that contextual effects would disappear as viewing time increased from 2 seconds to 4 seconds. Surprisingly, we found no effect. Our results showed that contextual effects persisted as the viewing time increased, and it is therefore concluded that timing does not influence the age estimation of strangers in groups. Opposing results have been found in previous facial identity studies and the influence of time. Past research has found that when participants are given more time to view stimuli, ensemble coding effects decrease as there is more time available to process information, supporting ensemble coding theory. One possible explanation for these results may be because of the differing types of tasks being used. These studies that support ensemble coding theory (see Lie et al., 2016 and Neumann et al., 2018), have used memory and true or false tasks. Having more processing time available in these memory identification tasks would increase the participants accuracy in determining if they have seen the target face previously, as they have longer to process information and are not needing to extract an average of the stimuli. Thus, resulting in ensemble coding effects disappearing. Our study differed as it relied on subjective measures, using a perception-based task. Our findings, alongside Carragher et al., 2019, propose that viewing time does not matter in perception-based tasks. Carragher et al., 2019, examined the influence of time on the 'cheerleader effect', which has been found to rely on the averaging process used in ensemble coding. Participants rated the attractiveness of a target face presented either alone or in a group, for varying lengths of time, up to 7000ms. Results found that faces were perceived to be more attractive when presented within a group than presented alone, regardless of viewing time. Concluding that contextual effects still occur at longer exposure durations up to 7 seconds. Results from Carragher et al., 2019 and our study here, indirectly suggest that timing does not influence the process of ensemble coding in perception-tasks. A proposed explanation for these findings is that the influence of time on ensemble coding is task dependant. With ensemble coding appearing to be used in memory-based tasks and not in perception-based tasks.

Implications

Our findings here are important in understanding the mechanisms behind estimating the age of strangers in groups. A large amount of research has been conducted on the age estimation of strangers more generally. However, recently research has found that age estimation of stranger's changes when the target face appears in a group. A very small amount of research on the effects of groups on age estimation has been conducted. Therefore, only a proposed explanation, ensemble coding, has been suggested to explain these results. Our findings here contribute to this emerging body of literature and suggest that the proposed explanation, ensemble coding, appears to be a plausible explanation for the contextual effects that occur in the age estimation of strangers in groups.

Limitations

Our study had limitations that may influence the generalisability of its findings. One limitation is that participants were only required to estimate the age of strangers in a group of up to five individuals. Past research on lower-level perception of shapes, where sets of up to 12 circles were presented, have found that ensemble coding benefits from individuals being presented with a larger number of stimuli. Increasing stimuli enables a more accurate summary statistic to be calculated (Robitaille & Harris, 2011). Previous age estimation in groups studies have only examined the effects of two flanker faces on age estimation and our study built on these findings by examining the effects of up to 4 flanker faces. To further determine if ensemble coding is used when estimating the age on strangers when in a group, future research should further increase the number of flankers presented to examine the influence this has on age estimation and if contextual effects still appear as the number of stimuli continues to increase. This will provide more evidence as to whether ensemble coding is a feasible explanation for the results found here and in Awad et al. (2020) and Pilz & Lou (2022) studies.

Another limitation is that the groups of faces presented in our study were all the same gender. While stimuli included both male and female faces, there was never a group of mixed gender faces presented to the participants. Viewing faces of different genders may have an influence on the averaging process used in ensemble coding, as there is more information to process. Future research should therefore examine whether having groups of mixed gender faces will influence whether contextual effects persists or if they will disappear, and if having mixed genders will make age estimates of a target face more or less accurate.

Conclusion

Our results show that individuals perceive target faces to appear younger when surrounded by other younger faces in the age estimation of strangers. Moreover, contextual effects increase as the number of faces presented increases. Indicating that a target face continues to appear younger as the size of the group presented increases. Additionally, contextual effects also persist as viewing time available increases, indicating that time does not impact age estimation. These findings are useful in understanding the mechanisms behind age estimation in groups and demonstrates that ensemble coding is a plausible explanation for the contextual effects found in the age estimation of strangers in groups. The results further indicate that the number of stimuli does influence ensemble coding in perception-based tasks, however, the factor of time is not relevant for the ensemble coding used in age estimation. Rather estimating the age of strangers in groups, relies on an averaging process of the stimuli regardless of how much processing time is available.

