Transmitting and decoding facial expressions of emotion during healthy aging: More similarities than differences

Marie L. Smith School of Psychological Sciences, Birkbeck College, University of London, London, UK

Daniel Grühn Department of Psychology, North Carolina State University, Raleigh, NC, USA

Ann Bevitt School of Psychological Sciences, Birkbeck College, University of London, London, UK

Mark Ellis School of Psychological Sciences, Birkbeck College, University of London, London, UK

Oana Ciripan School of Psychological Sciences, Birkbeck College, University of London, London, UK

Susan Scrimgeour School of Psychological Sciences, Birkbeck College, University of London, London, UK

Michael Papasavva School of Psychological Sciences, Birkbeck College, University of London, London, UK

Louise Ewing School of Psychological Sciences, Birkbeck College, University of London, London, UK; School of Psychology, University of East Anglia, Norwich, UK; Australian Research Council Centre of Excellence in Cognition and its Disorders, School of Psychology University of Western Australia, Crawley, Western Australia, Australia

Older adults tend to perform more poorly than younger adults on emotional expression identification tasks. The goal of the present study was to test a processing mechanism that might explain these differences in emotion recognition—specifically, age-related variation in the utilization of specific visual cues. Seventeen younger and 17 older adults completed a reverse-correlation emotion categorization task (Bubbles paradigm), consisting of a large number of trials in each of which only part of the visual information used to convey an emotional facial expression was revealed to participants. The task allowed us to pinpoint the visual features each group used systematically to correctly recognize the emotional expressions shown. To address the possibility that faces of different age groups are differently processed by younger and older adults, we included younger, middle-aged, and older adult face models displaying happy, fearful, angry, disgusted, and sad facial expressions. Our results reveal strong similarity in the utilization of visual information by younger and older adult participants in decoding the emotional expressions from faces across ages—particularly for happy and fear emotions. These findings suggest that age-related differences in strategic information use are unlikely to contribute to the decline of facial expression recognition skills observed in later life.

Citation: Smith, M. L., Grühn, D., Bevitt, A., Ellis, M., Ciripan, O., Scrimgeour, S., Papasavva, M., & Ewing, L. (2018). Transmitting and decoding facial expressions of emotion during healthy aging: More similarities than differences. *Journal of Vision*, *18*(9):10, 1–16, https://doi.org/10.1167/18.9.10.

https://doi.org/10.1167/18.9.10 Received August 15, 2017; published September 11, 2018 ISSN 1534-7362 Copyright 2018 The Authors

Introduction

Successful expression recognition is an integral part of adaptive emotional functioning. Theoretical models of key domains, such as empathy (e.g., Preston & de Waal, 2002), emotion understanding (e.g., Castro, Cheng, Halberstadt, & Grühn, 2016), and social communication (e.g., Haxby, Hoffman, & Gobbini, 2002), emphasize the importance of emotion recognition as a foundation for effective processing of affective information and successful social interaction. The empirical evidence in adulthood shows relatively consistent age differences on laboratory measures of emotion processing. Older adults tend to do worse than younger adults in accurately decoding emotions in static facial expressions (e.g., Isaacowitz et al., 2007; Ruffman, Henry, Livingstone, & Phillips, 2008). However, these differences do seem to differ by the emotion shown. In particular, small or no age differences are typically found for happy and disgusted expressions. Larger age differences are generally found for negative facial expressions (including fear, anger, and sadness; e.g., Ebner & Johnson, 2009) though the observed profiles vary substantially by study (Isaacowitz & Stanley, 2011).

Although there is a degree of consensus about the existence of age differences in correctly identifying emotional expressions, there is a debate about the underlying mechanisms (Isaacowitz & Stanley, 2011; Ruffman, 2011). Motivational, neurophysiological, cognitive, and visual scanning explanations have been proposed. The first of these accounts emphasizes potential age differences in motivations. In particular, based on ideas in socioemotional selectivity theory (Carstensen, Isaacowitz, & Charles, 1999), it is suggested that older adults might default to focus more on positive emotions and avoid negative emotions. Thus, older adults might be biased to attend less to negative expressions than to positive expressions. Neurophysiological explanations focus on age-related changes in brain regions responsible for emotion recognition in terms of structural differences or differences in the utilization of neurotransmitters (Ruffman et al., 2008). These accounts argue that declines in brain functionality associated with healthy aging are causing age differences in successful emotion recognition. Cognitive explanations argue that emotion recognition involves cognitive skills (e.g., processing speed) and as these skills show age-related declines more broadly, emotion recognition skills also necessarily decline (Garcia-Rodriguez et al., 2011).

A problem with the motivational, neurophysiological, and cognitive accounts is that they seem to lack a good explanation for the reported differences in decoding different negative emotions. Impairments in the recognition of sadness and anger appear to be the strongest and most consistent, with impairments in processing fear less consistent (see the meta-analysis by Ruffman et al., 2008; review by Isaacowitz et al., 2007). Here, systematic differences in viewing behavior might be able to provide an explanation. Visual scanning accounts argue that there may be age differences in gaze patterns when viewing expressive faces. This variability is critical because some patterns might be more functional in recognizing the correct emotion than others. For example, older adults, relative to younger adults, are observed to direct their gaze more to the lower halves (mouth) of faces than the upper halves (eyes: Murphy & Isaacowitz, 2010; Sullivan, Ruffman, & Hutton, 2007; Wong, Cronin-Golomb, & Neargarder, 2005). Such fixation to the lower half of the face may selectively and negatively impact recognition of emotions like anger and sadness, which primarily rely on examining the upper half rather than the lower half of the expressive face (Calder, Keane, Young, & Dean, 2000; Smith, Cottrell, Gosselin, & Schyns, 2005). For happy expressions, where performance in older adults tends to be spared, the most useful visual information is found in the lower half of the face. The critical information for fear comprises both the wide-opened eyes in the upper half of the face, which are the most consistently used and salient cue, alongside lower spatial frequency (SF) information from the open mouth, which is used more when the context renders this information to be informative (Smith & Merlusca, 2015).

Age differences in visual scanning might form part of the account of age differences in emotion recognition; however, they are likely to be only part of the story. For example, Murphy and Isaacowitz (2010) found that even after controlling for eye-fixation scores as well as cognitive and affective variables, age differences persisted in emotion recognition in faces. Thus, gaze pattern differences in themselves seem insufficient to fully account for age differences in emotion recognition (see also Sullivan et al., 2007). This may be because measuring the scanning pattern via successive gaze fixations does not indicate whether the information is also *used* in categorizing the emotion shown. Gaze pattern is a good proxy for information usage but is not identical to it. A stronger test would be to control the exact information from the face provided in order to investigate which pieces of information (i.e., facial features) are used to decode an emotion correctly, when made available to participants.

Smith et al. (2005) used a reverse correlation paradigm (Bubbles paradigm; Gosselin & Schyns, 2001) to investigate the facial features that young adults use to successfully decode the six basic facial expressions of emotion (plus neutral). On each trial they presented participants with subsampled versions of expressive faces, where the only parts of the expressive face visible to inform their emotion categorization decisions were those that sat behind randomly positioned Gaussian apertures (bubbles), and the rest of the face was hidden from view. By changing the location of the bubbles on each trial, they reverse-engineered the importance of different facial regions for correct categorization performance to create face maps of information use (e.g., if the eyes always led to correct fear categorization performance, this region would be indicated as a significant driver of correctly categorizing fear). The results pinpointed the diagnostic facial features for young adults identifying the different emotion expressions. For example, the broad smiling mouth was critical when decoding happy expressions, the wide-open eyes when decoding fear, and the wrinkles around the nose and mouth when decoding disgust.

The present study

The aim of the present study was to investigate whether there are any age differences in the utilization of facial features when decoding different emotional expressions. To do this, we again employed the bubbles reverse correlation approach (Gosselin & Schyns, 2001; Smith et al., 2005). As the bubbles task provides only fragments of the facial features to the decoder on each trial, and the stimuli are presented foveally, gaze patterns are practically eliminated in this task. Thus participants have access to all facial information without the need to move their eyes. If younger and older adults differ in the utilization and processing of facial features, different face maps will be generated for younger and older adults.

Past research is suggestive of a same-age bias that might drive younger and older adults to focus on different facial features in younger and older faces (for a review, see Rhodes & Anastasi, 2012). Furthermore, several recent studies have indicated that older faces provide a noisier visual signal for all observers that is considered to be less expressive due to the presence of wrinkles around key features (Folster, Hess, & Werheid, 2014; Hess, Adams, Simard, Stevenson, & Kleck, 2012). To control for the possibility that the age of the face stimulus is an important feature in decoding the emotional expression, we included younger, middle-aged, and older face models in our task. Should facial aging and in particular the presence of wrinkles and folds on the face alter the way in which emotions are transmitted with age, we would expect to see differences in the face maps generated for the categorization of expressions displayed by younger and older face models, which may interact further with participants' age.

