Compensating for age limits through emotional crossmodal integration

Laurence Chaby [1,2]*, Viviane Luherne-du Boullay [3], Mohamed Chetouani [2] and Monique Plaza [2]

[1] Institut de Psychologie, Sorbonne Paris Cité, Université Paris Descartes, Boulogne-Billancourt, France, [2] Groupe Intégration Multimodale, Interaction et Signal Social, Institut des Systèmes Intelligents et de Robotique, CNRS UMR 7222, Paris, France, [3] Université Paris 8 Vincennes Saint Denis, Saint-Denis, France

Social interactions in daily life necessitate the integration of social signals from different sensory modalities. In the aging literature, it is well established that the recognition of emotion in facial expressions declines with advancing age, and this also occurs with vocal expressions. By contrast, crossmodal integration processing in healthy aging individuals is less documented. Here, we investigated the age-related effects on emotion recognition when faces and voices were presented alone or simultaneously, allowing for crossmodal integration. In this study, 31 young adults ($M = 25.8$ years) and 31 older adults ($M = 67.2$ years) were instructed to identify several basic emotions (happiness, sadness, anger, fear, disgust) and a neutral expression, which were displayed as visual (facial expressions), auditory (non-verbal affective vocalizations) or crossmodal (simultaneous, congruent facial and vocal affective expressions) stimuli. The results showed that older adults performed slower and worse than younger adults at recognizing negative emotions from isolated faces and voices. In the crossmodal condition, although slower, older adults were as accurate as younger except for anger. Importantly, additional analyses using the "race model" demonstrate that older adults benefited to the same extent as younger adults from the combination of facial and vocal emotional stimuli. These results help explain some conflicting results in the literature and may clarify emotional abilities related to daily life that are partially spared among older adults.

Keywords: aging, emotion, faces, voices, non-verbal vocalizations, multimodal integration, race model

Introduction

Emotion recognition is a fundamental component of social cognition. The ability to discriminate and interpret others' emotional states from emotional cues plays a crucial role in social functioning and behaviors (Carton et al., 1999; Adolphs, 2006; Corden et al., 2006; Frith and Frith, 2012). From early and throughout lifespan, emotion recognition is an essential mediator of successful social interactions and well-being (Izard, 2001; Engelberg and Sjöberg, 2004; Kryla-Lighthall and Mather, 2009; Suri and Gross, 2012). Hence, impaired recognition of others' emotional states may result in severe social dysfunctions, including inappropriate social behaviors, poor interpersonal communication and reduced quality of life (Feldman et al., 1991; Shimokawa et al., 2001; Blair, 2005). Such difficulties have been observed not only in disorders characterized

Abbreviations: PASA, Posterior–Anterior Shift in Aging; fMRI, functional magnetic resonance imaging; ERP, event-related potentials; RT, response times; STS, superior temporal sulcus, p-STC, posterior superior temporal cortex; BDI, Beck Depression Inventory; MMSE, Mini Mental State Examination; CDFs, cumulative distribution functions.

by prominent social-behavioral deficits (i.e., autism spectrum disorders, schizophrenia, neurodegenerative dementia; e.g., Chaby et al., 2012; and see for review Kennedy and Adolphs, 2012; Kumfor et al., 2014) but also in normal aging, which is frequently associated with social withdrawal and loneliness (e.g., Szanto et al., 2012; Steptoe et al., 2013).

Although older adults report high levels of satisfaction and better emotional stability with advancing age (Reed and Carstensen, 2012; Sims et al., 2015), they have difficulties processing some types of emotional information, which is often marked by a decline in emotion recognition (Ruffman et al., 2008; Isaacowitz and Blanchard-Fields, 2012). Most past studies have identified age-related difficulties in the visual channel, particularly when participants were asked to recognize emotion from posed facial expressions (see for review, Chaby and Narme, 2009; Isaacowitz and Stanley, 2011). These posed expressions were created to convey a single specific emotion, typically with exaggerated individual features, without any distracting or irrelevant features. However, emotions are not usually expressed solely by the face during daily social interactions; typically, voice (including non-verbal vocalizations) is also an important social signal, which needs to be processed quickly and accurately to allow successful interpersonal interactions. The rare studies that have explored how the ability to recognize vocal emotion changes with age have been conducted on speech prosody using words or sentences spoken with various emotional expressions. Theses studies concluded that advancing age is associated with increasing difficulties in recognizing emotion from prosodic cues (Kiss and Ennis, 2001; Paulmann et al., 2008; Mitchell et al., 2011; Lambrecht et al., 2012; Templier et al., 2015). However, vocal emotions could also be experienced via non-verbal affect bursts (e.g., *screams or laughter*; see Scherer, 1994) that typically accompany intense emotional feelings and that might be considered as the vocal counterpart of facial expressions. The processing of non-verbal vocal affects in aging individuals has rarely been studied (see Hunter et al., 2010; Lima et al., 2014); thus, this issue needs to be further investigated.

Altogether, the above studies showed evidence of age-related decline of some basic emotions via unimodal visual or auditory channels. These changes might start early, at approximately 40 years, for both facial (Williams et al., 2009) and prosodic emotions (Paulmann et al., 2008; Mill et al., 2009; Lima and Castro, 2011), and decline may occur linearly with advancing age (see Isaacowitz et al., 2007). In particular, compared to young adults, older adults could experience difficulties recognizing fear, anger and sadness from faces but experience no deficits recognizing happy or neutral faces (see for review, Isaacowitz et al., 2007; Ruffman et al., 2008). The recognition of disgust also seems highly preserved in older adults (e.g., Calder et al., 2003). Data from voices are less coherent, as difficulties have been found in older adults only for anger and sadness (Ruffman et al., 2008) or for almost all emotions (e.g., Paulmann et al., 2008).

Different mechanisms have been proposed to explain these age-related changes in emotion recognition. One preeminent explanation concerns structural and functional brain changes associated with age. Multiple interconnected brain regions are implicated in visual and auditory emotional processing. These regions include the frontal lobes, particularly the orbitofrontal

cortex (Hornak et al., 2003; Wildgruber et al., 2005; Tsuchida and Fellows, 2012) and the temporal lobes, particularly the superior temporal gyrus (Beer et al., 2006; Ethofer et al., 2006). The amygdala is also involved in this processing (Iidaka et al., 2002; Fecteau et al., 2007). Prefrontal cortex atrophy (in particular atrophy of the orbitofrontal region; Resnick et al., 2003, 2007; Lamar and Resnick, 2004) is a known marker of normal aging and could explain the difficulties identifying some facial emotions, in particular anger. Moreover, although the amygdala does not decline as rapidly as the frontal regions, some studies have reported a linear reduction of its volume with age (Mu et al., 1999; Allen et al., 2005). When comparing elderly people with young adults, neuroimaging studies observed a less significant activation of this structure among the elderly during the processing of emotional faces, especially negative ones (Mather et al., 2004). This was coupled with increased activity in the prefrontal cortex (Gunning-Dixon et al., 2003; Urry et al., 2006; Ebner et al., 2012). Conversely, other studies found a decrease in functional connectivity between the amygdala and posterior structures, which may reflect a decline in the perceptual process (Jacques et al., 2009). Overall, these patterns of brain activity observed in neuroimaging studies during a variety of emotional tasks (including recognition) are consistent with the Posterior–Anterior Shift in Aging (PASA; for review, see Dennis and Cabeza, 2008), which reflects the effect of aging on brain activity.

Another explanation for older adults' lower performance on negative emotion recognition emerges within the framework of the socio-emotional selectivity theory (Carstensen, 1992). With advancing age, adults appear to concentrate on a few emotionally rewarding relationships with their closest partners, report greater emotional control, and reduce their cognitive focus on negative information. Based on these observations, it was suggested that "paradoxically," the recognition of negative emotion declines (Carstensen et al., 2003; Charles and Carstensen, 2010; Mather, 2012; Huxhold et al., 2013).

Losses in cognitive and sensory functions are also possible explanations for age-related changes in emotion recognition. Increasing age is often associated with a decline in cognitive abilities (e.g., Verhaeghen and Salthouse, 1997; for review, see Salthouse, 2009), as well as with losses in visual and auditory acuity (Caban et al., 2005; Humes et al., 2009), which could hamper higher-level processes such as language and perception (Sullivan and Ruffman, 2004). However, these sensory attributes are shown to be poor predictors of the age-related decline in visual or auditory emotional recognition (e.g., Orbelo et al., 2005; Mitchell, 2007; Ryan et al., 2010; Lima et al., 2014).

Previous research on age-related differences in the recognition of basic emotions has focused predominantly on a single modality, and thus little is known about age-related differences in crossmodal emotion recognition. However, in daily life, people perceive emotions through multiple modalities, such as speech, voices, faces and postures (e.g., Young and Bruce, 2011; Belin et al., 2013). This indicates that our brain merges information from different senses to enhance perception and guide our behavior (Ernst and Bülthoff, 2004; Ethofer et al., 2013). Evidence supporting this idea includes studies of brain-damaged patients, such as traumatic or vascular brain injuries and brain tumors. These studies found

similar impairments in processing emotions from faces and voices in a single modality, but found that brain-damaged patients experienced greater performance using both facial and vocal stimuli (e.g., Hornak et al., 1996; Borod et al., 1998; Calder et al., 2001; Kucharska-Pietura et al., 2003; du Boullay et al., 2013; Luherne-du Boullay et al., 2014).