References

Amilon, K., van de Weijer, J., & Schötz, S. (2007). The impact of visual and auditory cues in age estimation. *Speaker Classification II: Selected Projects*, 10-21.

Awad, D., Clifford, C. W., White, D., & Mareschal, I. (2020). Asymmetric contextual effects in age perception. *Royal Society Open Science*, *7*(12), 200936.

Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces: Computer graphic manipulation of shape and colour information. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *259*(1355), 137-143.

de Fockert, J., & Wolfenstein, C. (2009). Short article: Rapid extraction of mean identity from sets of faces. *Quarterly Journal of Experimental Psychology*, *62*(9), 1716-1722.

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and Validation. *Behavior Research Methods*, *42*, 351-362.

Elias, E., Dyer, M., & Sweeny, T. D. (2017). Ensemble perception of dynamic emotional groups. *Psychological Science*, *28*(2), 193-203

Fink, B., & Matts, P. J. (2008). The effects of skin colour distribution and topography cues on the perception of female facial age and health. *Journal of the European Academy of Dermatology and Venereology*, *22*(4), 493-498.

George, P. A., & Hole, G. J. (1998). The influence of feature-based information in the age processing of unfamiliar faces. *Perception*, *27*(3), 295-312.

Gunn, D. A., Rexbye, H., Griffiths, C. E., Murray, P. G., Fereday, A., Catt, S. D., ... & Christensen, K. (2009). Why some women look young for their age. *PloS one*, *4*(12), e8021.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751-R753.

Haberman, J., & Whitney, D. (2009). Seeing the mean: ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 718.

Li, H., Ji, L., Tong, K., Ren, N., Chen, W., Liu, C. H., & Fu, X. (2016). Processing of individual items during ensemble coding of facial expressions. *Frontiers in Psychology*, *7*, 1332

Liao, D., Ishii, L. E., Chen, J., Chen, L. W., Kumar, A., Papel, I. D., ... & Ishii, M. (2020). How old do I look? Exploring the facial cues of age in a tasked eye-tracking study. *Facial Plastic Surgery & Aesthetic Medicine*, *22*(1), 36-41.

Liu, R., Ye, Q., Hao, S., Li, Y., Shen, L., & He, W. (2023). The relationship between ensemble coding and individual representation of crowd facial emotion. *Biological Psychology*, 108593.

Luo, A. X., & Zhou, G. (2018). Ensemble perception of facial attractiveness. *Journal of Vision*, *18*(8), 7-7.

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, *36*, 630-633.

Neumann, M. F., Ng, R., Rhodes, G., & Palermo, R. (2018). Ensemble coding of face identity is not independent of the coding of individual identity. *Quarterly Journal of Experimental Psychology*, *71*(6), 1357-1366.

Pilz, K. S., & Lou, H. (2022). Contextual and own-age effects in age perception. *Experimental Brain Research*, *240*(9), 2471-2480.

Phillips, L. T., Slepian, M. L., & Hughes, B. L. (2018). Perceiving groups: The people perception of diversity and hierarchy. *Journal of Personality and Social Psychology*, *114*(5), 766.

Rhodes, M. G. (2009). Age estimation of faces: A review. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(1), 1-12.

Robitaille, N., & Harris, I. M. (2011). When more is less: Extraction of summary statistics benefits from larger sets. *Journal of Vision*, *11*(12), 18-18.

Thorley, C. (2021). How old was he? Disguises, age, and race impact upon age estimation accuracy. *Applied Cognitive Psychology*, *35*(2), 460-472.

Thorley, C, Acton B, Armstrong, J, Ford, S & Gundry, S. "Are estimates of faces' ages less accurate when they wear sunglasses or face masks and do these disguises make it harder to later recognise the faces when undisguised?." *Cognitive Research: Principles and Implications* 7, no. 1 (2022): 1-12.

Voelkle, M. C., Ebner, N. C., Lindenberger, U., & Riediger, M. (2012). Let me guess how old you are: effects of age, gender, and facial expression on perceptions of age. *Psychology and Aging*, *27*(2), 265.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, *69*, 105-129.