Information from different SF bands has been shown to be important for decoding emotional expressions in young adult faces (e.g., Schyns, Petro, & Smith, 2007; Smith et al., 2005; Smith & Merlusca, 2015; Smith & Schyns, 2009; Vuilleumier, Armony, Driver, & Dolan, 2003). To investigate whether this bias holds across age groups, in the bubbles task we sampled the visual information across five different SF bands ranging from fine details in high SFs to coarse information in the low SF bands. This feature of the design allowed us to examine the importance of visual information separately at each SF (as well as when these are combined) for older and younger adults.

Finally, the impact of the comparison (emotion) categories used in expression categorization tasks has been raised as an important methodological issue in this field. Recently, Smith and Merlusca (2015) showed that different visual information was used to categorize the same fearful face expressions when they were presented in comparison with one or two other emotions (e.g., fear vs. happy) rather than in a multiple categorization task (e.g., alongside all other basic emotional expressions). For this reason, we used the more "true to life" scenario of establishing the emotion in an expressive face from one of a set of possibilities (happy, fearful, angry, disgusted, or sad) rather than selecting a smaller set of facial expressions of emotion.

Methods

Participants

We initially recruited 20 younger adult participants mainly from the student population of the University of London. Two participants were subsequently excluded due to high scores on the Toronto Alexithymia Scale, a self-report questionnaire measure of alexithymia (TAS-20, >54; Bagby, Parker, & Taylor, 1994) measuring the ability to identify emotions in the self and others. Four participants were also excluded for excessively poor performance in the task generally (defined as less than 50% accuracy for any emotion category), which we considered likely to signal a lack of attention/engagement in the task. Similarly, we initially recruited 20 older adults from the local region of London. Older participants (aged 60 years and over) were mainly recruited from the London branch of the University of the Third Age, a voluntary adult learning organization. Three participants were subsequently excluded, one due to poor corrected vision as measured on the LogMAR chart (Bailey & Lovie, 1976; LogMAR ≥ 0.7 in both eyes), one due to a high score on the TAS-20 and one as a result of poor task performance. Subsequently we recruited three more young adults to provide a matched sample size to the older participants group. The final sample therefore comprised 17 younger adults aged 18 to 32 years (M = 24.8 years, SD = 4.9, three male) and 17 older adults aged 62 to 81 years (M = 70.1 years, SD = 5.0, six male). Younger adults (20/20 vision on the Sloan ETDRS Near Vision from Precision Vision, PLC; LogMAR 0.0) and older adults (20/32 vision, LogMAR 0.2) had normal to corrected-to-normal vision. All participants provided informed consent according to the declaration of Helsinki and were reimbursed for their time either at a rate of £8 per hour or via compulsory course credits.

Stimuli

Stimuli comprised grayscale versions of expressive faces posed by young (19–31 years), middle-aged (39–55 years), and older (69–80 years) models taken from the FACES database (Ebner, Riediger, & Lindenberger, 2010). Three male and three female models were selected for each age group for a total of 18 models. Each individual was shown displaying five facial expressions of happiness, fear, anger, disgust, and sadness (for a total of 90 face stimuli). Stimulus images were further standardized by cropping to a standard size with all nonface information (e.g., neck and shoulders) removed and by horizontally aligning the center of each pupil.

On each experimental trial, subsampled versions of these expressive faces were created by randomly sampling visual information from the original images using circularly symmetric Gaussian apertures, or "bubbles" (Gosselin & Schyns, 2001). Only the information located behind these apertures was visible to the participant and could therefore inform their categorization decisions; the rest was hidden from view. The random positioning of the apertures ensures that a different combination of visual information was presented to the participant on every trial. Across the course of each testing session, this random sampling approximates a uniform sampling of the input information space and allows a nonbiased exploration of the importance of all of the available visual information for the categorization task. The number of apertures was adjusted for each observer on a trial-per-trial basis to maintain 75% correct performance for each condition of interest (age × emotion, minimum = 40 bubbles, maximum = 250). A greater number of apertures means that more information has been shown to the participant (to counter poor performance).

To sample the utilization of visual information across different SF bands, each face image was first decomposed into five nonoverlapping SF bands of one octave each (120–60, 60–30, 30–15, 15–7.5, and 7.5–3.75 cycles per face) using the Pyramid toolbox for MATLAB (Simoncelli, 1999). In this way the visual information available in five nonoverlapping SF bands (ranging from coarse shapes to the fine details) was extracted from the original image (see Smith et al., 2005 for an illustration of the stimulus generation procedure). Each SF band was then independently sampled with randomly positioned apertures and then recombined to produce one experimental stimulus comprised of a mixture of high and low SF information in randomly determined locations (for further details of the methods and another illustration of the stimulus generation process, see Gosselin & Schyns, 2001). Thus on each trial participants saw a mixture of low, mid, and high SF information from randomly selected locations across the face, and had to base their emotion categorization decisions on this information. Stimuli were projected on a light gray background to the center of a screen at a distance of 65 cm from the participant so that the visual angle was $5.36° \times 3.7°$ in keeping with earlier studies (Smith et al., 2005; Smith & Merlusca, 2015).

Measures

Basic cognitive functioning was assessed with the 40-item Shipley vocabulary test (Shipley, Gruber, Martin, & Klein, 2009) as a measure of participants' crystallized intelligence and the digit symbol substitution task (Wechsler, 1981) to measure their perceptual speed. To ensure the participants did not have any difficulties in decoding emotional information in the self and others, participants completed the TAS-20 (Bagby et al., 1994).

Procedure

In three separate recording sessions (each completed on a different day), participants completed a total of 2,700 trials of the bubbles task. A randomly selected expressive face image was presented on each trial with each expression and individual model presented an equal number of times in a fully randomized design (i.e., there were 30 repetitions of each individual expressive face across the whole experiment). Each session comprised 900 trials made up of 60 repetitions of each age group (three) and emotion expression (five) combination. To maintain concentration and motivation, short breaks were provided every 90 trials (approximately every 4 to 5 min) consisting of generic motivational screens (e.g., "keep up the good work," odd-numbered blocks), interactive "puzzle-bubble games" (even-numbered blocks in Sessions 1 and 3, see Smith, Cesana, Farran, Karmiloff-Smith, & Ewing, 2017), or additional tests and questionnaires (even-numbered blocks in Session 2) including the TAS-20, the Shipley vocabulary test, and three administrations of the digit symbol substitution task. The experimenter remained in the testing cubicle with the participant throughout the experiment to encourage continued task engagement and to administer the additional tests and puzzle bubble games during breaks.

Each experimental trial began with a 500-ms fixation cross, which was immediately followed by the subsampled expressive face image. This face remained onscreen for 1500 ms and was replaced by a uniform gray screen until a response was given. Participants could respond at any time after stimulus onset and were instructed to respond as quickly as possible without making mistakes by pressing labeled buttons on the computer keyboard. Six response keys represented each emotional category (happy, sad, fear, disgust, and anger) and an "I don't know" option if participants felt that they could not make an accurate judgment based on the information presented. A short training phase at the start of each experimental session (including full face and subsampled stimuli) ensured that participants understood the task and were familiarized with the response keys. At the end of the third session participants completed a short postexperiment check, a categorization task with non-bubbled images, to establish their accuracy in categorizing the intact full-face expressive stimuli.

Results

Preliminary analyses

Consistent with general trends in the aging literature, older adults ($M = 37.8$, $SD = 2.0$) outperformed younger adults ($M = 29.2$, $SD = 4.6$) on the Shipley vocabulary task, $t(22.07) = 7.1$, $p < 0.001$, $d = 2.42$ (corrected for unequal variance between the groups). In contrast, younger adults ($M = 82.1$, $SD = 14.1$) performed better on the perceptual speed task, $t(32) = 3.4$, $p = 0.002$, $d = 1.16$, than older adults ($M = 66.8$, $SD = 12.1$). There were no significant age differences in self-reported alexithymia (scores on the TAS-20), $t(32) = 0.19$, $p = 0.85$, $d = 0.07$, between younger ($M = 37.2$, $SD = 8.7$) and older ($M = 37.9$, $SD = 10.8$) adults.

Bubble task: Performance

For the bubbles task, we varied the amount of information (number of bubbles) revealed on each trial for each of the 15 conditions (3 Model Age groups $\times$ 5 Emotions) independently via a staircase algorithm in an effort to equate performance levels at 75% correct. As a result, participants saw more information (more bubbles) for an experimental condition when their performance for that condition was low and less information when performance was high (standard methodology with the bubbles paradigm; e.g., Smith et al., 2005). Despite this, some differences in performance remained, reflecting well-reported variability in the ease with which some emotions can be decoded over others (see Figure 1). To examine potential performance differences, and in particular to establish the effect of healthy aging on the pattern of observed results, we examined the percentage of correctly identified emotional expressions for each age group of stimuli and participants. To this end, we ran a mixed ($3 \times 5 \times 2$) analysis of variance (ANOVA) with model age (young, middle-aged, and older faces) and emotion expressed (happy, fear, sad, disgust, or anger) as within-subjects factors and participants' age group (young or older adults) as a between-subjects factor.