Some studies in young adults have demonstrated that congruent emotional information processed via multisensory channels optimizes behavioral responses, which results in enhanced accuracy and faster response times (RT; De Gelder and Vroomen, 2000; Kreifelts et al., 2007; Klasen et al., 2011). In older adults, audiovisual performances have been shown to be equivalent or even improved relative to younger adults (Laurienti et al., 2006; Peiffer et al., 2007; Diederich et al., 2008; Hugenschmidt et al., 2009; DeLoss et al., 2013), with more rare exceptions showing reduced multisensory integration in older adults (Walden et al., 1993; Sommers et al., 2005; Stephen et al., 2010). Some of these studies have explored the effects of age on crossmodal emotional processing and found evidence for preserved multisensory processing in older adults when congruent auditory and visual emotional information were presented simultaneously (Hunter et al., 2010; Lambrecht et al., 2012).

Multisensory integration refers to the process by which unisensory inputs are combined to form a new integrated product (Stein et al., 2010). This process has been studied in humans using neuroimaging techniques, which show that different regions of the human brain are implicated in the integration of multimodal cues, including "convergence" areas such as the superior temporal sulcus (STS; Laurienti et al., 2005; James and Stevenson, 2011; Watson et al., 2014; see for review, Stein et al., 2014). Neuroimaging techniques such as functional magnetic resonance imaging (fMRI) generally show greater activity in response to bimodal stimulation. More precisely, in a series of fMRI experiments conducted by Kreifelts and collaborators (e.g., Kreifelts et al., 2007; see for review, Brück et al., 2011) the posterior superior temporal cortex (p-STC) emerges as a crucial structure for the integration of facial and vocal cues. In event-related potential (ERP) studies (e.g., Giard and Peronnet, 1999; Foxe et al., 2000; Molholm et al., 2002), multisensory enhancement is measured by comparing the ERP from the multisensory condition to the sum of the ERPs from each unimodal condition. Multisensory enhancement is also commonly measured in behavioral studies by calculating a redundancy gain between the crossmodal stimulus and the more informational unimodal stimulus. Another interesting method performed in studies using RT, is to test whether the redundant target effect (shorter RT under the crossmodal condition) reflects an actual multisensory integrative process by comparing the observed RT distribution with the distribution predicted by the "race model" (proposed by Miller, 1982; see also Colonius and Diederich, 2006). The "race model" assumes that a crossmodal stimulus presentation produces parallel activation (i.e., in a separate way) of the unimodal stimuli. According to this model, the shortening of RT for crossmodal relative to unimodal stimuli derives from the fact that either unimodal stimulus can produce a response. Thus, any violation of the race model (i.e., if the observed RTs in crossmodal trials are shorter than those predicted by the race model) indicates that the stimuli are not

processed in separate channels, which suggests an underlying integrative mechanism (see Laurienti et al., 2006; Girard et al., 2013; Charbonneau et al., 2013).

To date, the processing mechanisms responsible for multisensory enhancement in older compared to young adults remains unclear, and crossmodal emotional integration in aging evaluated by the race model has not been investigated. To characterize the age-related effect on emotional processing, we used emotional human stimuli (i.e., happy, angry, fear, sad and disgust) and a neutral expression in the form of unimodal (facial or vocal) or crossmodal (simultaneous congruent facial and vocal expressions) cues. Isolated facial expression was studied using pictures of posed facial expressions, and isolated vocal expression was studied using non-verbal affect stimuli. Our primary focus concerned crossmodal emotional processing in aging, and we aimed to explore whether older adults benefit from congruent crossmodal integration and to better understand the nature of this benefit. According to recent studies of multisensory integration mechanisms during aging (e.g., Lambrecht et al., 2012; Freiherr et al., 2013; Mishra and Gazzaley, 2013), we hypothesized that older adults benefit from congruent crossmodal presentation when identifying emotions. To assess this hypothesis, we calculated redundancy gains for scores and used the race model for RTs to determine the nature of multisensory integration achieved by combining redundant visuo-auditory information.

Materials and Methods

Participants

The study participants consisted of 31 younger (20–35; $M = 25.8$, $SD = 6.4$; 16 females) and 31 older adults (60–76; $M = 67.2$, $SD = 5.8$; 17 females); see **Table 1**. The participants spoke French and reported having normal or corrected-to-normal vision and good hearing abilities at the time of testing. All participants were living independently in the community and were in good general physical health. None of the participants had any history of psychiatric or neurological disorders, which might compromise cognitive function. They also had a normal score on the Beck Depression Inventory (Beck et al., 1996; BDI II, 21 item version; a score of less than 17 was considered to be in the minimal range). All elderly adults completed the Mini Mental State Examination (Folstein et al., 1975; MMSE), on which they scored above the cut-off score (26/30) for risk of dementia. Grade level was calculated with the Mill Hill Vocabulary Scale (French adaptation: Deltour, 1993), and this did not differ between groups ($p = 0.55$).

TABLE 1 | Participant demographic characteristics.

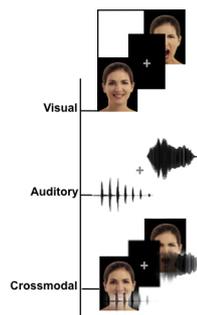|	Younger adults ($n = 31$)	Older adults ($n = 31$)
Age (years)	25.8 ± 6.4	67.2 ± 5.8
Education (years)	14.18 ± 1.6	13.55 ± 2.8
Mill-hill	36.87 ± 3.0	37.48 ± 4.8
Sex ratio (M/F)	16/15	17/14
BDI-II (/63)	5.65 ± 6.5	5.25 ± 3.8
MMSE (/30)	–	29.33 ± 0.6

FIGURE 1 | Schematic representation of the stimuli. Examples of the stimuli for the three different modalities, including visual (facial expressions), auditory (non-verbal affective vocalizations) and crossmodal stimuli (congruent facial and vocal emotions presented simultaneously).

The study was approved by the ethics committee of Paris Descartes University (Conseil d'Evaluation Ethique pour les Recherches en Santé, CERES, n IRB 2015100001072) and all participants gave informed consent.

Materials

Examples of stimuli and the task design for each condition are illustrated in **Figure 1**.

Visual stimuli. Visual stimuli consisted of pictures of human facial expressions obtained from the Karolinska Directed Emotional Faces database (Lundqvist et al., 1998). This database was chosen because it provided good examples of universal emotion categories with a high accuracy of labeling. The faces of 10 models (5 females, 5 males) expressing facial expressions of happiness, sadness, anger, fear, disgust or neutral constituted a set of 60 stimuli. All stimuli (presented on a black background) were 10 cm in height and subtended a vertical visual angle of 8° at a viewing distance of 70 cm.

Auditory stimuli. Auditory stimuli (**Figure 1**) consisted of non-verbal affective vocalizations (cry, laugh, etc.) obtained from The Montreal Affective Voices database (Belin et al., 2008). This database was chosen because it provided a standardized set of emotional vocalizations corresponding to the universal emotion categories without the potential confounds from linguistic content. The voices of 10 actors (5 females, 5 males) expressing happiness, sadness, anger, fear, disgust or neutral, vocalization constituted a set of 60 stimuli.

Crossmodal stimuli. Each emotional face was combined with an affective vocalization to construct 60 congruent expressions of faces and voices. The gender of the face and the voice were always congruent.

Procedure

Participants were tested individually in a single session that lasted approximately 45 min. The protocol was run using E-prime presentation software (Psychology Software Tools). Prior to the experiment, short facial-matching and vocal-matching tasks were administered to control for basic visual and auditory abilities in processing faces and voices. The subjects were asked to match the identity of non-emotional faces (i.e., six pairs of neutral faces obtained from the Karolinska Directed Emotional Faces Database) and non-emotional voices (i.e., six pairs of neutral voices obtained from the Montreal Affective Voices database). The stimuli were different from those used in the main task.

Then, after a short familiarization period, the experiment began. The experiment consisted of three blocks (visual, auditory, crossmodal) of 60 trials. Each trial started with the presentation of a fixation cross for 300 ms and was followed by the target stimulus, which was presented or repeated until the subject responded. Participants were asked to select (by clicking with the computer mouse) one label from a list of choices that best described the emotion presented. The six labels were displayed at the bottom of the computer screen and were visible throughout the test. There was an inter-trial interval of 700 ms. The order of the three blocks was counterbalanced across participants, and the order of trials was pseudo-randomized across each block. During the session, resting pauses were provided after every 10 trials, and the participants could take breaks if necessary between blocks. No feedback was given to the participants.

Statistical Analysis

Participants' accuracy (scores of correct responses) and corresponding RT (in milliseconds, ms) was computed for each condition. To control for outliers, trials with RT below 200 ms or greater than two standard deviations above the mean of each condition (0.90% of the trials in young adults; 1.25% of the trials in older adults) were excluded.