We found significant main effects of model age, $F(2, 64) = 1.43$, $p < 0.001$, $\eta^2 = 0.31$, and emotional expression, $F(3.02, 96.5) = 27.9$, $p < 0.001$, $\eta_p{}^2 = 0.47$, as well as a significant interaction between model age and emotional expression, $F(4.46, 142.76) = 6.5$, $p < 0.001$, $\eta_p{}^2 = 0.17$. The main effect of model age reflected slightly reduced performance for identifying the emotion in older faces ($M = 74.2\%$, $SD = 4.8\%$) than in younger faces, $t(33) = 4.5$, $p < 0.001$, $d = 1.08$, $M = 76.4$, $SD = 3.6$, and middle-aged faces, $t(33) = 4.3$, $p < 0.001$, $d = 1.03$, $M = 76.8$, $SD = 3.8$). There was no significant differences in correctly recognizing emotions in younger and middle-aged faces, $t(33) = 0.75$, $p = 0.46$, $d = 0.18$. The main effect of emotion reflected better performance for happy ($M = 80.3$, $SD = 2.0$) and fearful faces ($M = 79.7$, $SD = 3.0$) than for sad faces ($M = 75.6$, $SD = 5.9$) and all three showed better performance than faces depicting disgust ($M = 71.3$, $SD = 7.7$) and anger ($M = 72.2$, $SD = 6.8$). All pair-wise comparisons were significant, all ts > 2.9, ps < 0.007, ds > 0.7, with the exception of happy versus fear, $t(33) = 1.6$, $p = 0.12$, $d = 0.39$) and disgust versus anger, $t(33) = 0.7$, $p = 0.47$, $d = 0.17$). No other effect reached significance; in particular, there was no significant main effect of participant age group or significant interaction effect with age group (all Fs < 1.8, ps > 0.17, $\eta_p{}^2 < 0.05$), which confirms that performance on the task was equivalent across our younger and older participant groups.[1]

To disentangle the significant interaction between the expressed emotion and the age of the face model, we conducted separate follow-up analyses on each expressed emotion. For fear, $F(1.22, 40.5) = 0.24$, $p = 0.68$, $\eta_p{}^2 = 0.007$, there were no significant performance differences related to face age, with only minor differences for happiness, $F(2, 66) = 3.7$, $p = 0.031$, $\eta_p{}^2 = 0.1$, with younger faces enhanced with respect to middle-aged and older faces (80.9 vs. 79.9, $t[33] = 2.86$, $p = 0.007$, $d = 0.69$; 80.9 vs. 80.2, $t[33] = 1.8$, $p = 0.08$, $d = 0.44$, respectively). There were, however, strong effects related to face age for anger, $F(2, 66) = 13.73$, $p < 0.001$, $\eta_p{}^2 = 0.29$; sadness, $F(2, 66) = 9.5$, $p < 0.001$,

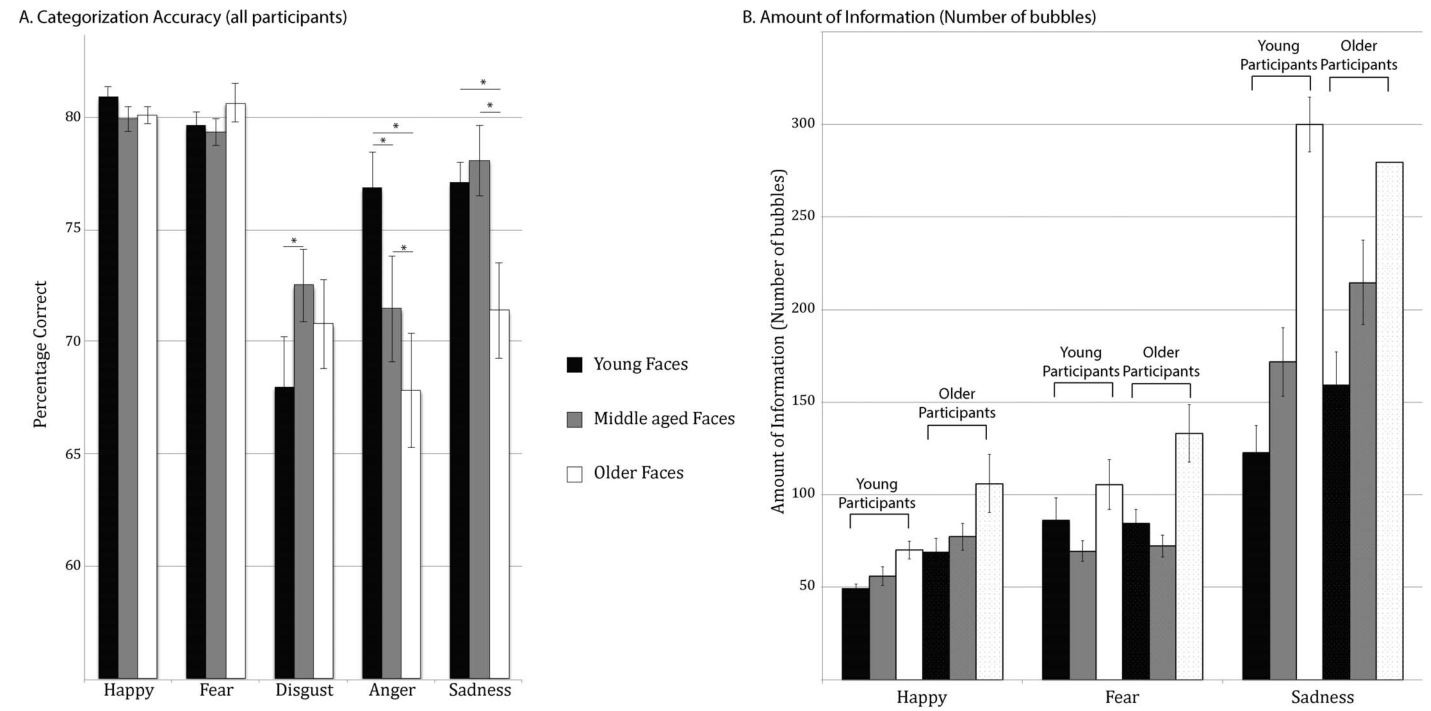

Figure 1. (A) Categorization performance (accuracy) in the bubbles task for younger, middle-aged, and older faces displaying five facial expressions of emotion. There were no significant differences in performance across the participant groups so averaged results are shown. (B) The amount of information required to achieve the performance levels shown in Panel A—that is, the average number of bubbles used to present the three emotional expressions where significant effects of participant group are observed (happy, fear, and sadness). Error bars represent the standard error of the mean.

$\eta_p{}^2 = 0.22$; and disgust, $F(2, 66) = 4.0$, $p = 0.022$, $\eta_p{}^2 = 0.11$. In particular, participants' performance was significantly reduced for older sad and angry faces compared to both young and middle-aged versions (ts > 2.9, ps < 0.006, $d = 0.70$) with no difference between young and middle-aged versions for either emotion (ts < 1.6, ps > 0.11, $d < 0.39$). Finally, for disgust, participants found middle-aged faces to be easiest, with significantly superior performance relative to the younger faces, $t(33) = 2.9$, $p = 0.006$, $d = 0.70$, and a trend for the same pattern for older exemplars, $t(33) = 1.8$, $p = 0.08$, $d = 0.43$. Figure 1A presents the percentage of correctly identified expressions by face age and expressed emotion (asterisks indicate significant differences, $p < 0.05$).

In sum, despite some differences as a result of the age of the expressed face and the particular emotional expression, younger and older adults performed equivalently well on the bubbles task in categorizing the emotional stimuli. However, it is important to interpret this result in the context of this particular paradigm. By its very design the bubbles approach strives to equate performance levels by modulating online the task difficulty (via the amount of information shown) to target a set performance level. It is therefore necessary to establish if equivalent performance across the participant age groups is a result of differences in the amount of visual information shown. To this end, we extracted the average amount of information shown (i.e., the average number of bubbles) for each experimental condition and ran a second mixed ($3 \times 5 \times 2$) ANOVA with the same factors.