First, the data were entered into an overall analysis of variance (ANOVA), with age (young adults, older adults) as a between-subjects factor and with modality (visual, auditory, crossmodal) and emotion (neutral, happiness, fear, anger, sadness, disgust) as within-subjects factors. Effect sizes are reported as partial eta-squared (η_p^2). ANOVAs were adjusted with the Greenhouse-Geisser non-sphericity correction for effects with more than one degree of freedom. To provide clarity, uncorrected degrees of freedom, the Greenhouse-Geisser epsilon (ε) and adjusted p values are reported. Planned comparisons or *post hoc* Bonferroni tests were conducted to further explore the interactions between age, modality and emotion. The alpha level was set to 0.05 (p values were corrected for multiple comparisons).

Second, to examine whether both groups showed redundancy gains, as reflected by the difference in the scores when the visual and auditory stimuli were presented together (crossmodal condition) compared to each modality alone (unimodal condition), we calculated a "redundancy gain" for each participant separately by subtracting the higher of the scores under the unimodal conditions from the score under the crossmodal condition [(*crossmodal score—best modality score) × 100*] (see Calvert et al., 2004; Girard et al., 2013). The significance of the difference in redundancy gain (*in percent*) between younger and older participants was tested using an independent samples t - test.

Finally, to further test the advantage of crossmodal over unimodal processing, we investigated whether the RTs obtained under the crossmodal condition exceeded the statistical facilitation predicted by the race model (Miller, 1982). In multisensory research, the race model inequality has become a standard tool to identify crossmodal integration using RT data

(Townsend and Honey, 2007). To analyze the race model inequality, we used RMItest software (http://psy.otago.ac.nz/miller), which implements the algorithm described in Ulrich et al. (2007). The procedure requires four steps. First, participants' RTs in each condition (i.e., visual, auditory and crossmodal) are converted to cumulative distribution functions (CDFs). Second, the race model distribution is calculated by summing the CDFs of observed responses to the two unimodal conditions (visual and auditory) to create a "predicted" multisensory distribution. Third, percentile points (i.e., in the present study: 5th, 15th, 25th, 35th, 45th, 55th, 65th, 75th, 85th, and 95th) are determined for every distribution of RT. Finally, in each group, the mean RT for the crossmodal condition and the "predicted" condition are compared for each percentile using a *t*-test. If significant values are obtained in the crossmodal condition relative to the predicted condition, we conclude that the race model cannot account for the facilitation of the redundant signal conditions, supporting the existence of an integrative process.

Results

Age-related Difference in Emotion Recognition

Mean performance and RTs for all conditions are presented in **Table 2**[1]. For both younger and older groups, the mean performance accuracy was greater than 80% for the visual, auditory and crossmodal conditions. However, we found significant main effects of age, indicating that older adults performed less accurately and more slowly than younger adults ($85.23 \pm 1.24\%$ vs. $92.58 \pm 0.51\%$, $F(1,60) = 30.17$, $p < 0.001$, $\eta_p^2 = 0.33$ for scores; 3619 ± 145 ms vs. 1991 ± 68 ms, $F(1,60) = 103.4$, $p < 0.001$, $\eta_p^2 = 0.63$ for RTs). Importantly, we found a significant effect of modality on the scores [$F(2,120) = 137.54$, $p < 0.001$, $\varepsilon = 0.92$, $\eta_p^2 = 0.7$] and the RTs [$F(2,120) = 62.48$, $p < 0.001$, $\varepsilon = 0.88$, $\eta_p^2 = 0.51$], indicating that participants responded more effectively under the crossmodal condition than under either unimodal condition (all $p < 0.001$). There was a significant effect of emotion on the scores [$F(5,300) = 92.11$, $p < 0.001$, $\varepsilon = 0.62$, $\eta_p^2 = 0.60$] and the RTs [$F(5,300) = 36.91$, $p < 0.001$, $\eta_p^2 = 0.38$]. Furthermore, main effects were accompanied by several two-way interactions: between group and modality (see **Figure 2**) on the scores [$F(2,120) = 6.98$, $p = 0.002$, $\varepsilon = 0.92$, $\eta_p^2 = 0.10$] and the RTs [$F(2,120) = 7.66$, $p = 0.001$, $\varepsilon = 0.88$, $\eta_p^2 = 0.11$]; between group and emotion on the scores [$F(5,300) = 8.13$, $\varepsilon = 0.61$, $p < 0.001$, $\eta_p^2 = 0.12$] and the RTs [$F(5,300) = 8.62$, $p < 0.001$, $\eta_p^2 = 0.12$], and between modality and emotion on the scores [$F(10,600) = 27.01$, $p < 0.001$, $\varepsilon = 0.55$, $\eta_p^2 = 0.31$] and the RTs [$F(10,600) = 10.52$, $p < 0.001$, $\varepsilon = 0.56$, $\eta_p^2 = 0.15$]. Importantly, there was a significant effect of the three-way interaction between group, modality and emotion on the scores [$F(10,600) = 3.23$, $p = 0.005$, $\varepsilon = 0.55$, $\eta_p^2 = 0.05$] and the RTs [$F(10,600) = 2.7$, $p = 0.016$, $\eta_p^2 = 0.04$]. This reveals the following (see **Table 2**): (a) in the visual and auditory modality, older adults have lower scores

---

[1]To control for potential gender differences, this variable was initially entered as a between-subject factor in the analyses. However, gender failed to yield any significant main effects ($F < 1$) or interactions ($p > 0.1$) so we collapsed across gender in the reported analysis.
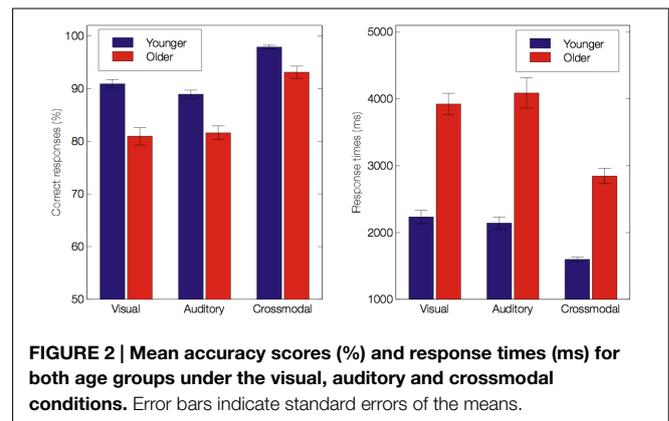

FIGURE 2 | Mean accuracy scores (%) and response times (ms) for both age groups under the visual, auditory and crossmodal conditions. Error bars indicate standard errors of the means.

than younger adults for the negative emotions only[2] (i.e., sadness, anger and disgust in the visual modality, $p < 0.01$; anger and fear in the auditory modality, $p < 0.01$), and (b) in the crossmodal condition, older adults perform more poorly than younger adults for anger only ($p < 0.001$). Concerning RTs, in the unimodal and crossmodal conditions, older adults were slower to identify all emotions (all $p < 0.001$) except for happiness ($p > 0.1$).

Integration of Crossmodal Emotional Information in Aging

To explore the ultimate crossmodal gain in the scores, we calculated a "redundancy gain" (i.e., the difference between the crossmodal condition and the unimodal condition with the higher score) for each participant in the two groups (see Materials and Methods section).

For the scores, our analysis indicated that the redundancy gain was greater for the older (8.82%) than for the younger adults (5.86%, $p = 0.007$). In the older group, all but two subjects showed a redundancy gain (29/31; one performed equally between the auditory modality and the crossmodal condition, and the other performed slightly better under the visual condition compared to the crossmodal condition). Moreover, there was a significant difference between the unimodal and crossmodal conditions for all emotions (all $p < 0.003$). In the younger group, all subjects except for one (30/31; who performed equally between the auditory condition and the crossmodal condition) showed a redundancy gain. Our analysis showed a significant difference between the unimodal and crossmodal conditions for negative emotions only (fear, sadness, anger and disgust) (all $p < 0.007$); for the neutral emotion and for happiness, performance ceilings may explain the lack of significant effects (all $p > 0.1$).
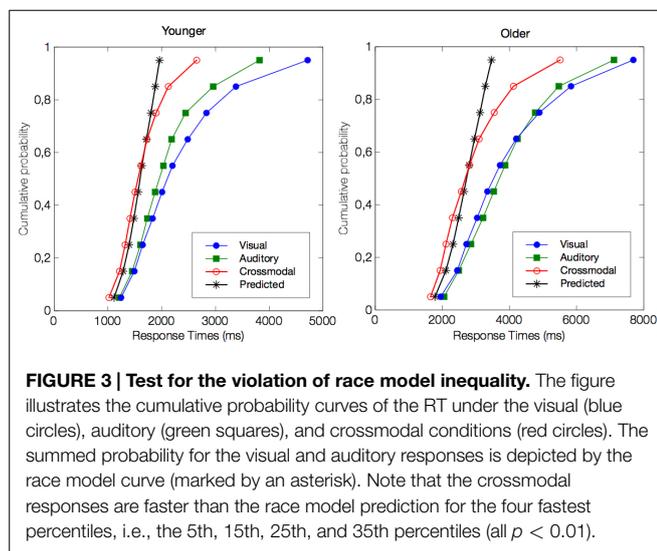
For RTs, we used the race model to explore crossmodal integration and to determine whether the observed crossmodal behavioral enhancement (i.e., shorter RTs) was beyond that predicted by statistical summation of the unimodal visual and auditory conditions (**Figure 3**). In the younger group, we observed a violation of the race model prediction for the 5th, 15th, 25th, and 35th percentiles of the RT distribution (all $p < 0.01$, but not for the slowest percentiles (all $p > 0.1$). These results support

---

[2]Note that a ceiling effect was observed for happiness in both groups and for neutral in the younger group.