Bubble task: Amount of information

As expected, in parallel with the accuracy results, we observed a main effect of stimulus age category, $F(2, 64) = 186$, $p < 0.001$, $\eta_p{}^2 = 0.85$) which reflected a relative increase in the amount of information that participants required to categorize older age faces ($M = 211$, $SD = 36$) compared to the younger and middle aged faces ($M = 154$, $SD = 37$; $M = 159$, $SD = 42$), $t[33] = 17.2, 14.8$, $p < 0.001$, $d = 4.17, 3.59$). We also observed a main effect of emotional expression, $F(3.04, 97.2) = 143$, $p < 0.001$, $\eta_p{}^2 = 0.82$, and a significant interaction of these two factors, $F(5.4, 172.8) = 26.3$, $p < 0.001$, $\eta_p{}^2 = 0.45$. Neither main effect interacted significantly with participant age group, $F < 0.3$, $p > 0.8$, $\eta_p{}^2 < 0.009$, but there was a three-way interaction of stimulus age, participant age, and emotional expression, $F(5.4, 172.8) = 2.4$, $p = 0.046$, $\eta_p{}^2 = 0.07$. Furthermore, we observed a main effect of participant age group, $F(1, 32) = 4.3$, $p = 0.046$, $\eta_p{}^2 = 0.12$, which indicated that older participants required significantly more visual information (i.e., more bubbles) in order to achieve performance levels equivalent to their younger counterparts (on average they needed 25 [$SD = 16$] more bubbles: $M = 162$ vs. 187, corresponding to around 15% more information from the face stimuli).[2]

To further explore the three-way interaction we considered each emotional expression in turn and conducted separate mixed (3×2) ANOVA with participant age as a between-subjects factor (young and old) and stimulus age group as a within-subjects factor (young, middle aged, and older). We observed a significant interaction of the two factors only for fearful and sad expressions (fear: $F[1.6, 52.1] = 3.8$, $p = 0.035$, $\eta_p^2 = 0.11$; sad: $F[2, 64] = 5.7$, $p = 0.005$, $\eta_p^2 = 0.1$; Fs < 1.83, ps > 0.18, $\eta_p^2 < 0.055$ otherwise). For fear there was no difference in information use for the young faces across the participant age groups, $t(32) = 1.5$, $p = 0.14$, $d = 0.53$, but a clear difference for older faces, $t(32) = 2.3$, $p = 0.027$, $d = 0.80$, with older participants requiring substantially more information, $M = 149$, $SD = 44$ vs. $M = 104$, $SD = 66$) and there was an indication of a similar trend for middle-aged faces, $t(32) = 1.7$, $p = 0.096$, $d = 0.59$.[3] For sadness, by contrast, older adults did not require more information for older sad or middle-aged faces, $t(32) = 0.74$, $p = 0.47$, $d = 0.25$; $t(32) = 1.6$, $p = 0.12$, $d = 0.55$. In fact this group used less information than their younger counterparts for the older sad faces ($M = 287$, $SD = 52$; $M = 302$, $SD = 69$). Figure 1B presents the amount of information required (number of bubbles) to correctly identify expressions by face age, expressed emotion, and participant age.

To summarize, older participants required more information to perform equivalently well on the bubbles task for the majority of emotional expressions and stimulus age categories. By design, the experimental paradigm works to standardize performance accuracy at a set level by modulating task difficulty in terms of the amount of information shown (the number of bubbles behind which expressive faces are presented). Despite performing to criterion (target level of 75% correct), the requirement for significantly more information by older adults indicates that they found the task more difficult and that their performance was worse than the younger participant group in almost every category. Exceptions to this included young faces displaying fearful expressions where their efficiency was on a par with younger adults, and for expressions of sadness on older and middle-aged faces where they actually required less information than their younger counterparts.

Bubbles task: Information utilization

In light of the task performance results we chose to selectively focus our attention on the information use results for the emotions where overall performance exceeded the target accuracy level (i.e., happy, fear, and sadness for young and middle-aged faces, anger for young faces).[4] For these conditions of interest, every trial was sorted as a function of whether or not the information presented to the participant resulted in a correct response in the expression categorization task. Observers tend to be correct if the information necessary to perform the task has been provided to them and inversely tend to be incorrect if this information is missing. As an example, to determine the specific information driving correct categorizations of fear, we summed together all of the information locations leading to correct categorizations of fear and subtracted from that the sum of all information locations leading to an incorrect response to generate classification images (equivalent to a least squares line fit on the data). These values were transformed into z scores using the nonface regions of the image space as a baseline. Regions statistically associated with correct categorization performance were determined by applying a $p < 0.05$ peak threshold alongside a $p < 0.05$ cluster extent threshold on these probabilities (see Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005, for the specially designed statistical tests that compensate for multiple comparisons in the image space).

Figure 2, Panel A visually represents the result of these statistical tests to depict the critical visual information used on a sample face image for the categorization of happy, fearful, sad, and angry emotional expressions. These figures allow us to view and compare the profiles observed in each participant group across the various stimulus age categories. These effective faces indicate the combined facial information that is statistically associated with correct categorization performance and they are the combination of the significant regions from a representative face image at each spatial scale.[5]

It is immediately clear from the considerable overlap in information use that younger and older participants are using very similar critical visual information to categorize happy and fearful faces. For *happy* expressions, it is apparent that the smiling mouth drives correct categorizations across both observer groups across spatial frequencies—that is, from fine detail around the teeth, to the broad outline of the smile and associated facial contours. For *fearful* expressions, as expected, the important role of the wide-open eyes in higher spatial frequencies is confirmed for both young and older observers across the different stimulus age categories (see also Supplementary Figure S1, for the detailed SF breakdown). At lower spatial frequencies, the fearful mouth cues are also important for both sets of observers. Young observers differ from older observers only in that they also make more use of the flared nostril for middle-aged and older faces, while

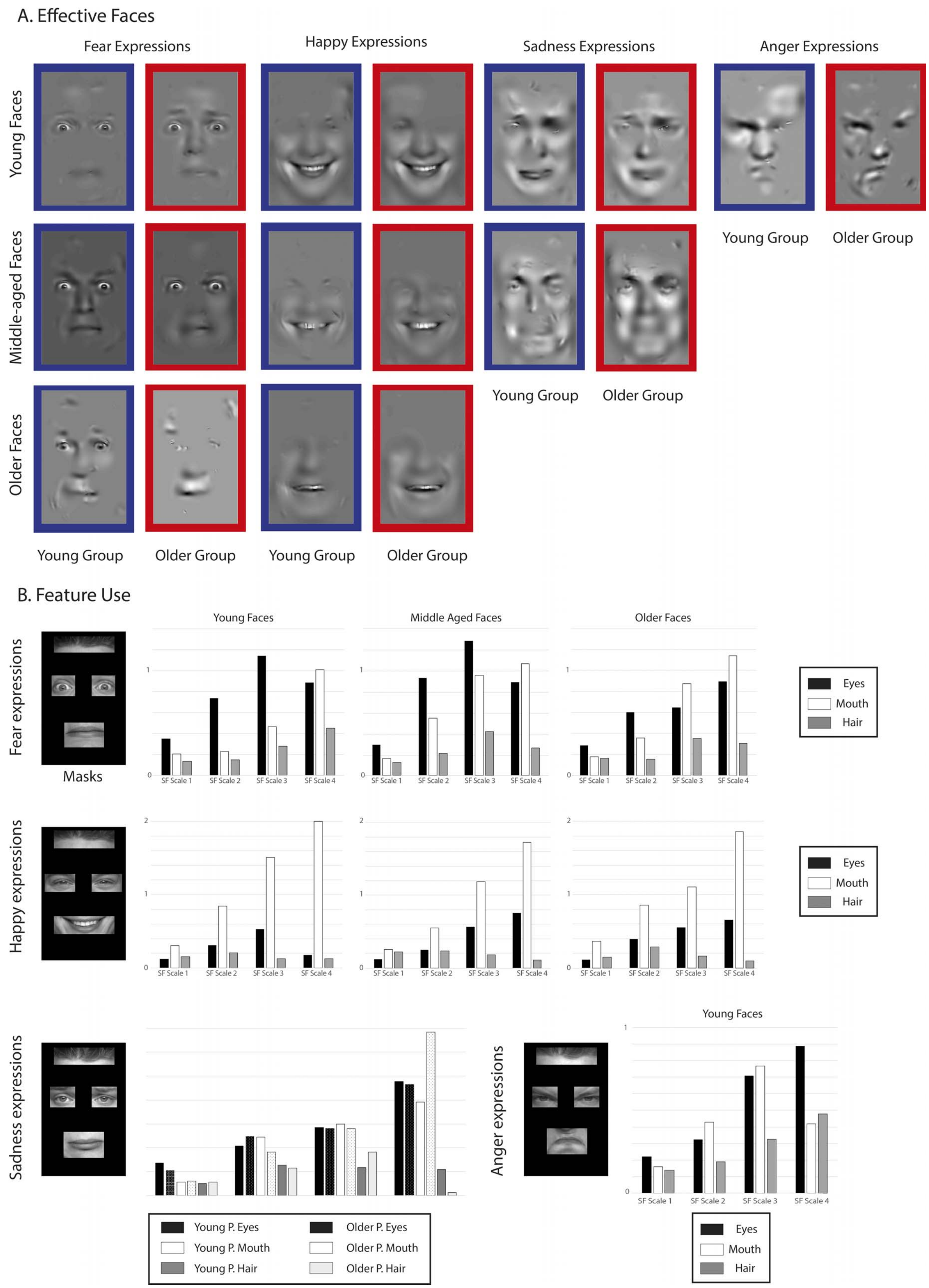

Figure 2. (A) Effective faces depicting the significant visual information used ($p < 0.05$ corrected) by younger observers (blue background) and older observers (red background) for each of the aged expressive faces. Effective faces comprise the combination of the regions in each SF scale that were significantly correlated with correct categorization performance, displayed on a sample face from the stimulus set. (B) Feature masks used to assess the degree to which individual participants use specific key features. Feature →