TABLE 2 | Mean accuracy scores (%) and response times (ms) by age group and emotion. Standard errors of the means are shown in parentheses.

		Mean accuracy (%)					
		Neutral	Happy	Fear	Sadness	Anger	Disgust
Older	Visual	90.9 (2.7)	99.3 (0.4)	83.2 (2.3)	73.5 (3.4)	66.1 (4.7)	72.5 (2.5)
	Auditory	92.6 (2.7)	95.2 (1.9)	73.2 (3.9)	94.5 (1.3)	47.7 (3.4)	86.4 (2.3)
	Crossmodal	98.7 (0.6)	100 (0.0)	89.6 (2.1)	95.8 (1.8)	76.8 (4.2)	97.7 (0.8)
Younger	Visual	97.1 (1.1)	99.7 (0.6)	87.7 (2.2)	90.9 (2.6)	84.5 (2.4)	85.4 (2.5)
	Auditory	96.7 (0.9)	98.0 (0.9)	86.7 (2.3)	94.2 (1.3)	62.9 (3.0)	94.8 (1.2)
	Crossmodal	99.6 (0.3)	99.0 (0.7)	97.4 (0.9)	97.4 (0.8)	94.5 (1.5)	99.3 (0.4)

		Response times (ms)					
		Neutral	Happiness	Fear	Sadness	Anger	Disgust
Older	Visual	3447 (227)	2502 (136)	4366 (225)	4805 (325)	4243 (280)	4173 (224)
	Auditory	4355 (342)	3195 (213)	4370 (265)	3824 (238)	4920 (354)	3871 (250)
	Crossmodal	2796 (158)	2188 (105)	3135 (138)	3115 (146)	3273 (190)	2567 (132)
Younger	Visual	1871 (107)	1555 (50)	2408 (162)	2464 (157)	2569 (177)	2538 (163)
	Auditory	2148 (157)	1967 (117)	2345 (153)	2082 (74)	2281 (124)	2026 (80)
	Crossmodal	1613 (49)	1370 (35)	1689 (54)	1722 (50)	1571 (45)	1624 (35)

FIGURE 3 | Test for the violation of race model inequality. The figure illustrates the cumulative probability curves of the RT under the visual (blue circles), auditory (green squares), and crossmodal conditions (red circles). The summed probability for the visual and auditory responses is depicted by the race model curve (marked by an asterisk). Note that the crossmodal responses are faster than the race model prediction for the four fastest percentiles, i.e., the 5th, 15th, 25th, and 35th percentiles (all $p < 0.01$).

the existence of a crossmodal integrative process. The temporal window in which this benefit was significant was from 1019 to 1410 ms. Similar to the responses by the younger group, the older group responses were shorter than those predicted by the race model for the 5th, 15th, 25th, and 35th percentiles of the RT distribution (all $p < 0.01$). The temporal window in which this benefit was significant was from 1647 to 2300 ms. Although the maximal enhancement occurred at different absolute RTs between the two populations, this peak enhancement occurred at the exact same percentile of the cumulative distribution curve.

Discussion

While a large body of evidence shows that older adults are less accurate than younger adults in recognizing specific emotions

from emotional faces, fewer studies have examined vocal emotion recognition, and hardly any studies have investigated the recognition of emotion from emotional faces and voices presented simultaneously (Hunter et al., 2010; Lambrecht et al., 2012). The purpose of this study was to compare unimodal facial and vocal emotion processing in older and younger adults and, in addition, to test whether older adults benefit from the combination of congruent emotional information from different channels, which reveals crossmodal integration. Our results first confirm that older adults experience difficulties in emotion recognition. They were less accurate and slower overall than younger adults in processing emotion from facial or non-verbal vocal expressions presented alone. Second, the participants similarly recognized facial and vocal cues, and both groups benefitted from the crossmodal condition. Third, age-related differences were modulated by emotion, as older adults were particularly affected in term of accuracy with regards to processing negative emotions under both the facial and vocal conditions. Finally, our results provide compelling evidence for the multisensory nature of emotional processing in aging. The important finding of this study was that older adults benefit to the same extent as younger adults from the combination of information presented in the visual and auditory modalities. This suggests that crossmodal processing represents a mechanism compensating for deficits in the visual or auditory channels that often affect older adults.

Effects of Age on Emotion Recognition Based on Unimodal Stimuli

Our findings indicated that emotion recognition based on unimodal stimuli changes with age. In the visual modality, our results support previous findings showing age-related difficulties in the ability to recognize emotion from facial cues (see for a meta-analysis, Ruffman et al., 2008). However, most of these studies

used the collection of posed black-and-white photographs of human faces from the 1970s Ekman dataset (e.g., Orgeta and Phillips, 2008; Hunter et al., 2010; Slessor et al., 2010) that has been criticized for its lack of ecological validity, which leads to questions about the generalizability of the results (Murphy and Isaacowitz, 2010). The present study used emotional expressions consisting of static color photographs of faces (see also, Ebner et al., 2010; Eisenbarth and Alpers, 2011), and this study confirmed the robustness of age-related difficulties. The fact that the same results were found using dynamic facial expressions (Lambrecht et al., 2012) confirms that widespread difficulties in recognizing emotion from facial cues are encountered by older adults. In the auditory modality, the ability to recognize emotion from non-verbal vocal cues also becomes less efficient with age. This result is in accordance with that of Hunter et al. (2010), who used non-verbal affective vocalizations. It is also in line with some recent studies using spoken words in a neutral context that showed impairments in decoding emotional speech with advancing age (Paulmann et al., 2008; Mitchell et al., 2011; Lambrecht et al., 2012). As normal variations of prosodic emotion ability could be associated with depression, relationship satisfaction or well-being in younger populations (Noller and Feeney, 1994; Emerson et al., 1999; Carton et al., 1999), the question remains whether and how the age-related decline in emotional vocal processing influences social interactions. However, it is important to note that in our study, the performance of the older group reached 80%, suggesting a relatively mild deficit. This suggest that non-verbal vocalizations, that are devoid of linguistic information, are however, effective at communicating diverse emotions in aging.

Age-related Difference in the Responses to Different Sensory Modalities and Specific Emotions

However, the main effect of age was tempered by a set of interactions, suggesting that age-related differences varied across modalities and across specific emotions. Specifically, in response to the visual and auditory stimuli, we found an age-related reduction in accuracy for negative expressions (i.e., fear, sadness, anger and disgust) and comparable performance for neutral and happy expressions. For the visual modality, this result is in accordance with individual studies using images of static faces showing different emotional expressions, which showed that certain discrete emotions, notably negative ones, are more sensitive to age-related variation (see for review, Ruffman et al., 2008). Studies regarding the auditory channel are more inconsistent because they are based on diverse paradigms. The results differ inasmuch as the studies did not isolate specific emotions (e.g., Orbelo et al., 2005; Mitchell, 2007; Mitchell et al., 2011), they investigated negative emotions only (e.g., Hunter et al., 2010), they explored a few contrasting emotions (e.g., Lambrecht et al., 2012) or they included several positive and negative emotions (e.g., Wong et al., 2005; Lima et al., 2014). Wong et al. (2005) found that older adults poorly recognized only sadness and happiness in speech; in contrast, using non-verbal vocalizations, Hunter et al. (2010) found that older adults poorly identified negative emotions (fear, anger,

sadness, disgust), whereas Lima et al. (2014) found that older adults performed poorly for all emotions (positive and negative ones). Note however, that for scores, interpretations about age-related difference in responses to specific emotions are limited because of the presence of ceiling effects for happy and neutral expressions. Interestingly, for RTs the effects seem to be more general since older adults were especially slow to respond to all emotions.

These divergent results across aging studies may be due to the individual variability of the samples and the use of different types of emotional stimuli with varying presentation times, which might influence the identification of the given emotion. In the present study, the stimuli were presented or repeated until the subject provided a response. The observed slower RT for all negative emotions contrasts with the findings of recent studies (Pell and Kotz, 2011; Rigoulot et al., 2013) using verbal emotional stimuli, which showed that listeners are generally faster at identifying fear, anger, and sadness and slower at identifying happiness and disgust. This suggests that non-verbal affective vocalizations are processed at different rates. Interestingly, this time window is consistent with a work by Pell (2005); when happy, sad, or neutral pseudo-utterances spoken in English were cut from the onset of the sentence to last 300, 600, or 1000 ms in duration, emotional priming of a congruent static face was only observed when vocal cues were presented for 600 or 1000 ms, but not for only 300 ms. Hence, vocal information enduring at least 600 ms maybe necessary to presumably activate shared emotion knowledge responsible for multimodal integration. More importantly, our data show that the participants did not find it easier to identify emotion from isolated facial or non-verbal vocal cues. By contrast, Hunter et al. (2010), who used facial and vocal non-verbal emotions, found that emotion recognition was easier in response to facial cues than vocal cues. However, our experiment used not only negative but also happy and neutral expressions, which potentially improved the performance of older adults in both the visual and auditory modalities.

Overall, these results are consistent with the fact that age-related emotional difficulties do not reflect general cognitive aging (Orbelo et al., 2005) but rather a complex change affecting discrete emotions; notably, the same authors also suggest that the age-related decline in emotional processing is not explained by sex effects or age-related visual or hearing loss. Nevertheless, assessing hearing and seeing abilities objectively could have informed the pattern of our findings and we can consider the lack of measuring these covariates as a limitation of the study. For example, recent findings (Ruggles et al., 2011; Bharadwaj et al., 2015) suggest that despite normal or near-normal hearing thresholds, a significant portion of listeners exhibit deficits in everyday communication (i.e., in complex environments such as noisy restaurants or busy streets).

These results could also be interpreted in terms of the socio-emotional selectivity theory, which states that aging increases emotional control, diminishes the impact of negative emotions and facilitates concentration on more positive social interactions (e.g., Charles and Carstensen, 2010; Huxhold et al., 2013). However, Frank and Stennett (2001) have noted that using only a few basic emotion categories allows participants to choose

their response based on discrimination and exclusion rules, which is less likely to be the case in a real-life setting. In particular, if happiness is the only positive emotion, participants can make the correct choice as soon as they recognize a smile. Therefore, a ceiling effect can be an alternative explanation to the socio-emotional selectivity theory. An alternative to examine possible valence-specific effects is the use of a similar number of positive and negative emotions (see Lima et al., 2014).

Integration of Crossmodal Emotional Information in Aging Individuals

The principal goal of the current study was to explore whether older adults benefit from congruent crossmodal integration and to better understand the nature of this benefit. In daily life, the combination of information from facial and vocal expressions usually results in a more robust representation of the expressed emotion (e.g., De Gelder and Vroomen, 2000; Dolan et al., 2001; De Gelder and Bertelson, 2003), which thus results in a more unified perception of the person (Young and Bruce, 2011).

In our study, emotional faces and voices come from different sensory modalities to build a unified and coherent representation of the same percept (i.e., an emotion) as defined by crossmodal integration mechanisms (Driver and Spence, 2000). We showed that whereas older adults exhibited slower RTs under the crossmodal condition, resulting in a different temporal window of multisensory enhancement, a multisensory benefit occurred to the same extent in the two groups. However, early studies of multisensory integration in aging individuals showed that compared to younger adults, older adults did not benefit from multisensory cues (Stine et al., 1990; Walden et al., 1993; Sommers et al., 2005) and experienced a suppressed cortical multisensory integration response that was associated with poor cortical integration (Stephen et al., 2010). By contrast, more recent studies point toward an enhancement of multisensory integration effects in older adults, notably reporting shorter RT in response to multisensory events (e.g., Mahoney et al., 2011, 2012; DeLoss et al., 2013).

Consistent with the latter works, the present study indicates that in younger and older adults, emotional information derived from facial and vocal cues is not reducible to the simple sum of the unimodal inputs and suggests that multisensory integration is maintained with increasing age and could play a compensatory role in normal aging. This is in accordance with a magneto-encephalography study (Diaconescu et al., 2013), which indicated that sensory-specific regions showed increased activity after visual-auditory stimulation in young and old participants but that inferior parietal and medial prefrontal areas were preferentially activated in older subjects. Activation of the latter areas was related to faster detection of multisensory stimuli. The authors proposed that the posterior parietal and medial prefrontal activity sustains the integrated response in older adults. This hypothesis is supported by the theory of PASA and that of cortical dedifferentiation, stating that healthy aging is accompanied by decreased specificity of neurons in the prefrontal cortex (Park and Reuter-Lorenz, 2009; Freiherr et al., 2013). This could explain why the

crossmodal RTs of older adults was longer than that of younger adults for each emotion.

Furthermore, a recent study using ERPs by Mishra and Gazzaley (2013) among healthy older adults (60–90 years old) suggested the existence of compensatory mechanisms susceptible to sustaining efficient crossmodal processing. The authors showed evidence that distributed audio-visual attention results in improved discrimination performance (faster RTs without any differences in accuracy in congruent stimuli settings) compared to focused visual attention. They noted that the benefits of distributed audio-visual attention in older adults matched those of younger adults. Interestingly, ERPs recoding during the task further revealed intact crossmodal integration in higher performing older adults, who had results similar to those of younger adults. As suggested by Barulli et al. (2013), attention, executive function and verbal IQ may play a role in the generation of a "cognitive reserve" that reduces the deleterious effects of aging and, thus, buffers against a diminished adaptive strategy (Hodzik and Lemaire, 2011). These results show the necessity of taking into account individual cognitive differences in aging. It is clear that significant cognitive decline is not an inevitable consequence of advancing age and that each cognitive domain is differentially affected. As aging can have diverse effects on cognitive functions, it is therefore important to emphasize the maintained functions rather than taking a customary approach that only underlines the loss of capacities among the elderly.

It should be noted however, as a possible limitation of the current study, that our stimuli are quite unnaturalistic since they combine non-dynamic (photographs) and dynamic (sound) stimuli. Although our participants did not report any incongruent perception of crossmodal stimuli, the use of emotional expressions that contain truly multimodal expressions (video and audio obtained from the same person), which are not posed, but enacted using the Stanislawski technique (see the Geneva Multimodal Emotion Portrayals, GEMEP; Bänziger et al., 2012) could be relevant.

Conclusion

In conclusion, our results suggest that despite a decline in facial and vocal emotional processing with advancing age, older adults integrate facial and vocal cues to yield a unified perception of the person. Given the changes in facial and vocal modality exhibited by older adults, it may be helpful for family members and caregivers to use multiple sensory modalities to communicate important affective information. Thus, supplementing facial cues with vocal information may facilitate communication, preventing older individuals from withdrawing from the community and reducing the development of affective disturbances such as depression. Future research is required to further examine whether crossmodal integration can benefit older adults who exhibit cognitive impairments (e.g., Mild Cognitive Impairments, Alzheimer's Disease). Such studies would be of particular interest in the context of recently developed assistive robotics platforms that prolong the ability of persons who have lost their autonomy to remain at home. For instance, serious games and socially aware

assistive robots have actually been designed without considering the age-specific effects on social signal recognition. Therefore, improving the efficiency and suitability of these interactive systems clearly requires a better understanding of crossmodal integration.

Author Contributions

Study concept and design was performed by LC, VL, MC, and MP. Data acquisition was conducted by VL. Data analysis was performed by LC, VL, and MC. All authors contributed to data interpretation and the final version of the manuscript, which all approved.

References

Adolphs, R. (2006). Perception and emotion how we recognize facial expressions. *Cur. Dir. Psychol. Sci.* 15, 222–226. doi: 10.1111/j.1467-8721.2006.00440.x

Allen, J. S., Bruss, J., Brown, C. K., and Damasio, H. (2005). Normal neuroanatomical variation due to age: the major lobes and a parcellation of the temporal region. *Neurobiol. Aging* 26, 1245–1260. doi: 10.1016/j.neurobiolaging.2005.05.023

Bänziger, T., Mortillaro, M., and Scherer, K. R. (2012). Introducing the Geneva multimodal expression corpus for experimental research on emotion perception. *Emotion* 12, 1161–1179. doi: 10.1037/a0025827

Barulli, D. J., Rakitin, B. C., Lemaire, P., and Stern, Y. (2013). The influence of cognitive reserve on strategy selection in normal aging. *J. Int. Neuropsychol. Soc.* 19, 841–844. doi: 10.1017/S1355617713000593

Beck, A. T., Steer, R. A., and Brown, G. (1996). *BDI-II, Beck Depression Inventory-II.* San Antonio, TX: Psychological Corporation.

Beer, J. S., John, O. P., Scabini, D., and Knight, R. T. (2006). Orbitofrontal cortex and social behavior: integrating self-monitoring and emotion-cognition interactions. *J. Cogn. Neurosci.* 18, 871–879. doi: 10.1162/jocn.2006.18.6.871

Belin, P., Campanella, S., and Ethofer, T. (2013). *Integrating Face and Voice in Person Perception.* New York, NY: Springer.