←
masks cover two key visual features (the eyes and the mouth) and a third baseline region (the hair). Within an expression, all masks have the same area. Bar charts: Average z scores in each of the chosen regions as a function of spatial scale for fear (top row), happiness (middle row), sadness (bottom left), and anger (bottom right) categorizations. The faces shown here are illustrative, example identities from the database used in the experiment (FACES; Ebner et al., 2010)

older observers only use this cue for the younger fearful faces. Similarly, there is a high degree of similarity in information use for *sadness* and *anger* expression categorizations, with both groups of participants using the furrowed brow and taut mouth when categorizing anger, and the downturned eyebrows and mouth for sadness. It's also worth noting that for both anger and sadness categorizations, observers make more use of the expressive eye on the left side of the face, than on the right. This is in line with studies that describe a bias for negative emotional information on the left (projecting primarily to the right hemisphere of the brain; e.g., Indersmitten & Gur, 2003; Jansari, Tranel, & Adolphs, 2000).

To formalize the similarities (and differences) in information use by the older and younger participants we employed the structural similarity index (SSIM), a popular technique that is used routinely to measure the low level visual similarity between two images (Wang, Bovik, Sheikh, & Simoncelli, 2004). SSIM values can range from −1, indicating they are completely dissimilar, to 1 indicating an identical perfect match. We used the SSIM index to compare the z scored classification images of the old and young observers for each emotion and stimulus facial age at each SF band. Note that the z scored classification images were used in their original form, before applying any significance threshold, to ensure similarities or differences were not artificially enhanced by the application of the threshold. SSIM values confirmed the numerical equivalence in information use across the two groups as shown in Table 1, (confidence intervals presented below each value in brackets)[6], with high overall values for fearful (SSIM = 0.72, s = 0.05), happy (SSIM = 0.75, s = 0.03), sadness (SSIM = 0.74, s = 0.01), and anger (SSIM = 0.76) categorizations. We note the relatively lower values (<0.5) for young fear faces in SF scale 3, which is driven by only older adults using the eyebrows alongside the wide-open eyes, and for middle aged fear faces in SF scale 4, which is driven by a use of the left eye in younger participants and a more diffuse use of the mouth cues in the older participants (see Supplementary Figure S1 for clear visualization of these differences).

We also employed the SSIM index to establish the consistency of information use within the participant groups as the age of the stimulus varied. We compared the z scored classification images (prior to application of the threshold) for each participant group evaluating the younger, middle-aged, and older versions of the same facial expressions. For example, we contrasted the classification images of young observers viewing young happy faces with those of young observers viewing middle-aged happy faces etc. Table 2 presents the SSIM indices for all possible comparisons averaged across the SF bands. Consistency is once again high for both participant groups (young: $M = 0.71$, $SD = 0.056$; older: $M = 0.73$, $SD = 0.027$).

	Spectral frequency band				
Model	1	2	3	4	Average
Fearful					
Young faces	0.80 [0.76, 0.83]	0.73 [0.67, 0.82]	0.48 [0.44, 0.78]	0.67 [0.19, 0.81]	0.67
Middle-aged faces	0.84 [0.79, 0.85]	0.71 [0.64, 0.82]	0.80 [0.61, 0.86]	0.47 [0.17, 0.82]	0.71
Older faces	0.78 [0.76, 0.82]	0.77 [0.67, 0.81]	0.82 [0.65, 0.85]	0.72 [0.32, 0.82]	0.77
Happy					
Young faces	0.83 [0.78, 0.86]	0.78 [0.7, 0.84]	0.70 [0.6, 0.80]	0.57 [0.4, 0.81]	0.72
Middle-aged faces	0.81 [0.79, 0.84]	0.77 [0.7, 0.82]	0.75 [0.52, 0.83]	0.71 [0.43, 0.83]	0.76
Older faces	0.87 [0.81, 0.88]	0.78 [0.71, 0.83]	0.74 [0.63, 0.83]	0.69 [0.37, 0.82]	0.77
Sad					
Young faces	0.78 [0.74, 0.81]	0.72 [0.64, 0.78]	0.69 [0.53, 0.79]	0.76 [0.31, 0.85]	0.74
Middle aged faces	0.84 [0.79, 0.85]	0.78 [0.71, 0.84]	0.74 [0.44, 0.78]	0.69 [0.53, 0.89]	0.75
Anger					
Young faces	0.80 [0.75, 0.82]	0.77 [0.71, 0.81]	0.75 [0.51, 0.82]	0.72 [0.32, 0.78]	0.76

Table 1. Similarity (structural similarity index [SSIM]) in the use of information between younger and older adults by the age and facial expressions of the models for each spectral frequency band. Confidence intervals (95%) are presented below in braces. *Note*. Similarity scores can range from −1 to +1.

	Younger adults			Older adults		
Expression	Y vs. M	Y vs. O	M vs. O	Y vs. M	Y vs. O	M vs. O
Happy	0.71 [0.67, 0.78]	0.74 [0.68, 0.79]	0.77 [0.67, 0.82]	0.70 [0.66, 0.78]	0.70 [0.67, 0.80]	0.77 [0.66, 0.81]
Fearful	0.62 [0.58, 0.74]	0.65 [0.61, 0.74]	0.76 [0.65, 0.77]	0.73 [0.62, 0.82]	0.73 [0.64, 0.79]	0.76 [0.65, 0.79]
Sadness	0.72 [0.63, 0.76]			0.72 [0.62, 0.78]		

Table 2. Pair-wise comparison in the similarity (structural similarity index [SSIM]) in the use of information between the young (Y), middle-aged (M), and older (O) models separately for each age group and emotional expression.

Bubbles task: Feature use

To further formalize the use of visual information by younger and older adults across the different stimulus age categories, and highlight any group differences, we computed individual classification images for each observer and submitted them to an additional standardized analysis. We selected two key visual features for the emotional faces (the eyes and the mouth), and a third uninformative hairstyle region to act as a control (see Figure 2, Panel B).[7] For each observer, we then extracted the average z scores from their classification image in each of these regions. Higher values in a particular region indicate that information in that region is more closely linked to correct categorization performance.[8] These values were then submitted to a series of mixed ANOVAs (one for happy, fear, sadness, and anger) with stimulus age (happy, fear expressions: three levels: young, middle aged, older; sad expressions: two levels: young and middle-aged; anger: one level: young), feature (three: eyes, mouth, and hair) and SF band (four: fine to coarse) as within-subjects factors and participant age group as the between-subjects factor.

For *fearful* face categorizations, no main effect of participant age group or interaction was observed (all Fs < 1.1, ps > 0.33, $\eta_p^2 < 0.034$). We did, however, observe a significant interaction of stimulus age with facial feature, $F(4, 128) = 3.1$, $p = 0.017$, $\eta_p^2 = 0.09$, alongside a trend for a main effect of stimulus age, $F(2, 64) = 2.5$, $p = 0.09$, $\eta_p^2 = 0.07$. Follow up t tests confirmed that this interaction was driven by significantly greater use of the eyes than the mouth for both younger and middle-aged fearful faces, $t(33) = 3.99$, $p < 0.001$, $d = 0.95$; $t(33) = 2.1$, $p = 0.04$, $d = 0.51$, with no such difference for older fearful faces, $t(33) = -0.3$, $p = 0.75$, $d = 0.07$. As expected, we also observed main effects of facial feature, $F(2, 64) = 43$, $p < 0.001$, $\eta_p^2 = 0.57$, and SF scale, $F(1.9, 62.6) = 39.1$, $p < 0.001$, $\eta_p^2 = 0.55$, as well as a significant interaction between both, $F(2.6, 83) = 7.04$, $p < 0.001$, $\eta_p^2 = 0.18$. Follow-up tests confirmed that overall, participants used the eyes and the mouth regions significantly more than the baseline hair region, $t(27) = 10.9$, 5.4, $p < 0.001$, $d < 1.31$, thus confirming their importance for the task and that participants used the eyes more than the mouth at higher spatial frequencies (SF Bands 1–3: $t[33] = 5.5$, 7.4, 3.0; $p < 0.005$, $d < 0.73$, respectively) but not at lower spatial frequencies where their use did not differ (SF Band 4: $t[33] = -1.1$, $p = 0.27$, $d = 0.27$). Figure 2, Panel B summarizes the use of each facial region at the spatial scales for each stimulus age group.