Belin, P., Fillion-Bilodeau, S., and Gosselin, F. (2008). The Montreal affective voices: a validated set of nonverbal affect bursts for research on auditory affective processing. *Behav. Res. Methods* 40, 531–539. doi: 10.3758/BRM.40.2.531

Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., and Shinn-Cunningham, B. G. (2015). Individual differences reveal correlates of hidden hearing deficits. *J. Neurosci.* 35, 2161–2172. doi: 10.1523/JNEUROSCI.3915-14.2015

Blair, R. J. R. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Conscious. Cogn.* 14, 698–718. doi: 10.1016/j.concog.2005.06.004

Borod, J. C., Cicero, B. A., Obler, L. K., Welkowitz, J., Erhan, H. M., Santschi, C., et al. (1998). Right hemisphere emotional perception: evidence across multiple channels. *Neuropsychology* 12, 446–458. doi: 10.1037/0894-4105.12.3.446

Brück, C., Kreifelts, B., and Wildgruber, D. (2011). Emotional voices in context: a neurobiological model of multimodal affective information processing. *Phys. Life Rev.* 8, 383–403. doi: 10.1016/j.plrev.2011.10.002

Caban, A. J., Lee, D. J., Gómez-Marín, O., Lam, B. L., and Zheng, D. D. (2005). Prevalence of concurrent hearing and visual impairment in US adults: the national health interview survey, 1997–2002. *Am. J. Public. Health* 95, 1940–1942. doi: 10.2105/AJPH.2004.056671

Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., et al. (2003). Facial expression recognition across the adult life span. *Neuropsychologia* 41, 195–202. doi: 10.1016/S0028-3932(02)00149-5

Calder, A. J., Lawrence, A. D., and Young, A. W. (2001). Neuropsychology of fear and loathing. *Nat. Rev. Neurosci.* 2, 352–363. doi: 10.1038/35072584

Calvert, G., Spence, C., and Stein, B. E. (2004). The Handbook of Multisensory Processes. Cambridge, MA: MIT Press.

Carstensen, L. L. (1992). Social and emotional patterns in adulthood: support for socioemotional selectivity theory. *Psychol. Aging* 7, 331–338. doi: 10.1037/0882-7974.7.3.331

Carstensen, L. L., Fung, H. H., and Charles, S. T. (2003). Socioemotional selectivity theory and the regulation of emotion in the second half of life. *Motiv. Emot.* 27, 103–123. doi: 10.1023/A:1024569803230

Carton, J. S., Kessler, E. A., and Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *J. Nonverbal. Behav.* 23, 91–100. doi: 10.1023/A:1021339410262

Chaby, L., Chetouani, M., Plaza, M., and Cohen, D. (2012). "Exploring multimodal social-emotional behaviors in autism spectrum disorders," in *Workshop on Wide Spectrum Social Signal Processing, 2012 ASE/IEEE International Conference on Social Computing*, Amsterdam. 950–954. doi : 10.1109/SocialCom-PASSAT.2012.111

Chaby, L., and Narme, P. (2009). Processing facial identity and emotional expression in normal aging and neurodegenerative diseases. *Psychol. Neuropsychiatr. Vieil.* 7, 31–42. doi: 10.1016/S0987-7053(99)90054-0

Charbonneau, G., Bertone, A., Lepore, F., Nassim, M., Lassonde, M., Mottron, L., et al. (2013). Multilevel alterations in the processing of audio–visual emotion expressions in autism spectrum disorders. *Neuropsychologia* 51, 1002–1010. doi: 10.1016/j.neuropsychologia.2013.02.009

Charles, S., and Carstensen, L. L. (2010). Social and emotional aging. *Annu. Rev. Psychol.* 61, 383–409. doi: 10.1146/annurev.psych.093008.100448

Colonius, H., and Diederich, A. (2006). The race model inequality: interpreting a geometric measure of the amount of violation. *Psychol. Rev.* 113, 148–154. doi: 10.1037/0033-295X.113.1.148

Corden, B., Critchley, H. D., Skuse, D., and Dolan, R. J. (2006). Fear recognition ability predicts differences in social cognitive and neural functioning in men. *J. Cogn. Neurosci.* 18, 889–897. doi: 10.1162/jocn.2006.18.6.889

De Gelder, B., and Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn. Sci.* 7, 460–467. doi: 10.1016/j.tics.2003.08.014

De Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cogn. Emot.* 14, 289–311. doi: 10.1080/026999300378824

DeLoss, D. J., Pierce, R. S., and Andersen, G. J. (2013). Multisensory integration, aging, and the sound-induced flash illusion. *Psychol. Aging* 28, 802–812. doi: 10.1037/a0033289

Deltour, J. J. (1993). *Echelle de Vocabulaire de Mill Hill de JC Raven. Adaptation Française et Normes Européennes du Mill Hill et du Standard Progressive Matrices de Raven (PM38).* Braine-le-Château: Editions l'application des techniques modernes.

Dennis, N. A., and Cabeza, R. (2008). "Neuroimaging of healthy cognitive aging", in *Handbook of Aging and Cognition*, 3rd Edn, eds T. A. Salthouse and F. E. M. Craik (Mahwah, NJ: Lawrence Erlbaum Associates), 1–54.

Diaconescu, A. O., Hasher, L., and McIntosh, A. R. (2013). Visual dominance and multisensory integration changes with age. *Neuroimage* 65, 152–166. doi: 10.1016/j.neuroimage.2012.09.057

Diederich, A., Colonius, H., and Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia* 46, 2556–2562. doi: 10.1016/j.neuropsychologia.2008.03.026

Dolan, R. J., Morris, J. S., and de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *P. Natl. A. Sci.* 98, 10006–10010. doi: 10.1073/pnas.171288598

Driver, J., and Spence, C. (2000). Multisensory perception: beyond modularity and convergence. *Curr. Biol.* 10, R731–R735. doi: 10.1016/S0960-9822(00)00740-5

du Boullay, V., Plaza, M., Capelle, L., and Chaby, L. (2013). Identification of emotions in patients with low-grade gliomas versus cerebrovascular accidents. *Rev. Neurol.* 169, 249–257. doi: 10.1016/j.neurol.2012.06.017

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

Ebner, N. C., Riediger, M., and Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men. Development and validation. *Behav. Res. Methods* 42, 351–362. doi: 10.3758/BRM.42.1.351

Eisenbarth, H., and Alpers, G. W. (2011). Happy mouth and sad eyes: scanning emotional facial expressions. *Emotion* 11, 860–865. doi: 10.1037/a0022758

Emerson, C. S., Harrison, D. W., and Everhart, D. E. (1999). Investigation of receptive affective prosodic ability in school-aged boys with and without depression. *Cogn. Behav. Neurol.* 12, 102–109.

Engelberg, E., and Sjöberg, L. (2004). Emotional intelligence, affect intensity, and social adjustment. *Pers. Indiv. Differ.* 37, 533–542. doi: 10.1016/j.paid.2003.09.024

Ernst, M. O., and Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends. Cogn. Sci.* 8, 162–169. doi: 10.1016/j.tics.2004.02.002

Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., et al. (2006). Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. *Neuroimage* 30, 580–587. doi: 10.1016/j.neuroimage.2005.09.059

Ethofer, T., Bretscher, J., Wiethoff, S., Bisch, J., Schlipf, S., Wildgruber, D., et al. (2013). Functional responses and structural connections of cortical areas for processing faces and voices in the superior temporal sulcus. *Neuroimage* 76, 45–56. doi: 10.1016/j.neuroimage.2013.02.064

Fecteau, S., Belin, P., Joanette, Y., and Armony, J. L. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *Neuroimage* 36, 480–487. doi: 10.1016/j.neuroimage.2007.02.043

Feldman, R. S., Philippot, P., and Custrini, R. J. (1991). "Social competence and nonverbal behavior," in *Fundamentals of Nonverbal Behavior*, eds R. S. Feldman and B. Rime (New York, NY: Cambridge University Press), 329–350.

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiat. Res.* 12, 189–198.

Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., and Schroeder, C. E. (2000). Multisensory auditory–somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cogn. Brain Res.* 10, 77–83. doi: 10.1016/S0926-6410(00)00024-0

Frank, M. G., and Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *J. Pers. Soc. Psychol.* 80, 75. doi: 10.1037/0022-3514.80.1.75

Freiherr, J., Lundström, J. N., Habel, U., and Reetz, K. (2013). Multisensory integration mechanisms during aging. *Front. Hum. Neurosc.* 7:863. doi: 10.3389/fnhum.2013.00863

Frith, C. D., and Frith, U. (2012). Mechanisms of social cognition. *Annu. Rev. Psychol.* 63, 287–313. doi: 10.1146/annurev-psych-120710-100449

Giard, M. H., and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J. Cogn. Neurosci.* 11, 473–490. doi: 10.1162/089892999563544

Girard, S., Pelland, M., Lepore, F., and Collignon, O. (2013). Impact of the spatial congruence of redundant targets on within-modal and cross-modal integration. *Exp. Brain Res.* 224, 275–285. doi: 10.1007/s00221-012-3308-0

Gunning-Dixon, F. M., Gur, R. C., Perkins, A. C., Schroeder, L., Turner, T., Turetsky, B. I., et al. (2003). Age-related differences in brain activation during emotional face processing. *Neurobiol. Aging* 24, 285–295. doi: 10.1016/S0197-4580(02)00099-4

Hodzik, S., and Lemaire, P. (2011). Inhibition and shifting capacities mediate adults' age-related differences in strategy selection and repertoire. *Acta. Psychol.* 137, 335–344. doi: 10.1016/j.actpsy.2011.04.002

Hornak, J., Bramham, J., Rolls, E. T., Morris, R. G., O'Doherty, J., Bullock, P. R., et al. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain* 126, 1691–1712. doi: 10.1093/brain/awg168

Hornak, J., Rolls, E. T., and Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia* 34, 247–261. doi: 10.1016/0028-3932(95)00106-9

Hugenschmidt, C. E., Peiffer, A. M., McCoy, T. P., Hayasaka, S., and Laurienti, P. J. (2009). Preservation of crossmodal selective attention in healthy aging. *Exp. Brain Res.* 198, 273–285. doi: 10.1007/s00221-009-1816-3

Hunter, E. M., Phillips, L. H., and MacPherson, S. E. (2010). Effects of age on cross-modal emotion perception. *Psychol. Aging* 25, 779–787. doi: 10.1037/a0020528

Humes, L. E., Busey, T. A., Craig, J. C., and Kewley-Port, D. (2009). The effects of age on sensory thresholds and temporal gap detection in hearing, vision, and touch. *Atten. Percept. Psychophys.* 71, 860–871. doi: 10.3758/APP.71.4.860

Huxhold, O., Fiori, K. L., and Windsor, T. D. (2013). The dynamic interplay of social network characteristics, subjective well-being, and health: the costs and benefits of socio-emotional selectivity. *Psychol. Aging* 28, 3–16. doi: 10.1037/a0030170

Iidaka, T., Okada, T., Murata, T., Omori, M., Kosaka, H., Sadato, N., et al. (2002). Age-related differences in the medial temporal lobe responses to emotional faces as revealed by fMRI. *Hippocampus* 12, 352–362. doi: 10.1002/hipo.1113

Isaacowitz, D. M., and Blanchard-Fields, F. (2012). Linking process and outcome in the study of emotion and aging. *Perspect. Psychol. Sci.* 7, 3–17. doi: 10.1177/1745691611424750

Isaacowitz, D. M., Löckenhoff, C. E., Lane, R. D., Wright, R., Sechrest, L., Riedel, R., et al. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychol. Aging* 22, 147–159. doi: 10.1037/0882-7974.22.1.147

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

Izard, C. (2001). Emotional intelligence or adaptive emotions? *Emotion* 1, 249–257. doi: 10.1037/1528-3542.1.3.249

Jacques, P. L. S., Dolcos, F., and Cabeza, R. (2009). Effects of aging on functional connectivity of the amygdala for subsequent memory of negative pictures a network analysis of functional magnetic resonance imaging data. *Psychol. Sci.* 20, 74–84. doi: 10.1111/j.1467-9280.2008.02258.x

James, T. W., and Stevenson, R. A. (2011). "The use of fMRI to assess multisensory integration," in *The Neural Basis of Multisensory Processes*, Chap. 8, eds M. M. Murray and M. H. Wallace (Boca Raton, FL: CRC Press), 1–20.

Kennedy, D. P., and Adolphs, R. (2012). The social brain in psychiatric and neurological disorders. *Trends Cogn. Sci.* 16, 559–572. doi: 10.1016/j.tics.2012.09.006

Kiss, I., and Ennis, T. (2001). Age-related decline in perception of prosodic affect. *Appl. Neuropsychol.* 8, 251–254. doi: 10.1207/S15324826AN0804_9

Klasen, M., Kenworthy, C. A., Mathiak, K. A., Kircher, T. T., and Mathiak, K. (2011). Supramodal representation of emotions. *J. Neurosci.* 31, 13635–13643. doi: 10.1523/JNEUROSCI.2833-11.2011

Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., and Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage* 37, 1445–1456. doi: 10.1016/j.neuroimage.2007.06.020

Kryla-Lighthall, N., and Mather, M. (2009). "The role of cognitive control in older adults' emotional well-being," in *Handbook of Theories of Aging*, 2nd Edn, eds V. Berngtson, D. Gans, N. Putney, and M. Silverstein (New York, NY: Springer Publishing), 323–344.

Kucharska-Pietura, K., Phillips, M. L., Gernand, W., and David, A. S. (2003). Perception of emotions from faces and voices following unilateral brain damage. *Neuropsychologia* 41, 1082–1090. doi: 10.1016/S0028-3932(02)00294-4

Kumfor, F., Irish, M., Leyton, C., Miller, L., Lah, S., Devenney, E., et al. (2014). Tracking the progression of social cognition in neurodegenerative disorders. *J. Neurol. Neurosurg. Psychiatry* 85, 1076–1083. doi: 10.1136/jnnp-2013-307098

Lamar, M., and Resnick, S. M. (2004). Aging and prefrontal functions: dissociating orbitofrontal and dorsolateral abilities. *Neurobiol. Aging* 25, 553–558. doi: 10.1016/j.neurobiolaging.2003.06.005

Lambrecht, L., Kreifelts, B., and Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion* 12, 529–539. doi: 10.1037/a0026827

Laurienti, P. J., Burdette, J. H., Maldjian, J. A., and Wallace, M. T. (2006). Enhanced multisensory integration in older adults. *Neurobiol. Aging* 27, 1155–1163. doi: 10.1016/j.neurobiolaging.2005.05.024

Laurienti, P. J., Perrault, T. J., Stanford, T. R., Wallace, M. T., and Stein, B. E. (2005). On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. *Exp. Brain Res.* 166, 289–297. doi: 10.1007/s00221-005-2370-2

Lima, C. F., Alves, T., Scott, S. K., and Castro, S. L. (2014). In the ear of the beholder: how age shapes emotion processing in nonverbal vocalizations. *Emotion* 14, 145–160. doi: 10.1037/a0034287

Lima, C. F., and Castro, S. L. (2011). Speaking to the trained ear: musical expertise enhances the recognition of emotions in speech prosody. *Emotion* 11, 1021–1031. doi: 10.1037/a0024521

Luherne-du Boullay, V., Plaza, M., Perrault, A., Capelle, L., and Chaby, L. (2014). Atypical crossmodal emotional integration in patients with gliomas. *Brain Cogn.* 92, 92–100. doi: 10.1016/j.bandc.2014.10.003

Lundqvist, D., Flykt, A., and Ohman, A. (1998). *The Averaged Karolinska Directed Emotional Faces.* Stockholm: Karolinska Institute, Department of Clinical Neuroscience, Section Psychology.

Mahoney, J. R., Li, P. C. C., Oh-Park, M., Verghese, J., and Holtzer, R. (2011). Multisensory integration across the senses in young and old adults. *Brain Res.* 1426, 43–53. doi: 10.1016/j.brainres.2011.09.017

Mahoney, J. R., Verghese, J., Dumas, K., Wang, C., and Holtzer, R. (2012). The effect of multisensory cues on attention in aging. *Brain Res.* 1472, 63–73. doi: 10.1016/j.brainres.2012.07.014

Mather, M. (2012). The emotion paradox in the aging brain. *Ann. N. Y. Acad. Sci.* 1251, 33–49. doi: 10.1111/j.1749-6632.2012.06471.x

Mather, M., Canli, T., English, T., Whitfield, S., Wais, P., Ochsner, K., et al. (2004). Amygdala responses to emotionally valenced stimuli in older and younger adults. *Psychol. Sci.* 15, 259–263. doi: 10.1111/j.0956-7976.2004.00662.x

Mill, A., Allik, J., Realo, A., and Valk, R. (2009). Age-related differences in emotion recognition ability: a cross-sectional study. *Emotion* 9, 619–630. doi: 10.1037/a0016562

Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.* 14, 247–279.

Mitchell, R. L. (2007). Age–related decline in the ability to decode emotional prosody: primary or secondary phenomenon? *Cogn. Emot.* 21, 1435–1454. doi: 10.1080/02699930601133994

Mitchell, R. L., Kingston, R. A., and Barbosa Bouças, S. L. (2011). The specificity of age-related decline in interpretation of emotion cues from prosody. *Psychol. Aging* 26, 406–414. doi: 10.1037/a0021861

Mishra J., and Gazzaley, A. (2013). Preserved discrimination performance and neural processing during crossmodal attention in aging. *PLoS ONE* 8:81894. doi: 10.1371/journal.pone.0081894

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., and Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cogn. Brain Res.* 14, 115–128. doi: 10.1016/S0926-6410(02)00066-6

Mu, Q., Xie, J., Wen, Z., Weng, Y., and Shuyun, Z. (1999). A quantitative MR study of the hippocampal formation, the amygdala, and the temporal horn of the lateral ventricle in healthy subjects 40 to 90 years of age. *Am. J. Neuroradiol.* 20, 207–211.