For *happy* expressions, we selected the same three regions (around the eyes, mouth, and noninformative hair region) and ran the ANOVA as before. We found no significant main effect of participants' age group, $F(1, 32) = 0.52$, $p = 0.48$, $\eta_p^2 = 0.016$, or significant interaction, $F < 1.4$, $p > 0.2$, $\eta_p^2 < 0.043$, beyond a marginal trend to interact with facial feature, $F(2, 64) = 2.3$, $p = 0.11$, $\eta_p^2 = 0.066$, driven by a trend for more use of the mouth by older adults than younger, $t(32) = 1.7$, $p = 0.09$, $d = 0.59$, with no difference in the eyes or hair region, $t < 0.7$, $p > 0.5$, $d < 0.24$. As expected we again found main effects of feature, $F(2, 52) = 183$, $p < 0.001$, $\eta_p^2 = 0.85$, and scale, $F(2.27, 72.6) = 52.4$, $p < 0.001$, $\eta_p^2 = 0.62$, and a significant two-way interaction, $F(3.5, 111.8) = 39$, $p < 0.001$, $\eta_p^2 = 0.55$, which did not interact further with participant age or model age (Fs < 1.55, ps> 0.16, $\eta_p^2 < 0.046$). The main effect of feature confirmed significantly more use of the mouth than either the eyes, $t(33) = 15.7$, $p < 0.001$, $d = 3.81$, or the hair region, $t(33) = 16.5$, $p < 0.001$, $d = 4.01$, with the eyes more useful than the hair, $t(33) = 4.1$, $p < 0.001$, $d = 0.99$. The interaction mediated these effects by indicating that for the highest two spatial scales (SF Bands 1 and 2), the eyes are not any more useful than the control hair region. Finally, we observed a significant interaction of stimulus age and facial feature, $F(4, 128) = 3.6$, $p = 0.008$, $\eta_p^2 = 0.12$, which was driven by significantly less use of the eyes than the hair region for middle aged faces in the first SF scale, $t(33) = -3.6$, 0.001, $d = 0.87$.

For *sadness,* we again observed main effects of facial feature, $F(1.7, 55.5) = 26.9$, $p < 0.001$, $\eta_p^2 = 0.48$; SF scale, $F(1.8, 58.8) = 25.6$, $p < 0.001$, $\eta_p^2 = 0.45$; and their interaction, $F(2.6, 84) = 7.6$, $p < 0.001$, $\eta_p^2 = 0.19$. Furthermore there was a trend for the three-way interaction of facial feature, SF scale, and participant age, $F(2.6, 84) = 2.29$, $p = 0.09$, $\eta_p^2 = 0.067$. The main effect of feature once again indicated that there was significantly more use of the eyes and the mouth regions than the hair baseline, $t(33) = 7.6$, 4.3, $p <$

0.001, $d > 1.04$, with no overall difference between the eyes and mouth, $t(33) = 0.23$, $p = 0.82$, $d = 0.05$. SF scale modulated this pattern with the eyes significantly more used than the mouth in SF Band 1, $t(33) = 6$, $p < 0.001$, $d = 1.46$, and no difference between the mouth and the hair region, $t(33) = 0.36$, $p = 0.72$, $d = 0.09$. The further interaction trend indicated more use of the eyes than the mouth in older participants for SF Scale 2 while the younger participants showed the reverse pattern (though neither difference was significant, Older: $t(16) = 1.8$, $p = 0.09$, $d = 0.62$; Younger: $t(16) = -0.9$, $p = 0.36$, $d = 0.31$.

Finally, for *anger* categorizations, the ANOVA again confirmed a main effect of facial feature, $F(2, 64) = 6.3$, $p = 0.03$, $\eta_p^2 = 0.16$, and a main effect of SF scale, $F(2, 64) = 8.8$, $p < 0.001$, $\eta_p^2 = 0.21$, and a trend for their interaction, $F(3.3, 106.6) = 2.1$, $p = 0.1$, $\eta_p^2 = 0.06$. However, there was no main effect of participant age group, $F(1, 32) = 1.6$, $p = 0.21$, $\eta_p^2 = 0.05$, or interactions, $F < 0.6$, $p > 0.6$, $\eta_p^2 < 0.02$. Overall both the furrowed eyes and taut mouth were more useful than the baseline region, $t(33) = 3.3$, 2.4, $p = 0.002$, 0.02, $d = 0.80$, 0.58 respectively, but did not differ themselves, $t(33) = 1.3$, $p = 0.21$, $d = 0.31$. This was modulated by the effect of SF scale whereby the mouth was used significantly less than the eye region in the coarsest SF Band 4, $t(33) = 1.94$, $p = 0.06$, $d = 0.47$.

Discussion

Relative to younger adults, older adults tend to perform relatively more poorly in identifying the correct emotional label for static facial expressions. The goal of the present study was to investigate one potential mechanism that might explain these performance differences; in particular, we investigated age differences in the utilization of visual information when decoding the emotional expression in young, middle-aged, and older faces for happy, fearful, angry, disgusted, and sad expressions. Results indicated that there were no significant group differences in performance accuracy in the bubbles task between the younger and older participant groups (despite expected variability in categorization performance as a function of the expression shown and the age of the model face); however, older adults required more visual information to perform at equivalent accuracy levels to the younger adults group. This is in line with lifespan models of neuromodulation (Li, Brehmer, Shing, Werkle-Bergner, & Lindenberger, 2006) suggesting that older adults' neural network may require more information to differentiate stimuli. Crucially, for the first time we show also that for the most part, younger and older adults were also very similar in using the same visual features to extract the correct emotional expression (disgust, older sadness, middle-aged and older anger excluded). Furthermore, both younger and older adults tended to be consistent in the information that they used to extract emotional information irrespective of the age of the transmitter for fearful, happy, and sad faces.

Performance was best and equivalent for happy and fearful expressions for all stimulus age categories. Fear was consistently categorized by means of the characteristic wide-open eyes across the high- to mid-SF bands, and the mouth in lower SF bands. This pattern was similar across all stimulus age groupings and present irrespective of participants' age. The importance of the eyes is directly in line with past research using the bubbles approach to explore the information used in facial expression categorization by healthy young adult participants (e.g., Smith et al., 2005; Schyns et al., 2007). Fearful eye whites selectively activate the amygdala, even outside of conscious awareness (Morris, deBonis, & Dolan, 2002; Whalen et al., 2004), and abnormal processing of this feature has been shown to occur in line with deficits in fear perception in an individual with bilateral amygdala lesions (Adolphs et al., 2005) and individuals with other developmental disorders (Aspergers syndrome: see Corden, Chilvers, Skuse, 2008; conduct disorder: see Dadds, Masary, Wimalaweera, & Guastella, 2008). Given the evolutionary importance of fearful face cues, and the reported functional role of specific expressive changes (e.g., wide-open eyes resulting in a greater field of view; Susskind et al., 2008), it is perhaps not surprising that these cues remain a consistent indicator of the underlying emotion irrespective of age (in the stimulus or in the perceiver).

Also consistent with past research (Smith et al., 2005; Smith & Schyns, 2009), happiness categorizations were primarily driven by the distinctive information changes created by the wide-open smiling mouth across all SF bands. Similarly to fearful eye whites, the wide-opened smile may also be an evolutionarily important social signal, a key distal indicator of a willingness to cooperate (Smith & Schyns, 2009), driving approach behavior between individuals (Schmidt & Cohn, 2001). Few studies find age-related deficits in the categorization of happiness from expressive faces, and indeed the reverse has even been observed, with older adults biased to see happiness, the so-called positivity effect (see Mather & Carstenen, 2005, for a review; Ebner & Johnston, 2010). Although in the current study we found older and younger observers used the same visual cues to correctly categorize happiness, there was an indication that older adults relied upon mouth cues to a significantly greater degree.

SSIM indices—indicating the similarity in information use—were high across all expressions (bar disgust where poor performance meant that no solution was possible). Where the similarity in information use did show signs of beginning to differ between age groups was for the specific case of young and middle-aged faces displaying fear. While both young and older adults made use of the eye information for this judgment, the use of a wider region surrounding the eyes at mid SFs (e.g., to include the eyebrows) was specific to older adults. Older adults also made more use of more diffuse mouth cues at coarser levels in middle-aged faces. An overreliance on solely the fearful eye whites for categorizing fear has been observed before in young individuals categorizing young fearful faces in the context of a seven-alternative forced-choice task (Smith et al., 2005). A recent study offers a possible explanation, in that fearful mouth cues vary in their importance for categorization performance depending on the extent to which they are useful for the particular experimental scenario (Smith & Merlusca, 2015). That is, lower SF mouth cues are particularly drawn on only when they are especially informative for the task being undertaken (e.g., when comparing fear to a single other emotion, such as happiness, as opposed to comparing fear to all other basic emotional expression categories where the mouth feature may be a less reliable cue).

For the remaining emotions (disgust, anger, and sadness) there was no difference in performance between the young and older participants, but the age of the transmitter did affect performance. For sadness, performance for younger and middle-aged faces was equivalent, with a clear drop for the older faces. Anger categorization became progressively more difficult with age of the stimulus with middle-aged faces categorized less well than younger faces, and older faces significantly worse again. This drop in performance with stimulus age has been reported previously (e.g., Ebner et al., 2010; Folster et al., 2014; Hess et al., 2012) and may in part be due to reduced potency of the expressions in older faces. The presence of wrinkles around key features in the face may weaken their impact and create confusing countersignals to the expression-specific signals (e.g., wrinkling around the nose and mouth is associated with disgust, creases in the center of the forehead is associated with anger; Smith et al., 2005). Disgust categorizations were difficult for all stimulus ages, though notably better for middle-aged faces. This is in accordance with studies reporting disgust to be among the most difficult expression to categorize (Ruffman et al., 2008). In all other cases, when a bubbles solution was possible, we identified patterns of information use that were similar across participant groups (as evidenced by the high SSIM values) and in line with previously established face cues.

Our feature analysis for the eyes and the mouth supported the finding of very similar visual information use in categorizing emotions for younger and older adults. Where differences did appear (e.g., happy, sadness) they did not reach statistical significance. Trends indicated a stronger reliance on the mouth for older adults than younger adults when categorizing happiness, and opposing strategies for using higher SF information when categorizing sadness (more use of the eyes than the mouth in older adults, with a reverse pattern in younger adults). Stimulus age drove significant differences in information use only for fearful face categorizations, with the eyes dominating information use for younger and middle-aged faces, but equivalent use of the eyes and mouth for the older aged faces.

Though the current study is the first to directly contrast information use during expression categorization judgments, the apparent absence of differences in face-processing strategies between younger and older adults is not without precedent. In their eye-tracking study, Ebner, He, and Johnson (2011) similarly did not find age-related differences in looking time to different parts of expressive faces (upper vs. lower regions) when accurately categorizing them. Furthermore, it is possible that methodological differences might in part account for those instances in which age-related gaze differences have been observed (e.g., Wong et al., 2005, involved free viewing of expressive faces rather than an active categorization task). Perhaps importantly, where gaze differences have been observed in an active task (e.g., Sullivan et al., 2007), the nature of the difference is in the relative amount of time spent looking at different face regions, with older observers spending relatively less (but still considerable time) on the upper portions of the face. The more precise bubbles solution from the current study confirms that both older and younger adults can, and do, use facial cues from across the face in making facial expression categorizations. There is also no indication that older adults selectively use information differently across SF bands (Govenlock, Taylor, Sekuler, & Bennett, 2010). Instead, we observed them making use of exactly the same high frequency cues as younger adults for both happy and fear decisions.

By setting the task difficulty at a fairly low level (targeting accuracy at 75% correct), it may be that we did not have sufficient scope to observe subtle differences in processing ability for the "easier" emotion categories (young faces displaying most emotions, middle-aged and older faces displaying happiness and fear). Older adults do typically perform above these levels, with some reports suggesting few deficits in performance accuracy during emotion categorization tasks (e.g., Ebner, He, & Johnson, 2011; Ebner, Johnson, & Fischer, 2012; Krendl & Ambady, 2010; Sullivan et al., 2007, Expt 2), finding slower reaction times to be more reliable indicators of deficits (Ebner, He, & Johnson, 2011; Ebner, Johnson, & Fischer, 2012). A more challenging task may yet establish the presence of subtler differences in information use between older and younger participants.

Conclusions and future avenues

Given that older adults used very similar facial cues to decode emotional expressions as successfully as younger adults, the findings of the present study would suggest that differences in the ability to use visual information from across the face are not likely a major contributor to reported age differences in successful emotion recognition. However, more research is needed to replicate and extend the current findings. For example, in the present study, we used images of strangers for the participants to decode the emotional expression. There are, however, theoretical models (Castro et al., 2016) that predict differences in emotion recognition abilities for known others—such as family members—and unknown others (e.g., strangers). Older adults reported, for example, more similar emotional reactions to a known other person than younger adults (Cheng & Grühn, 2016). This might indicate that older adults may use other knowledge sources in everyday life—based on years of experience—to decode emotional expressions of close others. Thus, one could speculate that older adults might be selectively more efficient in utilizing visual information in faces from known persons. Future research might benefit from disentangling the effect of decoding emotional expressions from known and unknown persons.

It is also possible that the nature of the paradigm (representing only parts of the stimulus) atypically highlighted and directed older adults' attention towards features that they would not normally have used. Previous studies with the bubbles paradigm and brain-damaged individuals speak against this argument. Adolphs et al. (2005), for example, showed that during a similar task a patient with bilateral amygdala damage did not use the eye information to categorize fearful face expressions and consequently could not accurately categorize fear (eye-tracking and bubbles results). However, when instructed to attend to the eye information explicitly they did just that, and their performance improved accordingly. Had the nature of the bubbles trials been sufficient to draw attention towards the eyes that would not normally have been directed there, no such effects would have been observed. Similarly, Caldara et al. (2005), observed a bias away from the eyes in an individual with prosopagnosia. That said, the bubbles paradigm might have triggered the development of specific strategies in performing the task. Future research could benefit from investigating the impact of strategy use by manipulating instructions for the task.

Finally, the current study used highly standardized stimuli to generate the facial maps of information usage. We acknowledge that in real life, faces rarely pose emotional expressions in such a standardized way, as static images and without other contextual cues that could facilitate recognition. Nevertheless the present results are interesting and important because our tightly controlled paradigm allowed us to reveal, for the first time, that despite general consensus to the contrary, older adults are able to extract and use facial cues from the eye region and do use them when categorizing facial expressions of emotion. Indeed our face maps reveal that they do so in a manner identical to that of their younger counterparts. These results essentially rule out differences in information use as an explanation for natural age-related declines in expression categorization abilities and direct future research to alternative accounts.

Keywords: facial expressions, emotion recognition, emotion, age differences, healthy aging

Acknowledgments

MLS and LE were supported by grant RPG-2013-019 from the Leverhulme Trust awarded to MLS, and additional support from Birkbeck institutional support funding. The authors wish to thank all of their participants for volunteering their time and effort to completing this study.

Commercial relationships: none.
Corresponding author: Louise Ewing.
Email: l.ewing@uea.ac.uk.
Address: University of East Anglia, School of Psychology, Norwich, UK.

Footnotes

[1] Note that our experimental paradigm deliberately set out to match performance accuracy across all conditions, making use of a gradient descent algorithm to control the amount of information shown. As a result, significant differences in performance across participant age groups is not expected.

[2] When entering processing speed as a covariate, the main effect of age was no longer significant, $F(1, 33) = 2.2$, $p = 0.146$, $\eta_p{}^2 = 0.067$, indicating that processing speed may be a potential explanatory variable for the age difference in information use. However, a larger sample with a continuous age range is necessary to appropriately disentangle potential mediators for the age effect.

[3] Using processing speed as a covariate, there remains a clear trend for the effect of age group when viewing older age fear faces, $F(1, 33) = 3.3$, $p = 0.079$, $\eta_p^2 = 0.096$.

[4] For the remaining expressions (disgust all face ages, sadness in older faces, anger in middle-aged and older faces) where performance often did not reach this target level, the nature of the staircase algorithm is such that the amount of information presented is high for most of the experiment. In this situation obtaining a bubbles solution is difficult because considerably more power (i.e., more experimental trials) is required to weigh the relative importance of information from different facial features when most of the face is revealed on each trial. For instance, the average number of presented bubbles for conditions where 75% criterion was reached and conditions where 75% criterion was not achieved was 118 ($SD = 54$) versus 263 ($SD = 28$), respectively.

[5] See Supplementary Figures S1 and S2 for a detailed depiction of the regions associated with correct performance across each SF scale.

[6] 95% confidence intervals were computed via a 100-iteration bootstrap, where participants were randomly sampled (with replacement) to contribute their data to the overall bubbles solution, which was then submitted as before to the SSIM algorithm.

[7] Masks were generated to encompass the key features in isolation, separately for each expression and face age. The total area of each mask is equivalent within an expression (i.e., for all fearful faces the size of the eyes mask matches the size of the mouth mask, etc.).

[8] Note that as the individual participant data is necessarily noisier applying statistical tests to threshold individual classification images may lead to misleading results. We therefore choose to analyze the unthresholded z scores without thresholding first.

References

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

Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *28*, 23–32.