Murphy, N. A., and Isaacowitz, D. M. (2010). Age effects and gaze patterns in recognising emotional expressions: an in-depth look at gaze measures and covariates. *Cogn. Emot.* 24, 436–452. doi: 10.1080/02699930802664623

Noller, P., and Feeney, J. A. (1994). Relationship satisfaction, attachment, and nonverbal accuracy in early marriage. *J. Nonverbal. Behav.* 18, 199–221. doi: 10.1007/BF02170026

Orbelo, D. M., Grim, M. A., Talbott, R. E., and Ross, E. D. (2005). Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. *J. Geriatr. Psychiatry. Neurol.* 18, 25–32. doi: 10.1177/0891988704272214

Orgeta, V., and Phillips, L. H. (2008). Effects of age and emotional intensity on the recognition of facial emotion. *Exp. Aging Res.* 34, 63–79. doi: 10.1080/03610730701762047

Park, D. C., and Reuter-Lorenz, P. (2009). The adaptive brain: aging and neurocognitive scaffolding. *Annu. Rev. Psychol.* 60, 173–196. doi: 10.1146/annurev.psych.59.103006.093656

Paulmann, S., Pell, M. D., and Kotz, S. A. (2008). Functional contributions of the basal ganglia to emotional prosody: evidence from ERPs. *Brain Res.* 1217, 171–178. doi: 10.1016/j.brainres.2008.04.032

Peiffer, A. M., Mozolic, J. L., Hugenschmidt, C. E., and Laurienti, P. J. (2007). Age-related multisensory enhancement in a simple audiovisual detection task. *Neuroreport* 18, 1077–1081. doi: 10.1097/WNR.0b013e3281e72ae7

Pell, M. D. (2005). Prosody–face interactions in emotional processing as revealed by the facial affect decision task. *J. Nonverbal. Behav.* 29, 193–215. doi: 10.1007/s10919-005-7720-z

Pell, M. D., and Kotz, S. A. (2011). On the time course of vocal emotion recognition. *PLoS ONE* 6:e27256.

Reed, A. E., and Carstensen, L. L. (2012). The theory behind the age-related positivity effect. *Front. Psychol.* 3:339. doi: 10.3389/fpsyg.2012.00339

Resnick, S. M., Lamar, M., and Driscoll, I. (2007). Vulnerability of the orbitofrontal cortex to age-associated structural and functional brain changes. *Ann. N. Y. Acad. Sci.* 1121, 562–575. doi: 10.1196/annals.1401.027

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., and Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: a shrinking brain. *J. Neurosci.* 23, 3295–3301.

Rigoulot, S., Wassiliwizky, E., and Pell, M. D. (2013). Feeling backwards? How temporal order in speech affects the time course of vocal emotion recognition. *Front. Psychol.* 4:367. doi: 10.3389/fpsyg.2013.00367

Ruffman, T., Henry, J. D., Livingstone, V., and Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: implications for neuropsychological models of aging. *Neurosc. Biobehav. Rev.* 32, 863–881. doi: 10.1016/j.neubiorev.2008.01.001

Ruggles, D., Bharadwaj, H., and Shinn-Cunningham, B. G. (2011). Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15516–15521. doi: 10.1073/pnas.1108912108

Ryan, M., Murray, J., and Ruffman, T. (2010). Aging and the perception of emotion: processing vocal expressions alone and with faces. *Exp. Aging. Res.* 36, 1–22. doi: 10.1080/03610730903418372

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiol. Aging* 30, 507–514. doi: 10.1016/j.neurobiolaging.2008.09.023

Scherer, K. R. (1994). "Affect bursts," in *Emotions*, eds S. H. M. van Goozen, N. E. van de Poll, J. A. Sergeant (Hillsdale, NJ: Lawrence Erlbaum), 161–193

Shimokawa, A., Yatomi, N., Anamizu, S., Torii, S., Isono, H., Sugai, Y., et al. (2001). Influence of deteriorating ability of emotional comprehension on interpersonal behavior in Alzheimer-type dementia. *Brain Cogn.* 47, 423–433. doi: 10.1006/brcg.2001.1318

Sims, T., Hogan, C., and Carstensen, L. (2015). Selectivity as an emotion regulation strategy: lessons from older adults. *Curr. Opin. Psychol.* 3, 80–84. doi: 10.1016/j.copsyc.2015.02.012

Slessor, G., Phillips, L. H., and Bull, R. (2010). Age-related changes in the integration of gaze direction and facial expressions of emotion. *Emotion* 10, 555–562. doi: 10.1037/a0019152

Sommers, M. S., Tye–Murray, N., and Spehar, B. (2005). Auditory-visual speech perception and auditory-visual enhancement in normal-hearing younger and older adults. *Ear Hearing* 26, 263–275. doi: 10.1097/00003446-200506000-00003

Stephen, J. M., Knoefel, J. E., Adair, J., Hart, B., and Aine, C. J. (2010). Aging-related changes in auditory and visual integration measured with MEG. *Neurosc. Lett.* 484, 76–80. doi: 10.1016/j.neulet.2010.08.023

Stein, B. E., Burr, D., Constantinidis, C., Laurienti, P. J., Alex Meredith, M., Perrault, T. J., et al. (2010). Semantic confusion regarding the development of multisensory integration: a practical solution. *Eur. J. Neurosci.* 31, 1713–1720. doi: 10.1111/j.1460-9568.2010.07206.x

Stein, B. E., Stanford, T. R., and Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nat. Rev. Neurosci.* 15, 520–535. doi: 10.1038/nrn3742

Stine, E. A. L., Wingfield, A., and Myers, S. D. (1990). Age differences in processing information from television news: the effects of bisensory augmentation. *J. Gerontol.* 45, 1–8. doi: 10.1093/geronj/45.1.P1

Sullivan, S., and Ruffman, T. (2004). Emotion recognition deficits in the elderly. *Int. J. Neurosci.* 114, 403–432. doi: 10.1080/00207450490270901

Suri, G., and Gross, J. J. (2012). Emotion regulation and successful aging. *Trends Cogn. Sci.* 16, 409–410. doi: 10.1016/j.tics.2012.06.007

Szanto, K., Dombrovski, A. Y., Sahakian, B. J., Mulsant, B. H., Houck, P. R., Reynolds, C. F. III., et al. (2012). Social emotion recognition, social functioning, and attempted suicide in late-life depression. *Am. J. Geriat. Psychiat.* 20, 257–265. doi: 10.1097/JGP.0b013e31820eea0c

Steptoe, A., Shankar, A., Demakakos, P., and Wardle, J. (2013). Social isolation, loneliness, and all-cause mortality in older men and women. *Proc. Natl. Acad. Sci. U.A.S.* 110, 5797–5801. doi: 10.1073/pnas.1219686110

Templier, L., Chetouani, M., Plaza, M., Belot, Z., Bocquet, P., and Chaby, L. (2015). Altered identification with relative preservation of emotional prosody production in patients with Alzheimer's disease. *Geriatr. Psychol. Neuropsychiatr. Vieil.* 13, 106–115. doi: 10.1684/pnv.2015.0524

Townsend, J. T., and Honey, C. (2007). Consequences of base time for redundant signals experiments. *J. Math. Psychol.* 51, 242–265. doi: 10.1016/j.jmp.2007.01.001

Tsuchida, A., and Fellows, L. K. (2012). Are you upset? Distinct roles for orbitofrontal and lateral prefrontal cortex in detecting and distinguishing facial expressions of emotion. *Cereb. Cortex* 22, 2904–2912. doi: 10.1093/cercor/bhr370

Ulrich, R., Miller, J., and Schröter, H. (2007). Testing the race model inequality: an algorithm and computer programs. *Behav. Res. Methods* 39, 291–302. doi: 10.3758/BF03193160

Urry, H. L., Van Reekum, C. M., Johnstone, T., Kalin, N. H., Thurow, M. E., Schaefer, H. S., et al. (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. *J. Neurosci.* 26, 4415–4425. doi: 10.1523/JNEUROSCI.3215-05.2006

Verhaeghen, P., and Salthouse, T. A. (1997). Meta-analyses of age–cognition relations in adulthood: estimates of linear and nonlinear age effects and structural models. *Psychol. Bull.* 122, 231–249.

Walden, B. E., Busacco, D. A., and Montgomery, A. A. (1993). Benefit from visual cues in auditory-visual speech recognition by middle-aged and elderly persons. *J. Speech. Lang. Hear. Res.* 36, 431–436. doi: 10.1044/jshr.3602.431

Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., and Belin, P. (2014). Crossmodal adaptation in right posterior superior temporal sulcus during face–voice emotional integration. *J. Neurosci.* 34, 6813–6821. doi: 10.1523/JNEUROSCI.4478-13.2014

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., et al. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage* 24, 1233–1241. doi: 10.1016/j.neuroimage.2004.10.034

Williams, L. M., Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., and Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *J. Clin. Exp. Neuropsychol.* 31, 257–277. doi: 10.1080/13803390802255635

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

Young, A. W., and Bruce, V. (2011). Understanding person perception. *Brit. J. Psychol.* 102, 959–974. doi: 10.1111/j.2044-8295.2011.02045.x