Bailey, I. L., & Lovie, J. E. (1976). New design principles for visual acuity letter charts. *American Journal of Optomology and Physiological Optics*, *53*, 740–745.

Calder, A. J., Keane, J., Young, A. W., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 527–551.

Caldara, R., Schyns, P., Mayer, E., Smith, M. L., Gosselin, F., & Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of Cognitive Neuroscience*, *17*, 1652–1666, http://dx.doi.org/10.1162/089892905774597254.

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, *54*, 165–181.

Castro, V. L., Cheng, Y., Halberstadt, A. G., & Grühn, D. (2016). EUReKA! A conceptual model of emotion understanding. *Emotion Review*, *8*, 258–268, https://doi.org/10.1177/1754073915580601.

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, *5*(9):1, 659–667, https://doi.org/10.1167/5.9.1. [PubMed] [Article]

Cheng, Y., & Grühn, D. (2016). Perceived similarity in emotional reaction profiles between the self and a close other as a predictor of emotional well-being. *Journal of Social and Personal Relationships*, *33*, 711–732, https://doi.org/10.1177/0265407515590278.

Corden, B., Chilvers, R., & Skuse, D. (2008). Avoidance of emotionally arousing stimuli predicts social–perceptual impairment in Asperger's syndrome. *Neuropsychologia*, *46*, 137–147, http://dx.doi.org/10.1016/j.neuropsychologia.2007.08.005.

Dadds, M. R., El Masry, Y., Wimalaweera, S., & Guastella, A. J. (2008). Reduced eye gaze explains "fear blindness" in childhood psychopathic traits. *Journal of the American Academy of Child & Adolescent Psychiatry*, *47*, 455–463, http://dx.doi.org/10.1097/CHI.0b013e31816407f1.

Ebner, N. C., He, Y. I., & Johnson, M. K. (2011). Age and emotion affect how we look at a face: Visual scan patterns differ for own-age versus other-age emotional faces. *Cognition & Emotion*, *25*, 983–997, https://doi.org/10.1080/02699931.2010.540817

Ebner, N. C., & Johnson, M. K. (2009). Young and older emotional faces: Are there age-group differences in expression identificiation and memory? *Emotion*, *9*, 329–339, https://doi.org/10.1037/a0015179.

Ebner, N. C., & Johnson, M. K. (2010). Age-group differences in interference from young and older emotional faces. *Cognition & Emotion*, *24*, 1095–1116, https://doi.org/10.1080/02699930903128395.

Ebner, N. C., Johnson, M. K., & Fischer, H. (2012). Neural mechanisms of reading facial emotions in young and older adults. *Frontiers in Psychology*, *3*, 223, https://doi.org/10.3389/fpsyg.2012.00223.

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, https://doi.org/10.3758/BRM.42.1.351.

Folster, M., Hess, U., & Werheid, K. (2014). Facial age affects emotional expression decoding. *Frontiers in Psychology*, *5*, 30, https://doi.org/10.3389/fpsyg.2014.00030.

García-Rodríguez, B., Fusari, A., Fernández-Guinea, S., Frank, A., Antonio Molina, J., & Ellgring, H. (2011). Decline of executive processes affects the identification of emotional facial expressions in aging. *Current Aging Science*, *4*, 70–75, http://dx.doi.org/10.2174/1874609811104010070.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, *41*, 2261–2271.

Govenlock, S. W., Taylor, C. P., Sekuler, A. B., & Bennett, P. J. (2010). The effect of aging on the spatial frequency selectivity of the visual system. *Vision Research*, *17*, 1712–1719, https://doi.org/10.1016/j.visres.2010.05.025.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*, 59–67.

Hess, U., Adams, R. B., Simard, A., Stevenson, M. T., & Kleck, R. E. (2012). Smiling and sad wrinkles: Age-related changes in the face and the perception of emotions and intentions. *Journal of Experimental Social Psychology*, *48*, 1377–1380, https://doi.org/10.1016/j.jesp.2012.05.018.

Indersmitten, T., & Gur, R. C. (2003). Emotion processing in chimeric faces: Hemispheric asymmetries in expression and recognition of emotions. *Journal of Neuroscience*, *1*, 3820–3825.

Isaacowitz, D. M., Loeckenhoff, C., Lane, R., Wright, R., Sechrest, L., Riedel, R., ... Costa P. T. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging*, *22*, 147–259, https://doi.org/10.1037/0882-7974.22.1.147.

Isaacowitz, D. M., & Stanley, J. T. (2011). Bringing an ecological perspective to the study of aging and recognition of emotional facial expressions: Past, current, and future methods. *Journal of Nonverbal Behavior*, *35*, 261–278, https://doi.org/10.1007/s10919-011-0113-6.

Jansari, A., Tranel, D., & Adolphs, R. (2000). A valence-specific lateral bias for discriminating emotional facial expressions in free field. *Cognition and Emotion*, *14*, 341–353, https://doi.org/10.1080/026999300378860.

Krendl, A. C., & Ambady, N. (2010). Older adults' decoding of emotions: Role of dynamic versus static cues and age-related cognitive decline. *Psychology and Aging*, *25*, 788–793, https://doi.org/10.1037/a0020607.

Li, S., Brehmer, Y., Shing, Y. L., Werkle-Bergner, M., & Lindenberger, U. (2006). Neuromodulation of associative and organizational plasticity across the lifespan: Empirical evidence and neurocomputational modeling. *Neuroscience and Biobehavioral Reviews*, *30*, 775–790.

Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Science*, *9*, 496–502, https://doi.org/10.1016/j.tics.2005.08.005.

Morris, J. S., deBonis, M., & Dolan, R. J. (2002). Human amygdala response to fearful eyes. *NeuroImage*, *17*, 214–222.

Murphy, N. A., & Isaacowitz, D. M. (2010). Age effects and gaze patterns in recognizing emotional expressions: An in-depth look at gaze measures and covariates. *Cognition and Emotion*, *24*, 436–452, https://doi.org/10.1080/02699930802664623.

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*, 1–72.

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*, *138*, 146–74, https://doi.org/10.1037/a0025750.

Ruffman, T. (2011). Ecological validity and age-related change in emotion recognition. *Journal of Nonverbal Behavior*, *35*, 297–304, https://doi.org/10.1007/s10919-011-0116-3.

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews*, *32*, 863–881, https://doi.org/10.1016/j.neubiorev.2008.01.001.

Schmidt, K. L., & Cohn, J. F. (2001). Human facial expressions as adaptations: Evolutionary questions in facial expression research. *American Journal of Physical Anthropology*, *33*, 3–24.

Schyns, P. G., Petro, L., & Smith, M. L. (2007). Dynamics of visual information integration in the brain to categorize facial expressions. *Current Biology*, *17*, 1580–1585, https://doi.org/10.1016/j.cub.2007.08.048.

Simoncelli, E. P. (1999). *Image and multi-scale pyramid tools* [Computer software]. Retrieved from ftp://ftp.cis.upenn.edu/pub/eero/matlabPyrTools.tar.gz

Shipley, W. C., Gruber, C. P., Martin, T. A., & Klein, A. M. (2009) *Shipley-2 Manual*. Los Angeles, CA: Western Psychological Services.

Smith, M. L., Cesana, M. L., Farran, E., Karmiloff-Smith, A., & Ewing, L. (2017). A 'spoon full of sugar' helps the medicine go down: How a participant friendly version of a psychophysics task significantly improves task engagement, performance and data quality in a typical adult sample. *Behavior Research Methods*, *50*, 1011–1019, https://doi.org/10.3758/s13428-017-0922-6.

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*, 184–189, https://doi.org/ 10.1111/j.0956-7976.2005.00801.x.

Smith, M. L., & Merlusca, C. (2015). How task shapes the uses of information during facial expression categorizations. *Emotion*, *14*(3), 478–487, https://doi.org/10.1037/a0035588.

Smith, F. W., & Schyns, P. G. (2009). Smile through your fear and sadness. *Psychological Science*, *20*, 1202–1208, https://doi.org/ 10.1111/j.1467-9280.2009.02427.

Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *Journal of Gerontology: Psychological Sciences*, *62B*, P53–P60.

Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., & Anderson, A. K. (2008). Expressing fear enhances sensory acquisition. *Nature Neuroscience*, *11*, 843–850, https://doi.org/10.1038/nn.2138.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, *6*, 624–631, https://doi.org/ 10.1038/nn1057.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*, 600–612, http://doi.org/10.1109/TIP.2003.819861.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale—Revised (WAIS–R)*. New York: Psychological Corporation.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., … Johnstone, T. (2004, December 17). Human amygdala responsivity to masked fearful eye whites. *Science*, *306*, 2061, https://doi.org/10.1126/science.1103617.

Wong, B., Cronin-Golomb, A., & Neargarder, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology*, *19*, 739–749, https://doi.org/10.1037/0894-4105.19.6.739.